On-Call Simulation Modeling Training

Model Development, Calibration and Review

presented to

Caltrans

presented by

Cambridge Systematics, Inc.

October 1, 2018

Think Forward

Today's Agenda

- Background and objectives
- Model Development Review
- Model Validation Datasets
- Changing FHWA Guidance on Calibration
- Calibration Process Review
- Model Review Process

Project Background and Objectives

- Senate Bill (SB) 375 (Sustainable Community), SB 743 (CEQA Reform), and SB 391 (California Transportation Plan) require a more robust quantitative and analytic evaluation to describe the relative performance of transportation policies, strategies, and programs.

- SB 1, now in force, created the Solutions for Congested Corridors Program. Caltrans will be collaborating with regional partners to identify and develop fixes for these corridors, which cannot be analyzed using static methods alone.

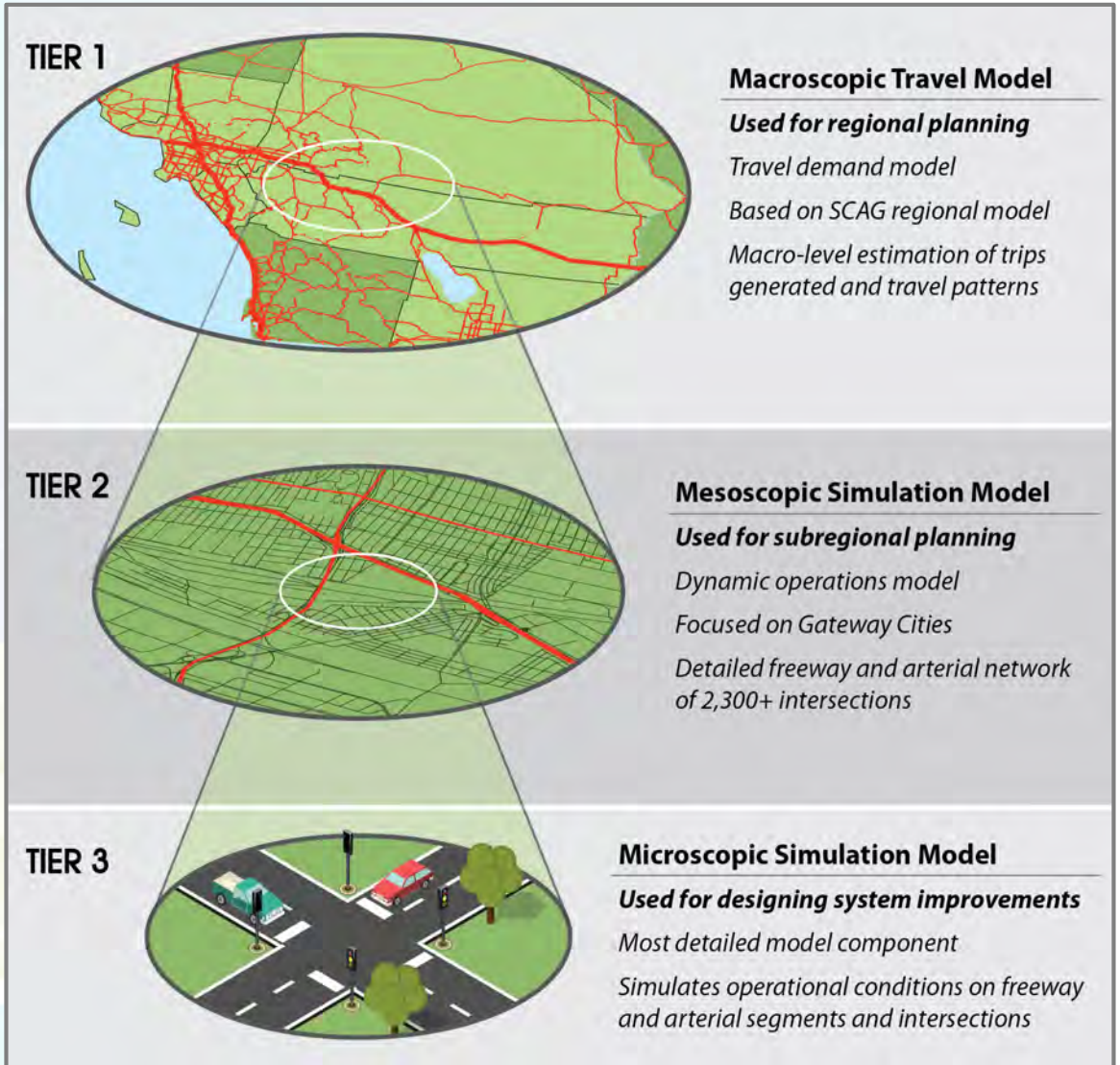
- On-call traffic simulation training will enable Caltrans to meet the mandate of these bills by educating Caltrans staff about how to perform complex analyses of our facilities for critical planning, operations, and capital improvement projects using the latest generation of traffic analysis tools.

Webinar Four-part Series

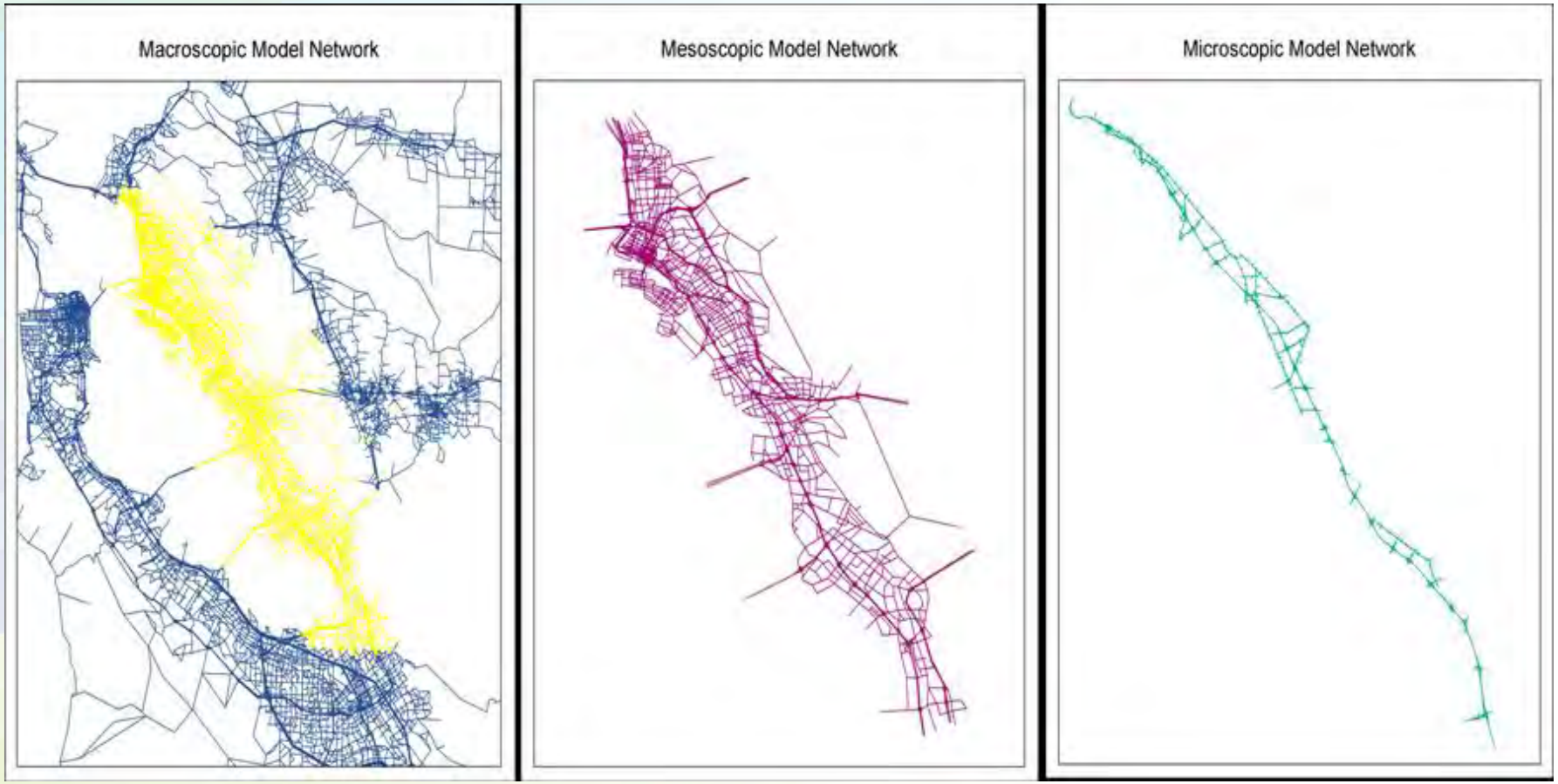
- Webinar 1 - Transportation Analysis and Simulation Overview – Delivered in January 2018
- Webinar 2 – Scoping a Simulation Project – Delivered in June 2018
- Webinar 3 - How to Develop, Calibrate & Review Models – Today
- Webinar 4 - How to Interpret and Communicate Model Results and How to Produce Output for Environmental Studies from Simulation Models – Documentation of calibration assumptions and results; documentation of overall analysis results; with and without project; tables, charts, graphics and maps summaries of performance measures. Key output from simulation models, key needs for environmental studies (volumes, delays, speeds, queueing, travel time, VMT, etc.), how to use and interpret simulation model output for environmental studies.
  Delivery date tentative January 2019.

Macro, Meso and Micro Simulation Modeling

- ***Macro*** - Long range traffic forecasts, regional patterns and mode shift
  - » Usually Demand Modeling
  - » Rarely Simulation
- ***Meso*** - Traveler information, HOT lanes, congestion pricing, regional diversion
- ***Micro*** - Detailed analysis of physical improvements and traffic control strategies, congested conditions



Analysis Resolutions



What is Calibration?

- Calibration is the adjustment of model parameters to improve the model's ability to reproduce time-dynamic system performance observed under specific operational conditions.

What is Validation?

- Process where the analyst checks the overall model-predicted traffic performance for a street/road system against field measurements of traffic performance, such as traffic volumes, travel times, average speeds, and average delays.

MODEL DEVELOPMENT

Traffic Analysis Tools: Volume III Microsimulation

- Microsimulation Model Development Process
  - From forthcoming update version

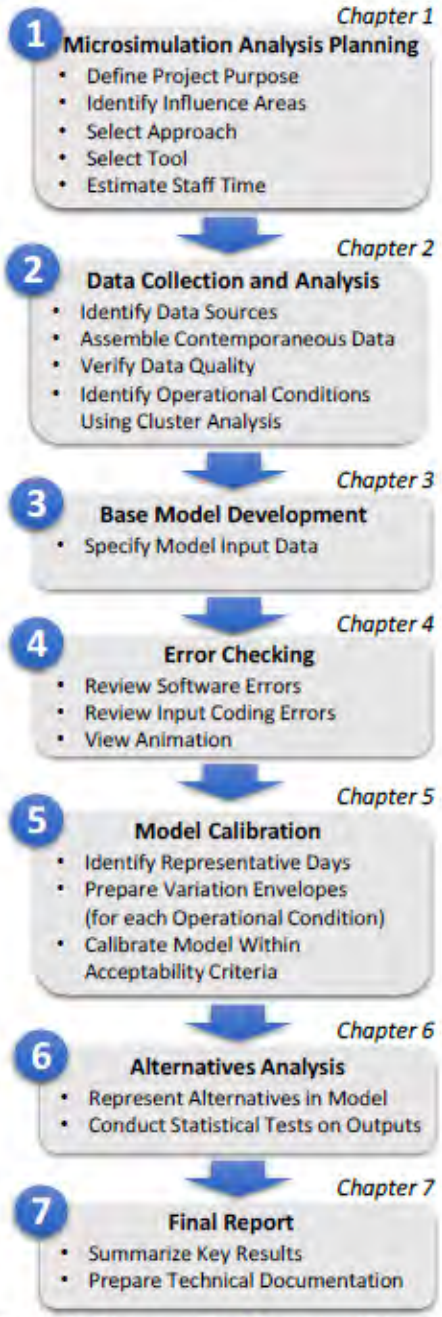


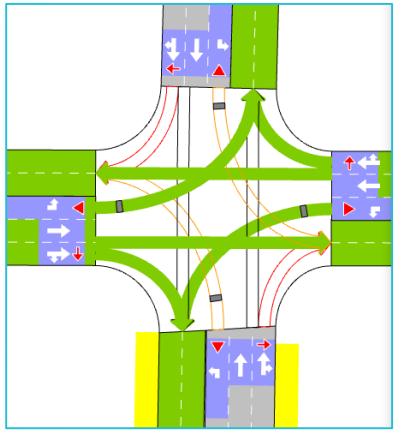
Figure 1. Diagram. The Microsimulation Analytical Process (Source: FHWA)

Review of Basic Network Elements

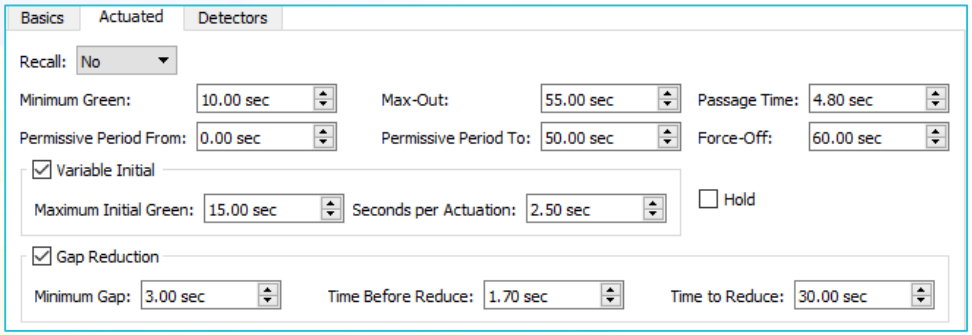
- Geometry
  - Network connectivity
  - Section/Link parameters
  - Turn/Connector parameters
    - Reduced speeds
    - U-turn prohibitions

Review of Basic Network Elements

- Signal Controls
  - Physical movements coded for each phase
  - Proper coding of timing and phase settings
  - Fixed vs Actuation
  - Actuation settings properly coded?
  - Coordination?



- Stop Controls
  - Driver reactions



Review of Basic Network Elements

- Demands
  - Zones / Centroids / Parking Lots
  - Review OD preparation and adjustment process
  - Traffic Demands
    - Static Routes
    - OD Matrices
    - Vehicle Types
    - Scaling Factors
- Warm-up time
- Cool-down time

Review of Basic Network Elements

- ITS Elements
- Existing Traffic Management Strategies
- Transit

Pre-Calibration Error Checking

- Test via diagnostic runs
  - Low volume simulations
- Are vehicle movements realistic?
- Are signals functioning correctly?
- Are vehicles routed logically?

MODEL VALIDATION DATASETS

Traditional Types of Data

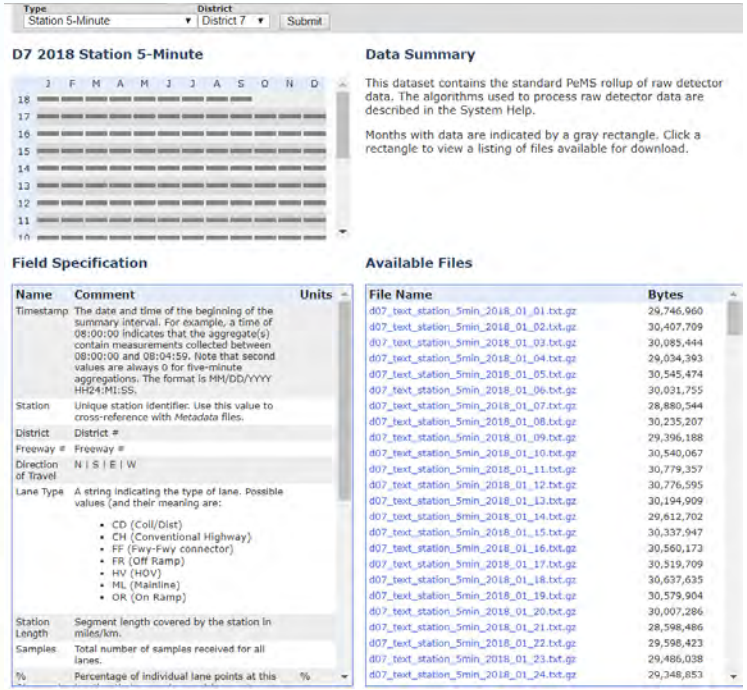
Physical Geometry	Traffic Control	Travel Demand	Operational Data
• Rectified aerial photography • Base GIS mapping files	• Signs • Signal control timing plans • Ramp Meter Controls	• Traffic counts • Vehicle classification counts • Transit Schedules	• Speeds • Travel times • Congestion • Queuing measurements • Other field observations

Using PeMS Data – Downloading Data

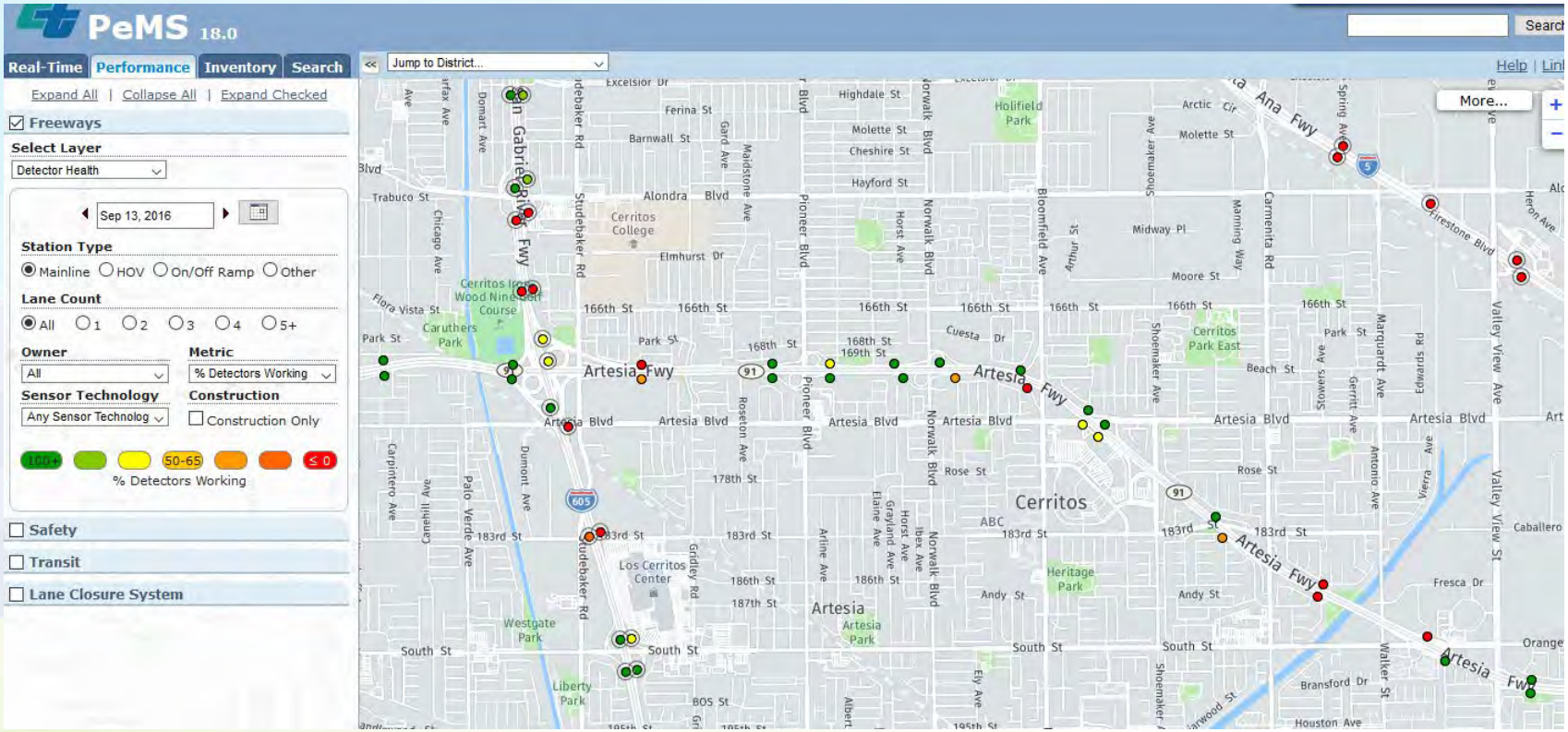
- PeMS Data Clearinghouse provides a single access point for downloading PeMS data sets.

- Flow and Speed
  - 5-minute data
  - AM, PM peak period or 24 hours
  - Week or Weekday (Tuesday, Wednesday, Thursday), no holidays
  - Month or Year

- Health of Detectors
  - Use flow and speed, only if it observed data is good.



PeMS Detector Health

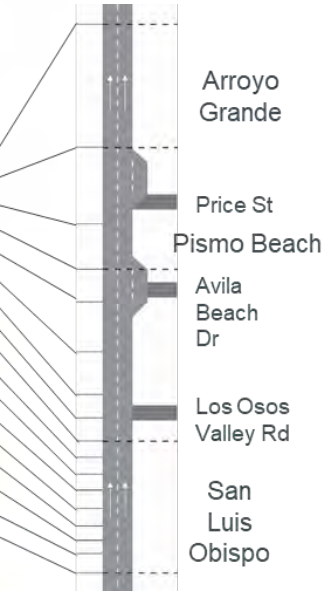


http://pems.dot.ca.gov/?dnode=search&content=cnt_search&view=p#33.87487,-118.08187,15

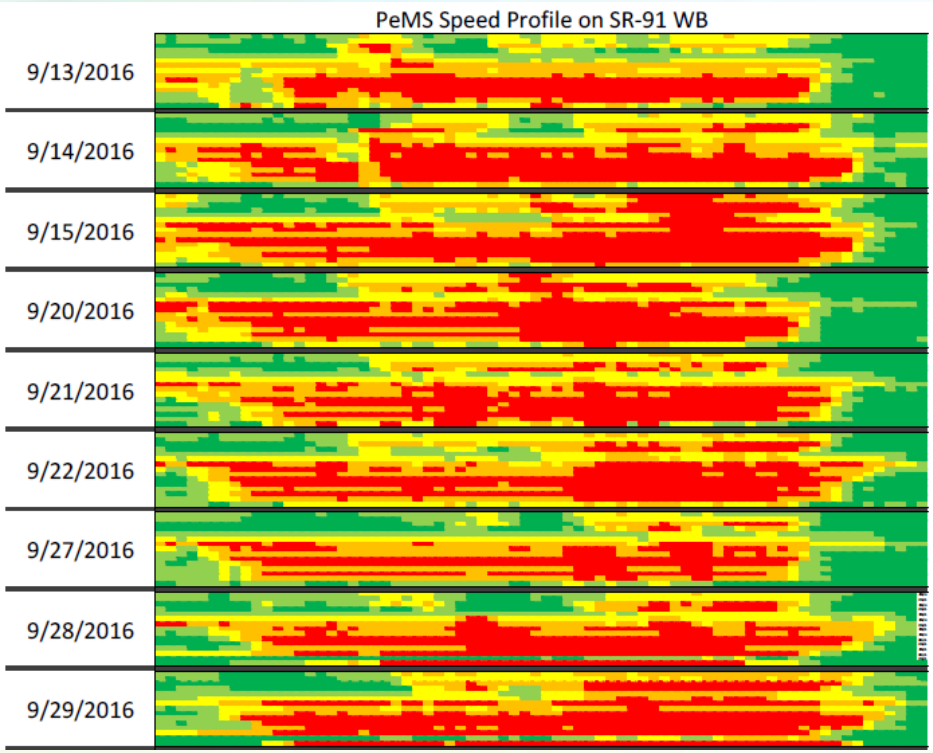
Using PeMS Data – Analyzing Data

- Once data is selected for the region and time of interest, heat maps can be developed to visualize speeds, congestion, flows, and quality of data for a specific time period.

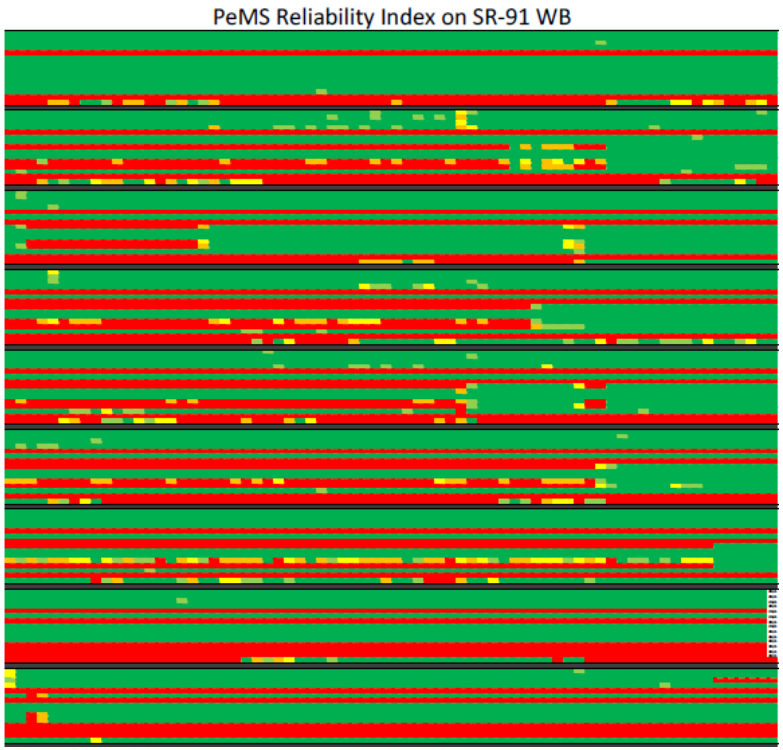
	Hour																							
Name	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
BROADWAY HWY 135 101 SB VDS MLSB	67	67	67	67	67	67	67	66	66	65	65	65	65	65	64	63	63	62	64	67	67	67	67	67
OAK PARK BLVD 101 NB VDS MLSB SB	66	66	66	66	66	66	65	63	62	63	63	63	63	63	63	63	60	57	61	62	64	65	66	66
SHELL BEACH RD 101 NB VDS MLSB S	65	65	65	65	65	65	65	65	64	64	64	63	63	63	62	59	56	56	61	64	64	65	65	65
SPYGLASS DR 101 SB VDS MLSB SB	69	69	69	69	69	69	68	67	67	66	65	65	64	64	61	51	38	35	55	66	67	67	68	68
AVILA BEACH DR 101 NB VDS MLSB S	70	70	70	69	70	70	70	70	70	69	69	67	67	67	65	59	44	38	60	69	69	69	70	70
SAN LUIS BAY DR 101 SB VDS MLSB	69	69	69	69	69	69	68	67	67	66	65	65	64	64	63	61	59	58	64	67	67	68	68	69
SO HIGUERA ST EXIT 101 NB VDS ML	67	67	67	67	67	67	67	67	66	65	64	64	63	63	62	58	54	55	64	66	66	67	67	67
SO HIGUERA ST ON RAMP ST 101 NB	67	67	66	66	67	67	67	67	67	66	65	65	64	64	63	59	53	56	64	67	66	66	67	67
LOS OSOS VALLEY RD 101 SB VDS ML	69	69	69	69	69	68	67	67	67	67	66	66	66	65	64	61	55	60	65	67	67	67	68	68
PRADO RD CMS 101 NB VDS MLSB SB	67	66	66	66	66	67	67	66	65	65	65	64	63	62	61	59	55	58	65	67	67	67	67	67
PRADO RD 101 NB VDS MLSB SB	67	66	66	66	66	67	67	67	66	66	66	65	64	63	63	59	55	57	66	67	67	67	67	67
MADONNA RD 101 SB VDS MLSB SB	69	69	69	69	69	68	67	67	66	66	66	65	64	64	63	61	59	61	66	67	67	67	68	68
MARSH ST 101 SB VDS MLSB SB	69	69	69	69	69	68	67	66	65	65	64	64	63	63	62	61	61	61	65	67	67	68	68	68
BROAD ST 101 SB VDS MLSB SB	69	70	70	70	69	69	68	66	65	64	64	63	63	63	62	60	60	60	64	66	67	68	69	69
TORO ST 101 NB VDS MLSB SB	66	66	66	66	67	67	67	65	65	66	66	65	65	65	65	63	63	63	66	67	67	67	67	67
GRAND AVE IN SLO AT 101 SB VDS M	69	69	68	69	68	69	67	65	65	66	67	67	66	66	65	65	66	66	67	68	68	68	68	69
MONTEREY ST 101 NB VDS MLSB SB	66	65	65	65	66	66	66	65	65	66	66	66	66	66	65	65	66	66	67	67	66	66	66	66
FOX HOLLOW RD 101 NB VDS MLSB SB	67	66	66	66	67	67	66	65	65	65	65	65	65	64	64	64	64	64	66	67	67	67	67	67



Speeds and Reliability Index Comparison



WB SR-91 corridor speed profile during PM peak period



Detector health along WB SR-91 corridor during PM peak hour

Emerging Types of Data – Connected Devices and Location Based Services

Operational Data	OD Data	Route Choices
• Gov't & 3rd Party Databases	• Infrequent pings • Where do devices rest?	• Frequent pings • What routes do devices travel?

Transportation Databases - Federal

Database	Description	Extents
FHWA Highway Performance Monitoring System (HPMS)	Data on the extent, condition, performance, use and operating characteristics of the nation's highways. Also captures characteristics of some arterial and collector functional systems.	Nationwide
FHWA National Bridge Inventory (NBI)	Condition data on more than 600,000 bridges located on Interstate Highways, U.S. highways, State and county roads, and publicly-accessible bridges on Federal lands.	Nationwide
FHWA National Household Travel Survey (NHTS) Add-On	Supplementary survey data purchased by State DOTs, MPOs, and COGs for their local areas.	Survey Partners (also known as Add-Ons) exist nationwide
U.S. Department of Homeland Security Weigh-in-Motion (WIM) Stations	WIM devices record axle weights and gross vehicle weights as vehicles drive over a measurement site.	Nationwide
National Performance Management Research Data Set (NPMRDS)	Historical traffic information using anonymous, real-time aggregated GPS probe data from a wide array of commercial vehicle fleets, connected cars and mobile apps.	Nationwide (for National Highway System)

Transportation Databases State and Regional

Database	Description	Extents
Caltrans Performance Measurement System (PeMS)	Real-time and historical traffic data collected from nearly 40,000 individual freeway detectors.	All major metropolitan areas in California
Location-Based Services Data	Set of mobile phone location based services data used to glean insights into linked trips and tours, robust demographics, and travel purpose.	Southern California
Arterial Performance Measurement Tool (APMT)	Establishes baseline performance conditions for selected subregional arterial corridors, such as travel demand, productivity, mobility and reliability.	Specific to Los Angeles County
Caltrans Automated Pavement Condition Survey (APCS)	Condition data collected at highway speeds using specialized vehicles with inertial profilers, transverse laser system, and high resolution cameras for all lanes of the entire State Highway for compliance reporting for MAP-21.	Within the State of California

Transportation Databases State and Regional

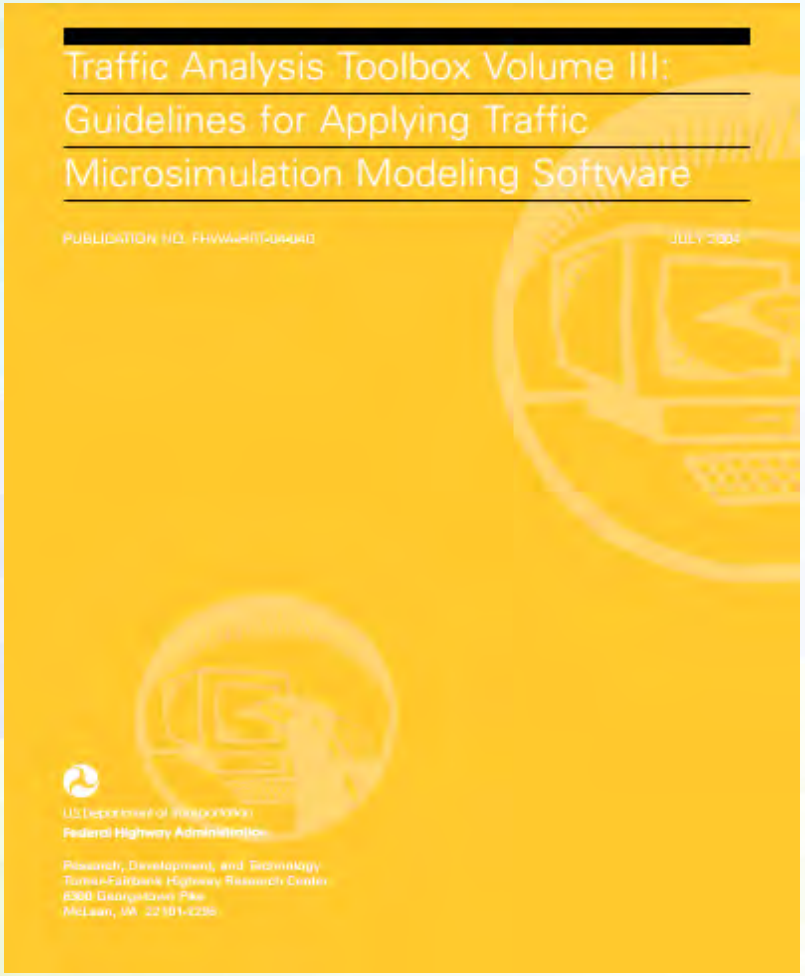
Database	Description	Extents
Statewide Integrated Traffic Records System (SWITRS)	Data gathered from collision scenes by California Highway Patrol staff and members of its Allied Agencies.	Within the State of California
California Vehicle Inventory and Use Survey (CA-VIUS)	State-level vehicle inventory survey that has collected information about commercial vehicle operations from establishments that operate trucks on California's roadways.	Mostly geared towards trucks that operate in California
Caltrans Traffic Counts	Individual Caltrans Districts have calculated the volumes hourly, daily, and monthly to derive an annual average daily traffic count.	Within the State of California
Truck Activity Monitoring System (TAMS)	Uses inductive loop signature technology to obtain high resolution truck data at Inductive Loop Detector sites and Weigh-in-Motion sites.	Various locations across California

Transportation Databases – Private Sector & Other

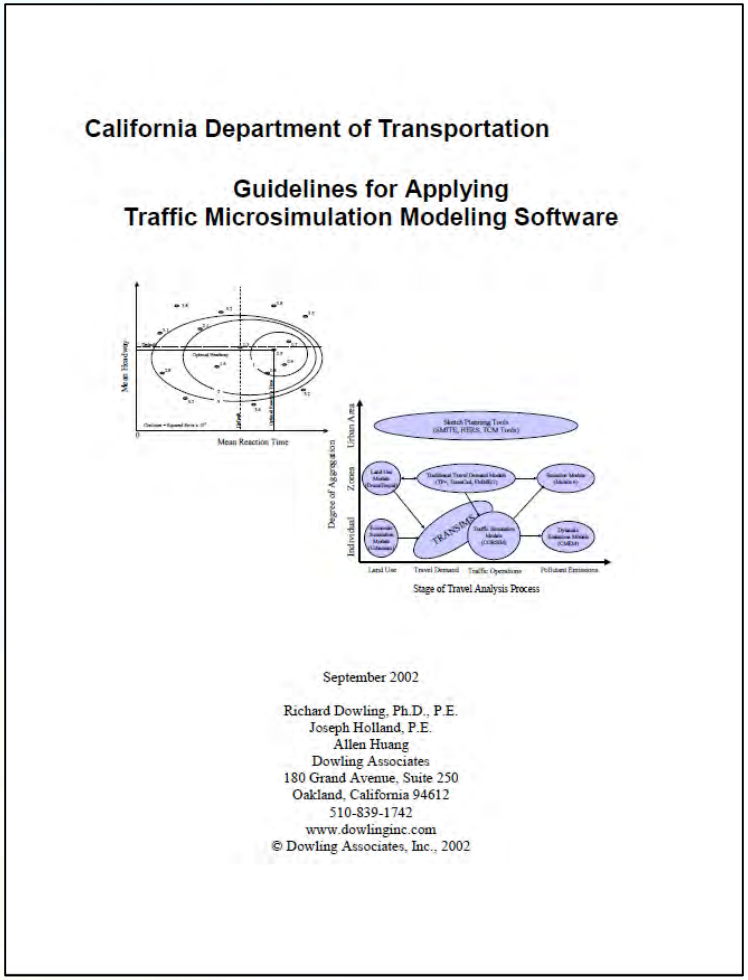
Database	Description	Extents
INRIX, HERE, Waze	Real-time, historical and predictive traffic information using anonymous, real-time aggregated GPS probe data from a wide array of commercial vehicle fleets, connected cars and mobile apps.	Nationwide
Streetlight Data	Collection of anonymized location records created by mobile phones, GPS devices, connected cars, commercial trucks, fitness trackers, etc.	Nationwide
Airsage	Collection of real-time mobile signals, GPS and other location data to produce and process billions of anonymous data points every day.	Nationwide
American Transportation Research Institute (ATRI)	Real-time anonymized freight truck GPS data (e.g., periodic time, location, speed) sourced through unique industry partnerships.	Nationwide

CHANGING FHWA GUIDANCE ON CALIBRATION

Current References



JULY 2004



September 2002

Calibration Criteria: Current Microsim State of the Practice

- Traffic Analysis Tools Volume III (2004)
  - "FHWA criteria"
  - (Actually Wisconsin)

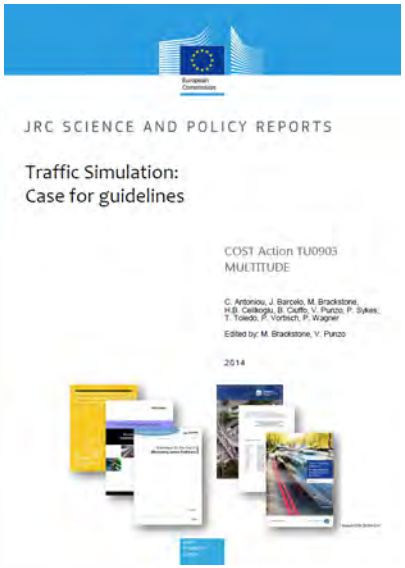
Criteria and Measures	Calibration Acceptance Targets
Hourly Flows, Model Versus Observed	
Individual Link Flows	
Within 15%, for 700 veh/h < Flow < 2700 veh/h	> 85% of cases
Within 100 veh/h, for Flow < 700 veh/h	> 85% of cases
Within 400 veh/h, for Flow > 2700 veh/h	> 85% of cases
Sum of All Link Flows	Within 5% of sum of all link counts
GEH Statistic < 5 for Individual Link Flows*	> 85% of cases
GEH Statistic for Sum of All Link Flows	GEH < 4 for sum of all link counts
Travel Times, Model Versus Observed	
Journey Times, Network	
Within 15% (or 1 min, if higher)	> 85% of cases
Visual Audits	
Individual Link Speeds	
Visually Acceptable Speed-Flow Relationship	To analyst's satisfaction
Bottlenecks	
Visually Acceptable Queuing	To analyst's satisfaction

Other State References

- Other State References
  - Oregon DOT's Protocol for VISSIM Simulation
  - Washington DOT's Protocol for VISSIM Simulation
  - Maryland DOT's VISSIM Modeling Guide
  - Virginia DOT's Traffic Operations and Safety Analysis Manual & VISSIM User Guide,
  - Wisconsin DOT's Traffic Engineering, Operations & Safety Manual - Chapter 16 Section 20 Microscopic Simulation Traffic Analysis
  - Florida DOT's Traffic Analysis Handbook
  - Iowa DOT's Microsimulation Guidance
  - Arkansas DOT's Guidelines for the Use of Microsimulation on AHTD Projects
  - Delaware Valley Regional Planning Commission's VISSIM Standards Project
  - Louisiana DOT's Microsimulation Modeling Policy

Other Valuable References

- International References
  - MULTITUDE (2014)
- Other FHWA Traffic Analysis Toolbox Reports: https://ops.fhwa.dot.gov/trafficanalysistools/index.htm
  - Volume XIII: ICM AMS Guide
  - Volume XIV: Guidebook on DTA Modeling
  - And others



Key Concerns with Previous FHWA Guidance

- Fully Integrate Time-Dynamic Representation of Congestion
- Require Better Representation of Recurrent and Non-Recurrent Conditions
- Remove Subjective Calibration Criteria
- Emphasize Accurate Bottleneck Modeling

Forthcoming Volume III Update

- New Volume III Update Report is completed, but not yet published
- Workshops introducing this new methodology are being conducted by FHWA
- Publication date still TBD

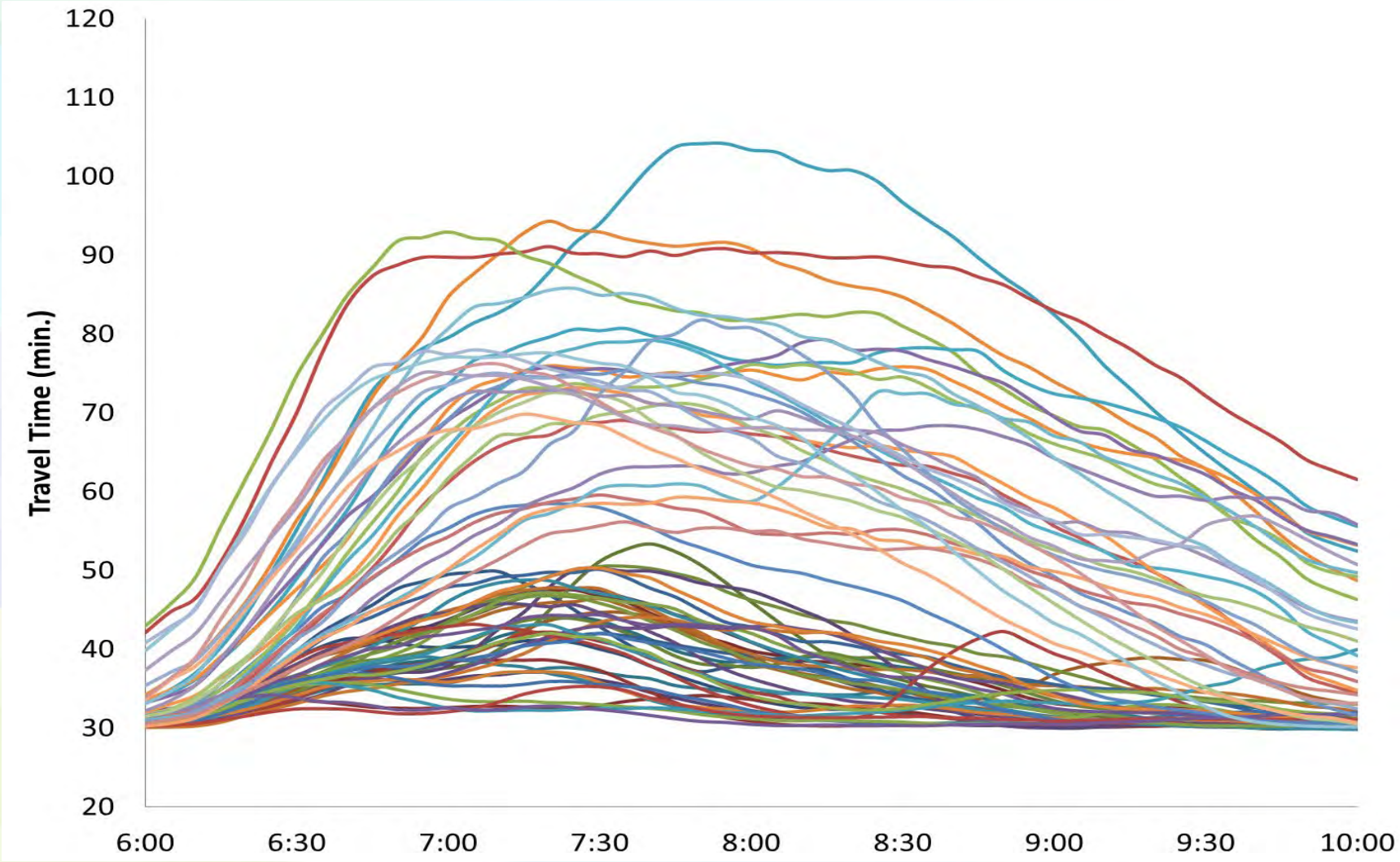
New FHWA Guidance

- Select Performance Measures
  - Calibrate selectively, only for key performance measures.
  - Performance measures for calibration must have good observed data.
- Calibration Process (For each travel condition)
  1. Identification of representative days.
  2. Preparation of variation envelopes.
  3. Assessment of simulation results against four calibration criteria:
     - Control for Maximum Number of Outliers
     - Control for Minimum Number of Inliers
     - Bounded Average Error
     - Bounded Systematic Error

Source: FHWA Volume III Update Workshop

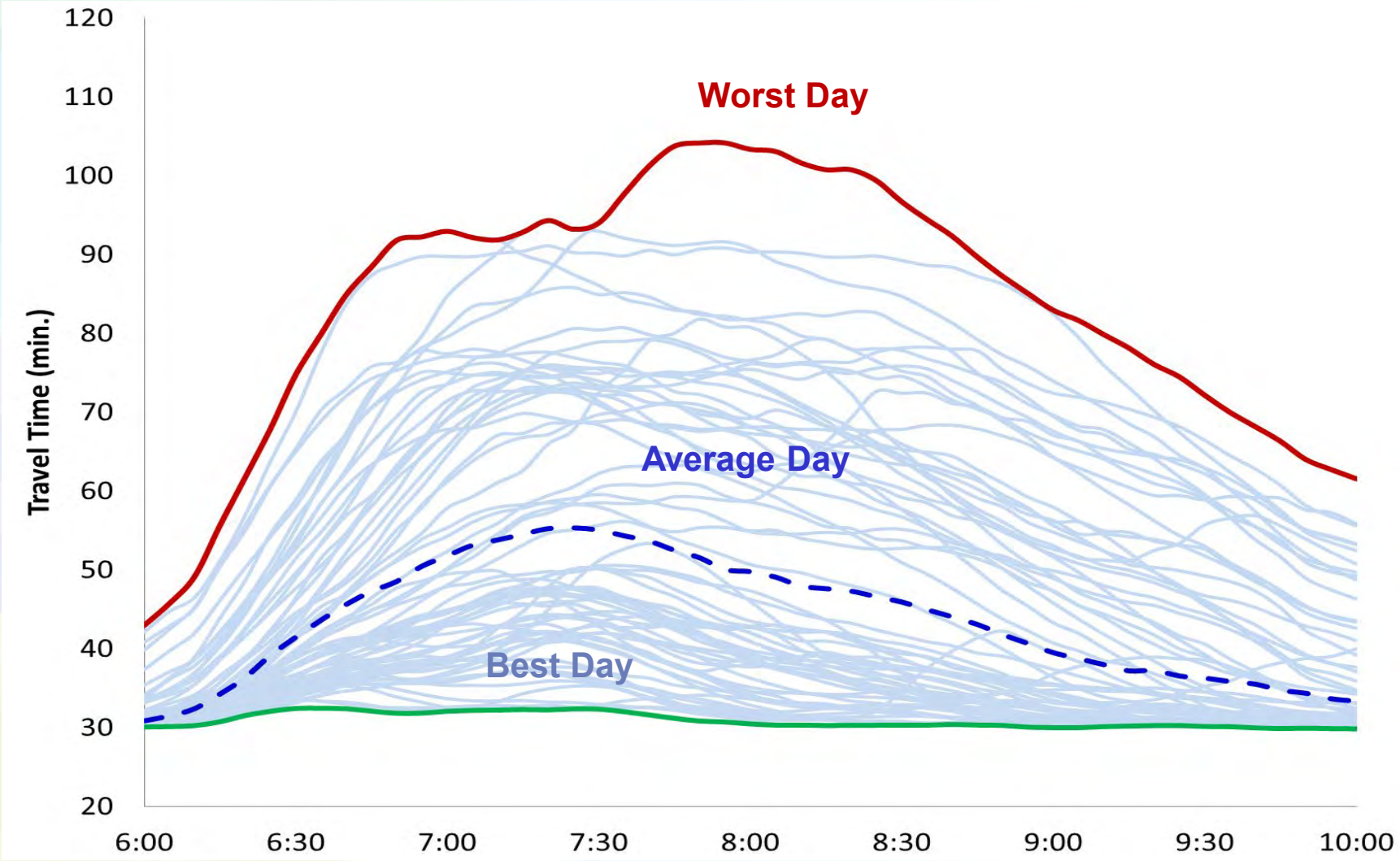
Real Systems Have Good Days & Bad Days

2012 Southbound AM Peak Travel Times, I-405 Corridor (WA)



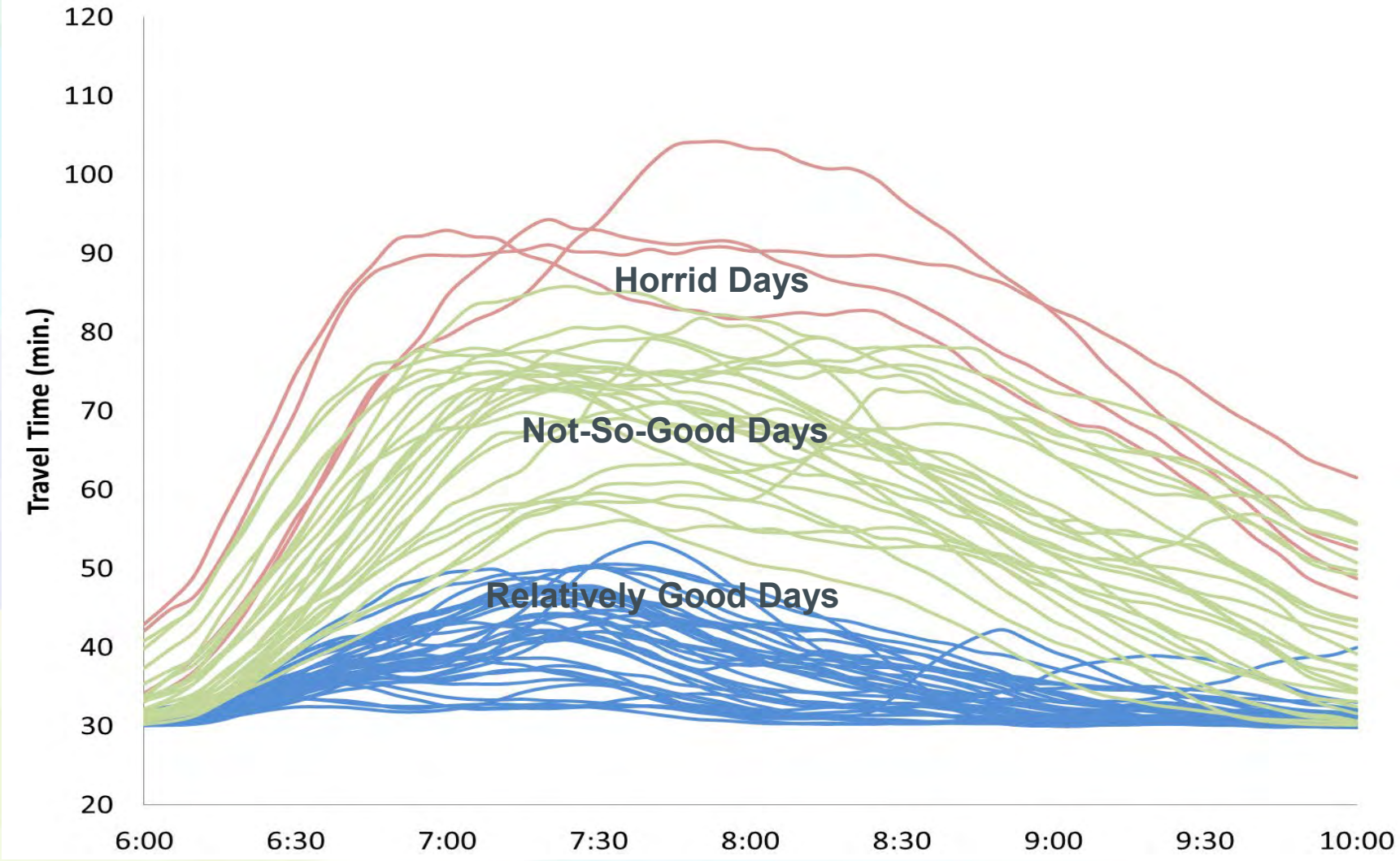
Source: FHWA & Noblis "TAT Volume III Guidelines for Microsimulation" presentation

Even An Average Day Captures Only a Fraction of System Dynamics



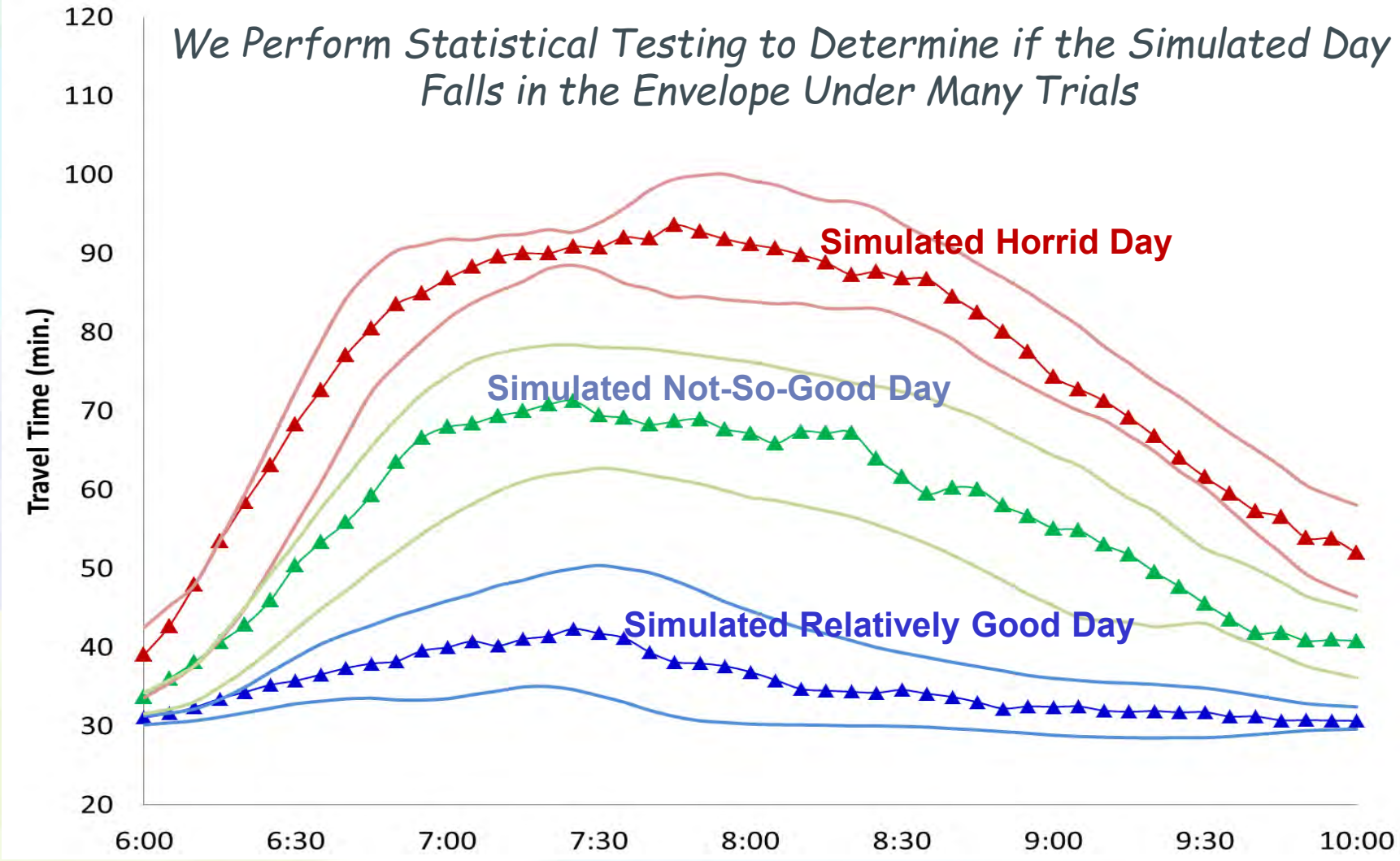
Source: FHWA & Noblis "TAT Volume III Guidelines for Microsimulation" presentation

Use Cluster Analysis to Identify Distinct, Dissimilar Operational Conditions



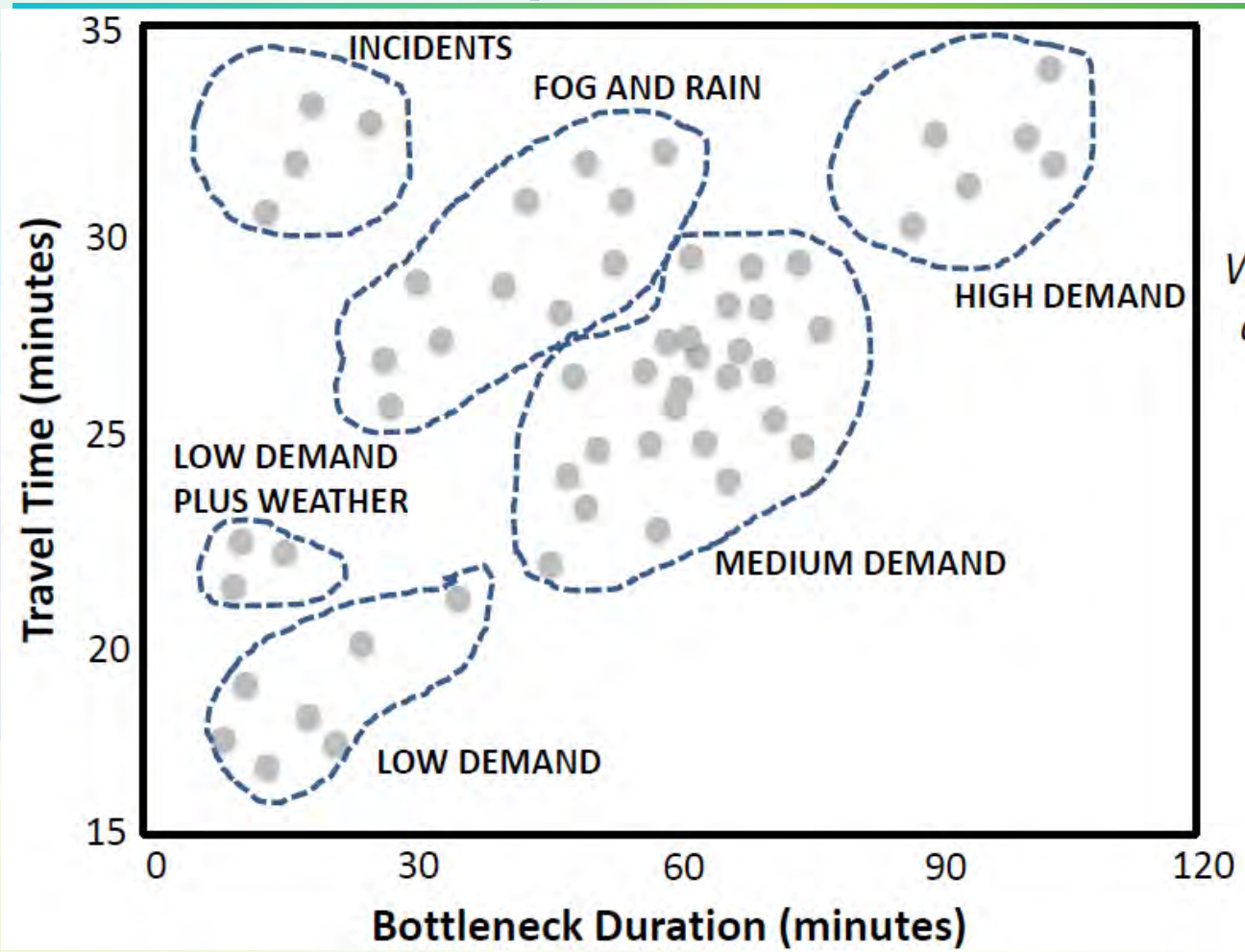
Source: FHWA & Noblis "TAT Volume III Guidelines for Microsimulation" presentation

Simulations Are Calibrated to Lie Within the Statistical Envelope



Source: FHWA & Noblis "TAT Volume III Guidelines for Microsimulation" presentation

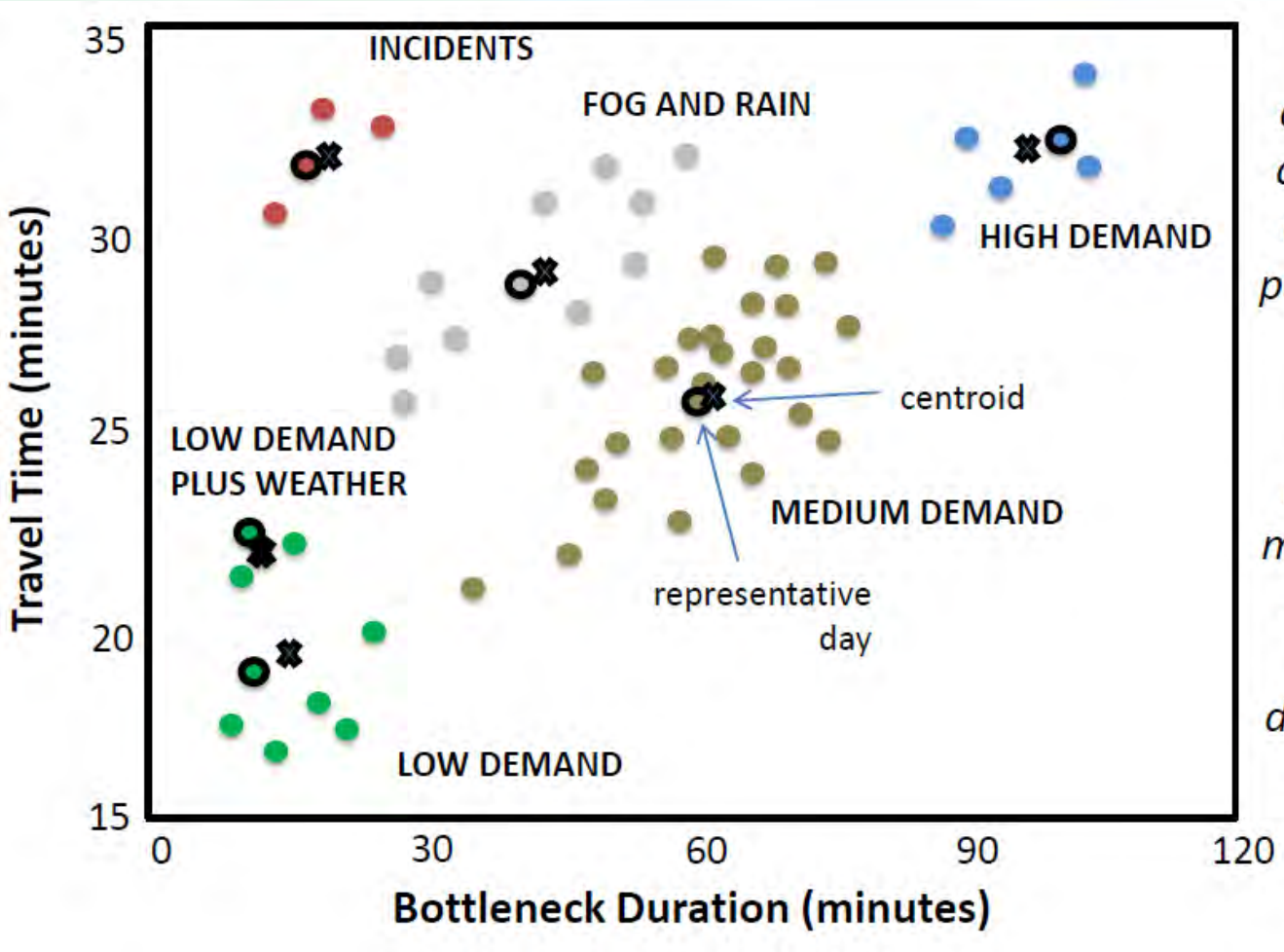
Develop Clusters of Conditions



We group observed data using cluster analysis to identify distinct travel conditions.

Source: FHWA Volume III Update Workshop

Select Representative Days



We calculate a centroid for each cluster that is the arithmetic mean performance in the cluster.

Then, we use relatively simple methods to find an actual day (i.e., representative day) that is closest to the centroid.

Source: FHWA Volume III Update Workshop

Example: San Diego ICM AMS Scenarios

Summary of Best Matching Incident Results

#	Baseline Cluster by Direction and Time Period	Information from Baseline Cluster Analysis			Baseline Period	Post-Deployment Period			
		Days in Cluster	Total Cluster Day Impact (min.)	% of Total Analysis Time Period	Date	Date	DSS Event ID	DSS Plan Type Implemented	DSS Response ID
1	NB PM 4	25	63.25	24.0	10/12/12	7/7/14	639956	Ramps, Signals, ATIS	19536
2	SB AM 2	39	108.03	37.5	10/2/12	2/9/15	754666	Signals, ATIS	27929
3	NB PM 5	3	18.75	2.9	11/21/12	2/19/15	760369	Signals, ATIS	28292
4	SB AM 3	8	34.64	7.7	10/1/12	5/7/15	804238	Ramps, Signals, ATIS	30028
5	n/a, hypothetical	-	-	-	-	5/26/15		None. Managed lanes opened.	
6	SB AM 1	29	49.88	27.9	1/30/13	5/27/15	817649	Signals	30332
7	NB PM 2	8	23.36	7.7	1/15/13	6/9/15	842085	Ramps, Signals	30451
8	NB PM 1	17	41.82	16.3	1/28/13	6/16/15	845922	Ramps, Signals, ATIS	30617
9	NB PM 3b	36	99.72	34.6	1/30/13	5/5/14	853963	Ramps, Signals, ATIS	31039

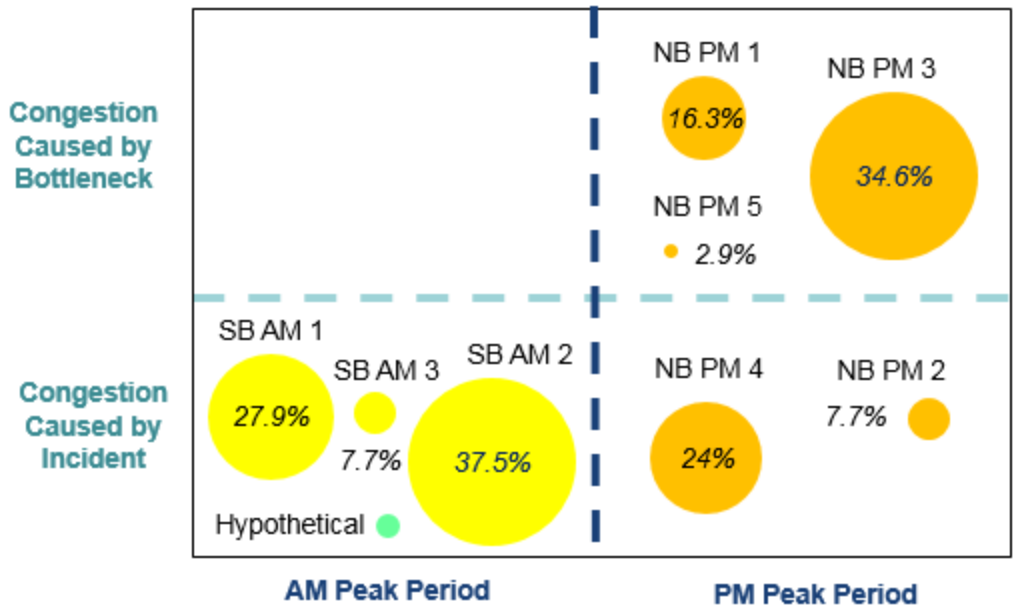
Experimental Design for Analysis of Different Operational Conditions

Highest Priority
- Scenarios with greatest frequency and impact

Lower Priority
- Scenarios with low likelihood but major impact
- Scenarios with frequent occurrence but limited impact

Lowest Priority
- Scenarios with low frequency and low impacts

Operational Condition Dartboard
Scenario Frequency



Note: The size of each circle represents the percent of total analysis time period.

CALIBRATION PROCESS REVIEW

Overall Calibration Process

- Establish calibration objectives.
- Identify the performance measures and critical locations against which the models will be calibrated.
- Determine the statistical methodology to be used to compare modeled results to the field data.
- Determine the strategy for model calibration and identify parameters that are the focus of adjustments.
- Assemble field data previously collected for comparison to model outputs
- Conduct model calibration runs following the strategy and conduct statistical checks.
- Validation: Test or compare the calibrated model with a data set not used for calibration

Model Calibration Methods

- Traffic Network Calibration
- Route Choice Calibration
- Demand Calibration

- Ideally in that order
- Reality is an iterative process

Network Calibration Parameters

- Network Representation
  - » Accurately represented
    - – Posted speeds, lane channelization, grades, visibility, speeds on curves, etc.
- Driver Behavior
  - » Car following parameters, lane changing parameters, desired speed distributions, etc.
- Controls
  - » Driver reaction to signals / stop controls
  - » Gap acceptance model parameters

Route Choice Parameters

- Cost functions
- User costs
- Vehicle type restrictions
- DTA use and settings
- Value of time / Willingness to pay

Demand Calibration

- Initial OD source / estimation
- OD Matrix Estimation (ODME) process (if used)
  - Static
  - Dynamic
- Trip Length Frequencies
- Temporal profiles

Calibration Parameter Checks

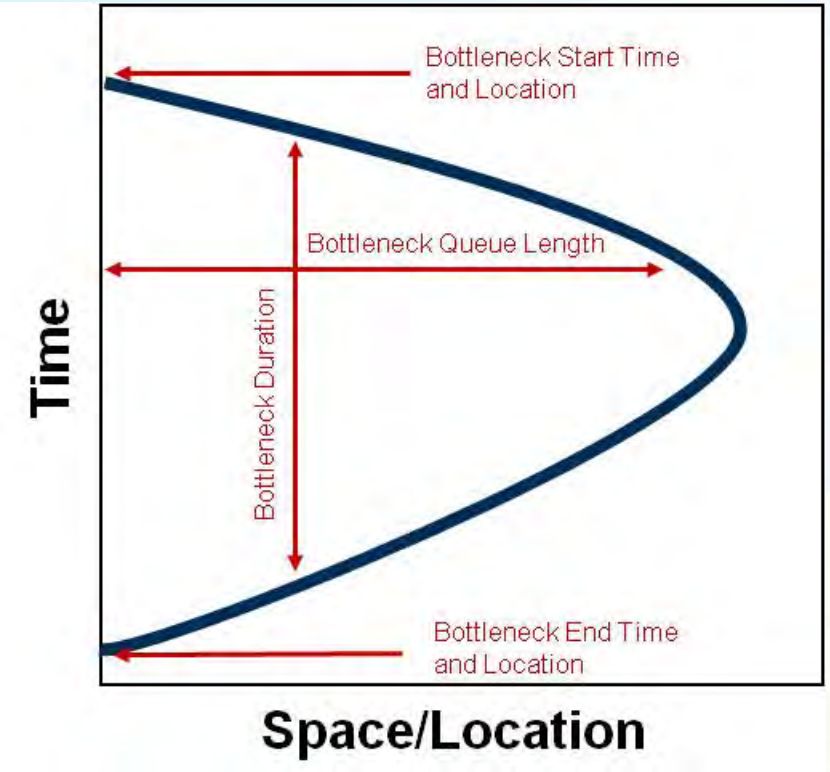
- Use of heavy local adjustments is not good calibration
  - Some are evitable
- Global and/or roadtype adjustment are preferred
  - Or rules based application of parameter changes
- Adjusting too many parameters is not good practice
  - Adjust key parameters the modeler has confidence in
- Avoid overfitting of model to field data
- Are enough random seed runs being performed?
  - Statistical significance of results

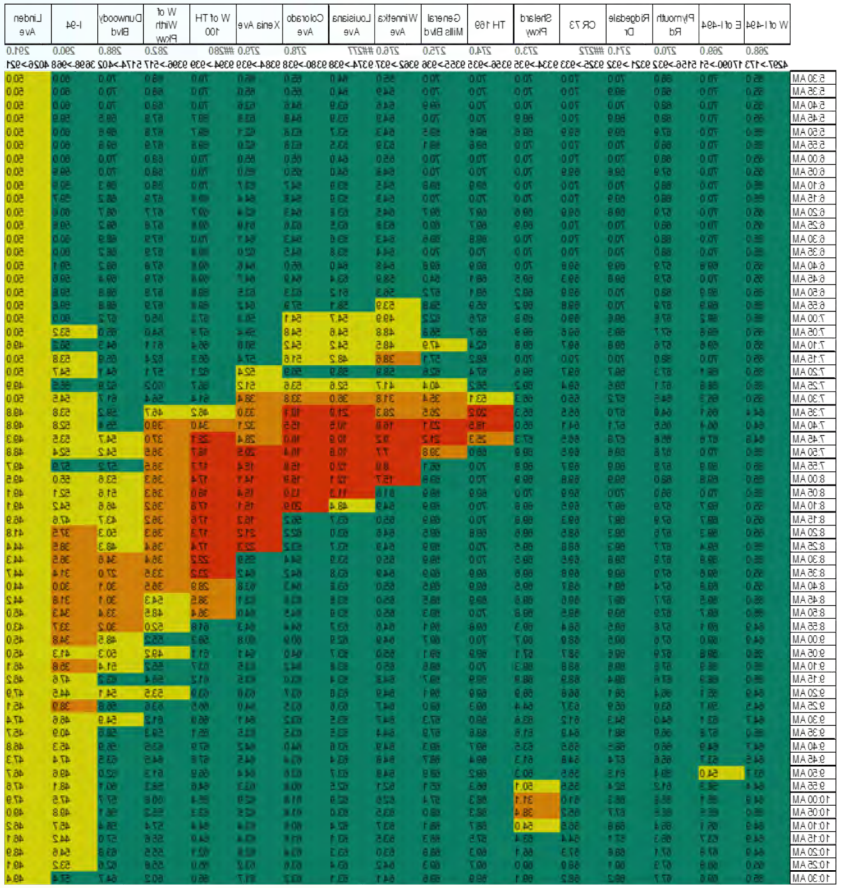
MODEL REVIEW PROCESS

Review of Calibration Results

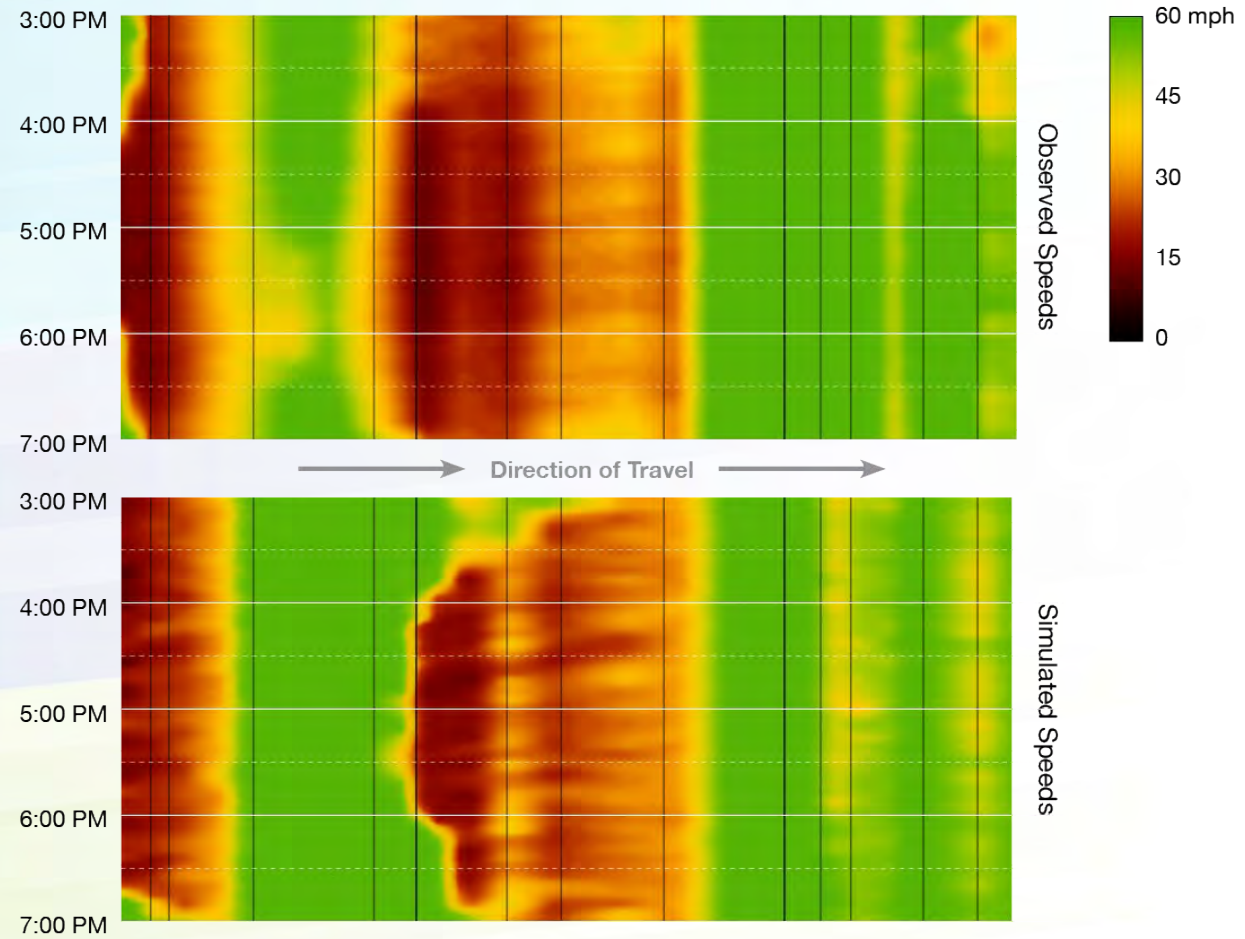
- Simulation Runs
  - Replicate the existing calibration results
  - Review / identify problematic areas
    - Areas of poor fit
    - Low speeds
    - High densities
    - Low flow rate
    - Prolonged queuing
    - Unintuitive routes

Example Speed Contour Diagram



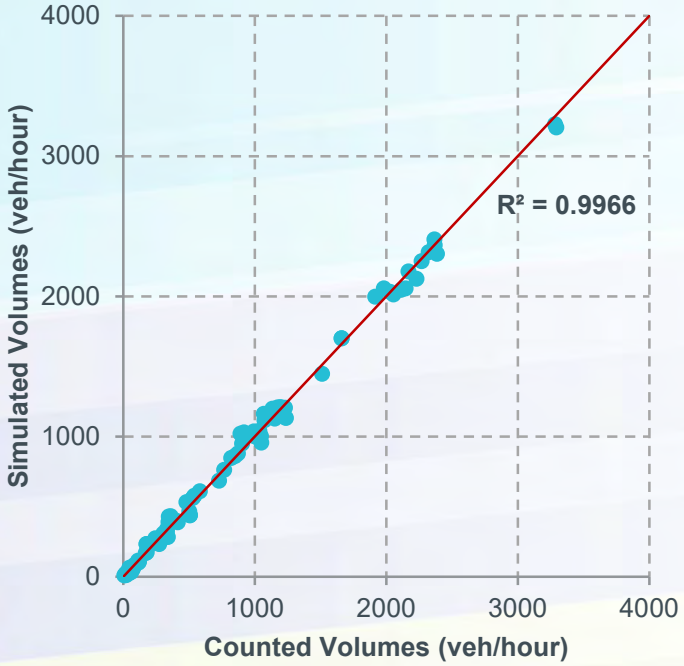


Speed & Bottleneck Validation



Volume Validation Scatterplot: Local Micro Example

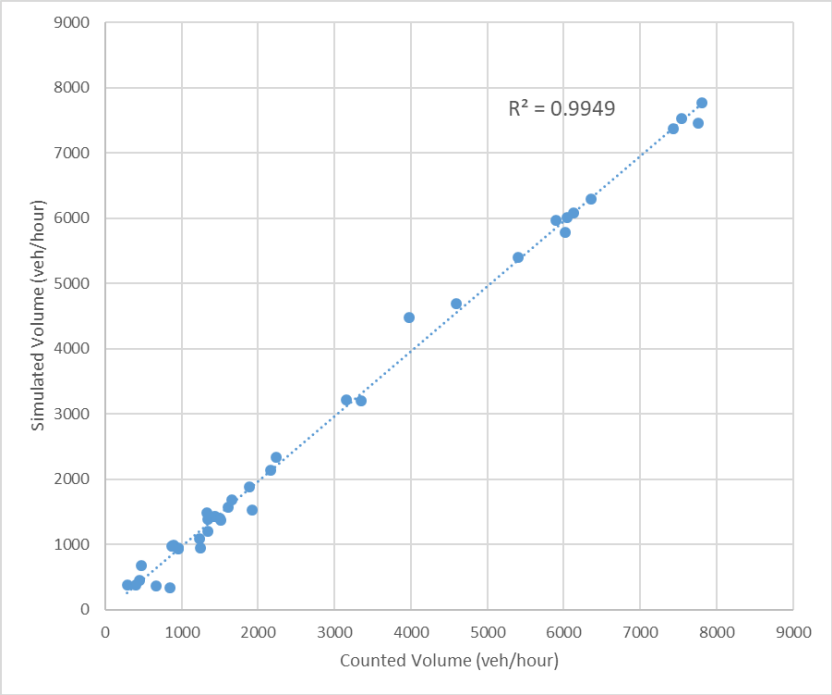
Link Flows

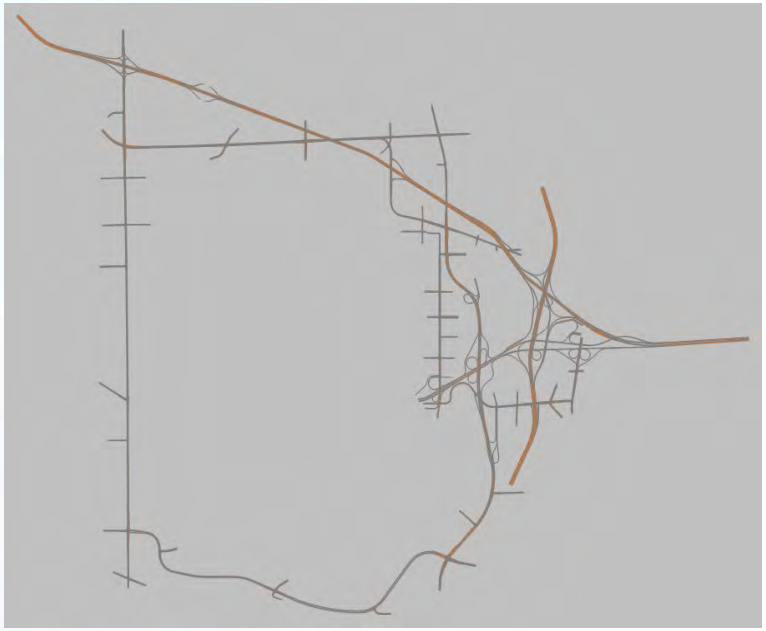


Criteria	Counts Satisfying Criteria	Percentage Satisfying Criteria	Target
GEH < 2	67	77.9%	n/a
GEH < 5	86	100%	> 85%
GEH < 10	86	100%	n/a

Criteria	Counts Satisfying Criteria	Percentage Satisfying Criteria	Target
Within 15%, for 700 veh/h < Flow < 2700 veh/h	41	100%	> 85%
Within 100 veh/h, for Flow < 700 veh/h	43	100%	> 85%
Within 400 veh/h, for Flow > 2700 veh/h	2	100%	> 85%

Volume Validation Scatterplot: Subregional Micro Example



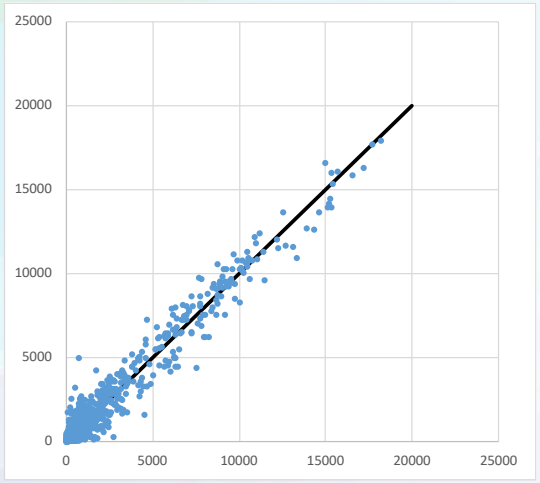


	Counts Meeting Criteria	Percentage of Counts
GEH < 2	22	55%
GEH < 5	**34**	**85%**
GEH < 8	35	88%
GEH < 10	38	95%

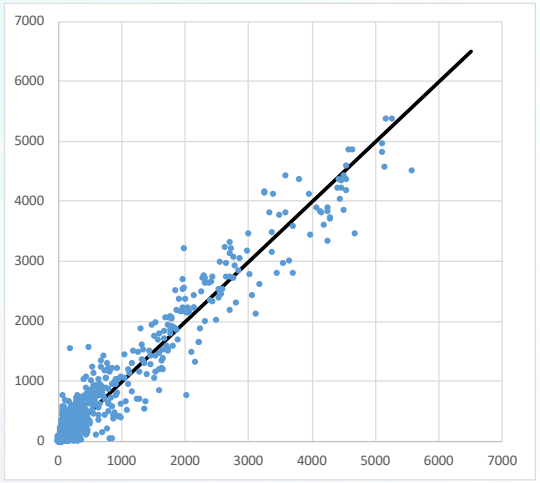
Criteria and Measures	Counts Meeting Criteria	Percentage of Counts
>85% within 100 veh/h, for Flow < 700 veh/h (>85%)	3	60%
>85% within 15%, for 700 veh/h < Flow < 2700 veh/h	18	86%
>85% within 400 veh/h, for Flow > 2700 veh/h	13	93%

Volume Validation Scatterplot: Regional Meso Example

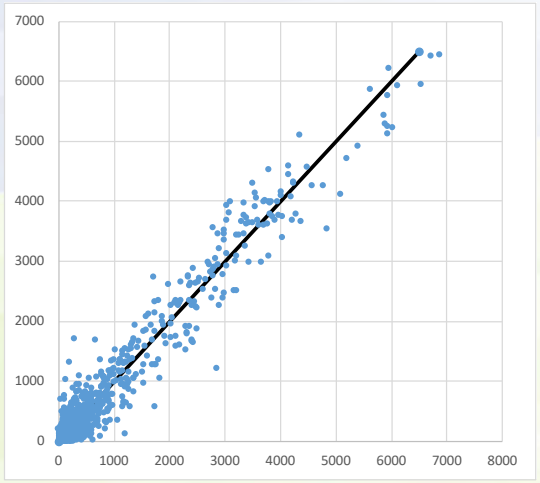
AM6-9 Period Flow - Entire DTA



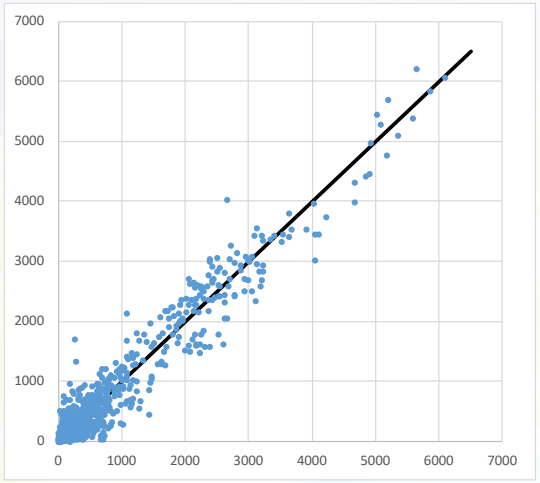
AM6-7 Hour Flow - Entire DTA

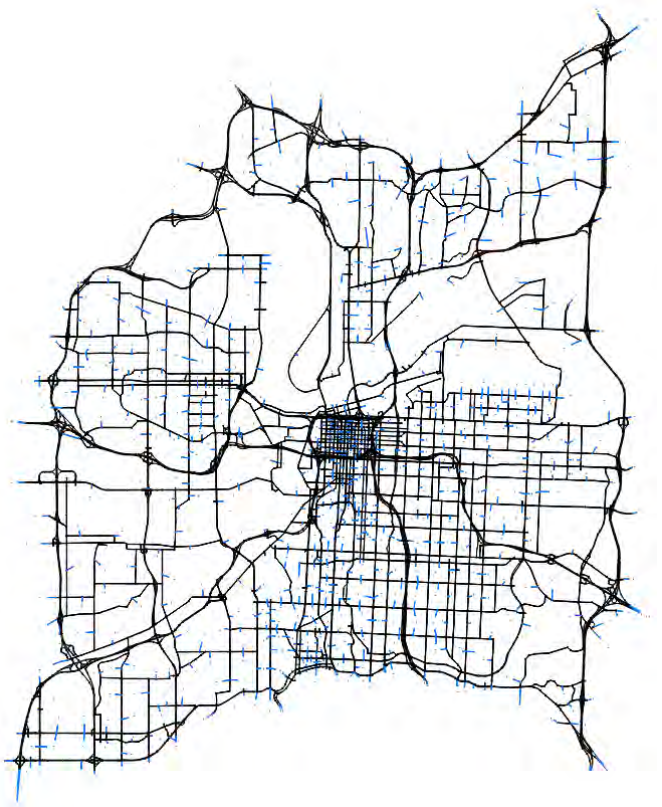


AM7-8 Hour Flow - Entire DTA



AM8-9 Hour Flow - Entire DTA





Volume Validation Regional Meso Example Total Error and RSME %

AM Peak Period - All Counts

Fac#	Facility	Counts	Sum Count	Sum Model	Average Count	Average Model	Avg Error
1	Freeways	148	1,235,701	1,223,770	8349	8269	-1%
2	Expressways	26	112,550	104,913	4329	4035	-7%
3	Princ.Art	130	203,504	191,701	1565	1475	-6%
5	Minor Art.	158	120,446	114,681	762	726	-5%
7	Collector	98	57,701	50,216	589	512	-13%
8	Ramp	273	220,404	260,226	807	953	18%
9	Fwy-Fwy Ramp	63	105,742	117,568	1678	1866	11%
>0	All Roads	896	2,056,048	2,063,075	2295	2303	0%

RSME% by Hour - All Counts

Facility	6-7 AM Hour	7-8 AM Hour	8-9 AM Hour	3 Hour AM Peak
Freeways	0.16	0.15	0.15	0.12
Expressways	0.35	0.20	0.27	0.22
Princ.Art	0.65	0.41	0.41	0.42
Minor Art.	1.01	0.86	0.83	0.83
Collector	0.77	0.77	0.83	0.72
Ramp	0.82	0.57	0.60	0.59
Fwy-Fwy Ramp	0.61	0.50	0.55	0.52
All Roads	0.38	0.33	0.34	0.30

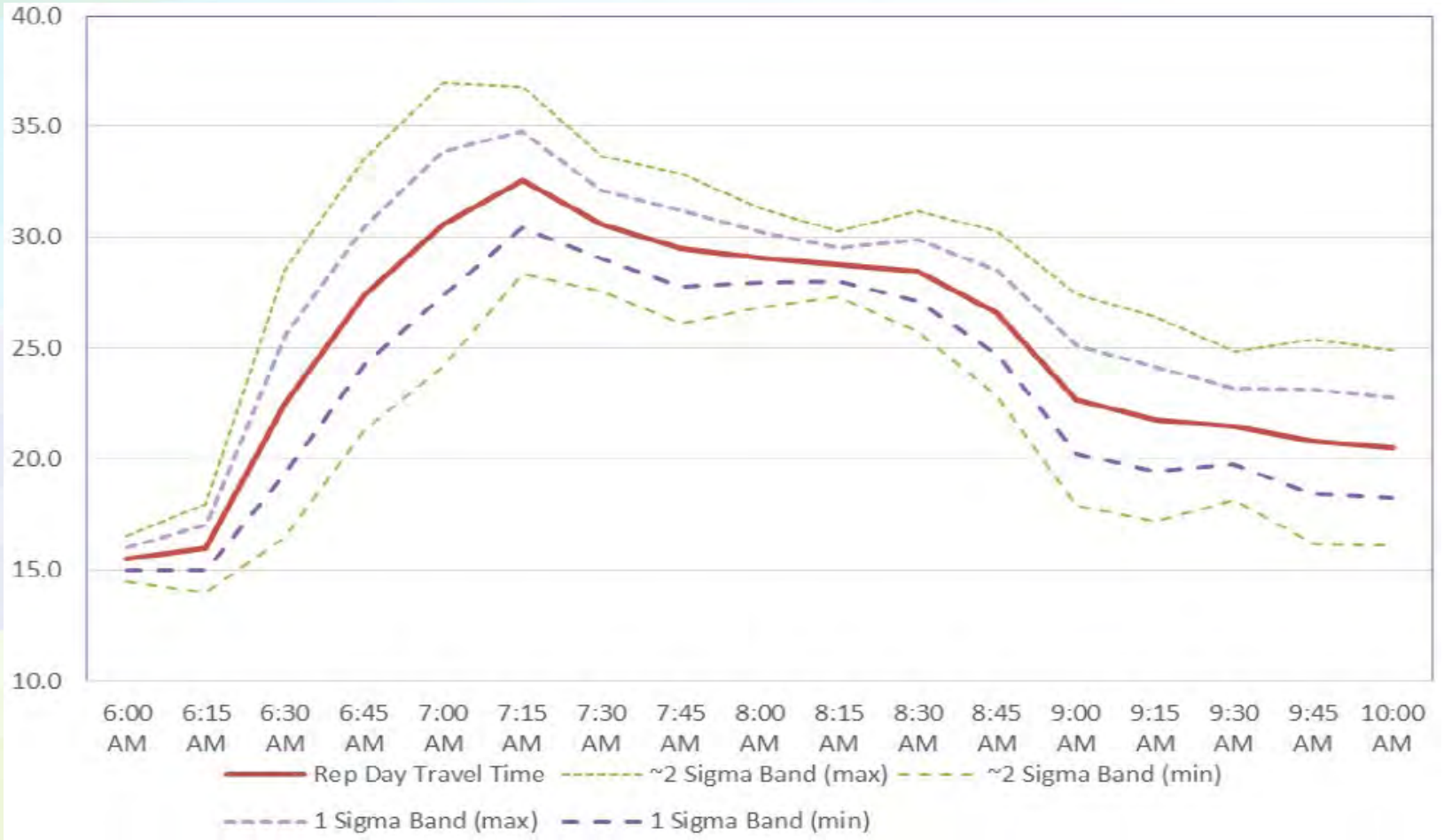
AM Peak Period - Focus Area Counts

Fac#	Facility	Counts	Sum Count	Sum Model	Average Count	Average Model	Avg Error
1	Freeways	31	227,697	229,431	7345	7401	1%
2	Expressways	12	49,723	48,087	4144	4007	-3%
3	Princ.Art	17	18,822	12,563	1107	739	-33%
5	Minor Art.	30	25,709	19,232	857	641	-25%
7	Collector	31	18,873	12,116	609	391	-36%
8	Ramp	64	57,590	58,866	900	920	2%
9	Fwy-Fwy Ramp	20	43,868	49,873	2193	2494	14%
>0	All Roads	205	442,281	430,168	2157	2098	-3%

RSME% by Hour - Focus Area Counts

Facility	6-7 AM Hour	7-8 AM Hour	8-9 AM Hour	3 Hour AM Peak
Freeways	0.16	0.19	0.18	0.16
Expressways	0.40	0.17	0.31	0.23
Princ.Art	0.66	0.48	0.56	0.51
Minor Art.	0.87	0.81	0.78	0.76
Collector	0.89	0.73	0.72	0.69
Ramp	0.85	0.56	0.60	0.57
Fwy-Fwy Ramp	0.37	0.35	0.41	0.34
All Roads	0.40	0.35	0.38	0.33

New FHWA Guidance: Time Variant Variations



New FHWA Guidance: Four Criteria

- Criterion I:

Criterion I: Control for Time-Variant Outliers

This criterion constrains the number of outliers in simulated results.

CRITERION I: 95% of simulated outputs fall within the ~2 Sigma Band, $c_r(t) \pm 1.96\sigma(t)$. Note that if fewer than 20 time intervals are used to characterize time-dynamics, Criterion I is relaxed to allow for one simulated result outside the ~2 Sigma Band.

- Criterion II:

Criterion II: Control for Time-Variant "Inliers"

This criterion ensures the majority of time-variant simulated results fall close to the representative day, and that during the most congested time periods the simulated results are close to the observed data.

CRITERION II: Two-thirds of the simulated results (and both critical time intervals) fall within the 1 Sigma Band for this operational condition.

New FHWA Guidance: Four Criteria

- Criterion III:

Criterion III: Bounded Dynamic Absolute Error (BDAE)

This criterion ensures that, on average, simulated results are close to the observed representative day. The criterion involves a test to ensure that the average simulated absolute error from the representative day over all time intervals is less than or equal to differences from the representative day seen across all days in the operational condition. Let:

$c_r(t)$ Observed value of representative day during time interval t

$c_i(t)$ Observed value of non-representative day within the cluster during time interval t

$\tilde{c}_r(t)$ Simulated performance measure during time interval t

N_T Number of time intervals

$N_{cluster}$ Number of days in the cluster representing this operational condition

Next, calculate the BDAE Threshold:

$$BDAE\ Threshold = \frac{\sum_{i \neq r} \sum_{t} \frac{|c_r(t) - c_i(t)|}{N_T}}{N_{cluster} - 1} \quad (12)$$

CRITERION III is met when:

$$\frac{\sum_{t} |c_r(t) - \tilde{c}_r(t)|}{N_T} \leq BDAE\ Threshold \quad (13)$$

New FHWA Guidance: Four Criteria

- Criterion IV:

Criterion IV: Bounded Dynamic Systematic Error

This criterion ensures that the simulated data are not excessive over- or under-estimators. In this case, the criterion utilizes a similar test to Criterion III but with respect to average simulated error (not absolute).

CRITERION IV is met when:

$$\left| \frac{\sum_t c_r(t) - \tilde{c}_r(t)}{N_T} \right| \leq \frac{1}{3} \times BDAE\ Threshold \qquad (14)$$

Review of Calibration Results

- Test Simulation Runs
  - Replicate the existing calibration results
  - Review / identify problematic areas
    - Areas of poor fit
    - Low speeds
    - High densities
    - Low flow rate
    - Prolonged queuing
    - Unintuitive routes

Additional Validations

- Ideally validate to a second set of data
  - A similar representative day
- Model stability tests
  - Test different random seeds
  - ‘Local link removal’ test
- Model reasonableness tests
  - Does the model respond in a logical way to a change?

ADDITIONAL QUESTIONS?