Solano Transportation Authority (STA) |
Napa Valley Transportation Authority (NVTA)

SOLANO NAPA ACTIVITY-BASED MODEL (SNABM)

Model Documentation | August 27, 2020

PREPARED FOR:
SOLANO TRANSPORTATION AUTHORITY (STA) | NAPA VALLEY TRANSPORTATION AUTHORITY (NVTA)

SUBMITTED BY:
RSG & TJKM

55 Railroad Row
White River Junction, VT 05001
802.295.4999
www.rsginc.com

CONTENTS

1.0 INTRODUCTION 1
2.0 MODEL DESIGN 3
2.1 GENERAL MODEL DESIGN...3
MODEL SEGMENTATION...3
2.2 CT-RAMP DESIGN DETAILS 10
POPULATION SYNTHESIS...12
LONG TERM DECISIONS 12
DAILY DECISIONS...12
TOUR-LEVEL DECISIONS...13
TRIP-LEVEL DECISIONS...14
3.0 DATA 15
3.1 CALIFORNIA HOUSEHOLD TRAVEL SURVEY 15
SNABM CALIBRATION TARGETS PREPARATION...........................16
3.2 TRANSIT ON-BOARD SURVEY...22
3.3 TRANSIT RIDERSHIP DATA...22
DEVELOPMENT OF MODE CHOICE TARGETS 23
3.4 TRAFFIC COUNT DATA...24
4.0 MODEL INPUTS 27
4.1 POPULATIONSIM IMPLEMENTATION.....................................27
POPULATIONSIM VALIDATION 29
4.2 ZONAL DATA...32
4.3 HIGHWAY NETWORK...35
4.4 TRANSIT NETWORK 37
4.5 WALK ACCESS BUFFERS...38

4.6 NONRESIDENTIAL DEMAND 41
5.0 MODEL CALIBRATION 42
5.1 CALIBRATION OVERVIEW 42
ROUND 1 43
ROUND 2 44
ROUND 3 46
5.2 FINAL CALIBRATION RESULTS 46
OVERVIEW 46
LONG-TERM TRAVEL CHOICES 49
TOUR LEVEL CHOICES 53
TRIP LEVEL CHOICES 62
6.0 MODEL VALIDATION 66
6.1 HIGHWAY VALIDATION 66
6.2 TRANSIT VALIDATION 71
6.3 FUTURE YEAR RUN 72
7.0 SUMMARY 77
8.0 APPENDIX A: SNABM USER GUIDE 78
8.1 SOFTWARE REQUIREMENTS 78
8.2 SETUP DETAILS 79
8.3 CONFIGURATION 81
8.4 RUNNING SNABM 83

LIST OF FIGURES

FIGURE 1 SNABM TAZ SYSTEM WITHIN SOLANO/NAPA COUNTIES 5
FIGURE 2 TM1.5 CT-RAMP MODEL DESIGN 10
FIGURE 3 CT-RAMP VISUALIZER TOOL 16
FIGURE 4 COUNT LOCATIONS FOR DAILY VALIDATION 25
FIGURE 5 COUNT LOCATIONS FOR PEAK HOUR VALIDATION 26
FIGURE 6 SF-MTC POPULATION SYNTHESIZER - HOUSEHOLD SIZE SUB-MODEL 28
FIGURE 7 BASE YEAR (2015) POPULATIONSIM VALIDATION -NAPA/SOLANO COUNTIES 30
FIGURE 8 2040 POPULATIONSIM VALIDATION - NAPA/SOLANO COUNTIES 31
FIGURE 9 ZERO EMPLOYMENT ZONE (TAZ 852) INVESTIGATION 33
FIGURE 10 VALLEJO HIGH SCHOOL - STUDENT POPULATION DISTRIBUTION 34
FIGURE 11 SNABM 2015 BASE YEAR NETWORK 36
FIGURE 12 SNABM TRANSIT LINES 38
FIGURE 13 SHORT TRANSIT WALK ACCESS PERCENT 39
FIGURE 14 LONG TRANSIT WALK ACCESS PERCENT 40
FIGURE 15 PERSON TYPE DISTRIBUTION - NAPA/SOLANO 45
FIGURE 16 PERSON TYPE DISTRIBUTION - BAY AREA 47
FIGURE 17 PERSON TYPE DISTRIBUTION - NAPA/SOLANO 47
FIGURE 18 HOUSEHOLD SIZE DISTRIBUTION - BAY AREA 47
FIGURE 19 OVERVIEW AGGREGATE SUMMARIES - NAPA/SOLANO 48
FIGURE 20 TRAVEL RATES - NAPA/SOLANO 49
FIGURE 21 AUTO OWNERSHIP DISTRIBUTION - NAPA/SOLANO 50
FIGURE 22 WORK TOUR LENGTH FREQUENCY DISTRIBUTION - BAY AREA 50
FIGURE 23 UNIVERSITY TOUR LENGTH FREQUENCY DISTRIBUTION - BAY AREA 50
FIGURE 24 SCHOOL TOUR LENGTH FREQUENCY DISTRIBUTION - BAY AREA 51
FIGURE 25 AVERAGE MANDATORY TOUR LENGTHS 52
FIGURE 26 DISTRICT - DISTRICT FLOW OF WORKERS (ACS) 53
FIGURE 27 DISTRICT - DISTRICT FLOW OF WORKERS - SNABM 53
FIGURE 28 DAILY ACTIVITY PATTERN - TOTAL [BAY AREA] 54
FIGURE 29 MANDATORY TOUR FREQUENCY - TOTAL [NAPA/SOLANO] 54
FIGURE 30 PERSONS BY INDIVIDUAL NON-MANDATORY TOURS - TOTAL [NAPA/SOLANO] 55
FIGURE 31 TOTAL TOUR RATE (ACTIVE PERSONS) - NAPA/SOLANO 56
FIGURE 32 HOUSEHOLDS BY JOINT TOUR COMBINATION - BAY AREA 56
FIGURE 33 NON-MANDATORY TOUR LENGTH DISTRIBUTION - NAPA/SOLANO 57
FIGURE 34 AVERAGE NON-MANDATORY TOUR LENGTHS (MILES) - NAPA/SOLANO 57
FIGURE 35 TOUR DEPARTURE PROFILE - NAPA/SOLANO 58
FIGURE 36 TOUR ARRIVAL PROFILE - NAPA/SOLANO 58
FIGURE 37 TOUR DURATION PROFILE - NAPA/SOLANO 58
FIGURE 38 TOUR AGGREGATE DEPARTURE - NAPA/SOLANO 59
FIGURE 39 TOUR AGGREGATE ARRIVAL - NAPA/SOLANO 59
FIGURE 40 TOUR MODE [ZERO AUTO] - BAY AREA 60
FIGURE 41 TOUR MODE [AUTO DEFICIENT] - BAY AREA 60
FIGURE 42 TOUR MODE [AUTO SUFFICIENT] - BAY AREA 61
FIGURE 43 TOUR MODE [TOTAL] - BAY AREA 61
FIGURE 44 TOUR MODE [TOTAL] - NAPA/SOLANO 62
FIGURE 45 STOP FREQUENCY - NAPA/SOLANO 62
FIGURE 46 STOP LOCATION OUT OF DIRECTION DISTANCE (MILES) - NAPA/SOLANO 63
FIGURE 47 AVERAGE OUT OF DIRECTION DISTANCE (MILES) - NAPA/SOLANO 63
FIGURE 48 TRIP DEPARTURE PROFILE - NAPA/SOLANO 64
FIGURE 49 AGGREGATE TRIP DEPARTURE - NAPA/SOLANO 64
FIGURE 50 TRIP MODE - BAY AREA 65
FIGURE 51 TRIP MODE - NAPA/SOLANO 65
FIGURE 52 SCATTERPLOT OF DAILY VOLUMES 70
FIGURE 53 POPULATION PROJECTIONS (2015-2040) 73
FIGURE 54 EMPLOYMENT PROJECTIONS (2015-2040) 73

LIST OF TABLES

TABLE 1 MTC TM1.5 AND SNABM TAZS BY COUNTIES 4
TABLE 2 SAMPLE RATES BASED ON DISTANCE FROM NAPA/SOLANO COUNTIES 6
TABLE 3 MODEL PERSON TYPES 7
TABLE 4 WORK LOCATION SIZE SEGMENTS 7
TABLE 5 SNABM ACTIVITY TYPES 7
TABLE 6 INSERT TIME PERIODS 8
TABLE 7 SNABM TRIP MODES 9
TABLE 8 TM1.5 CT-RAMP DESIGN OUTLINE 11
TABLE 9 SNABM CALIBRATION AND VALIDATION DATASETS 15
TABLE 10 OVERVIEW SUMMARIES 17
TABLE 11 LONG-TERM SUMMARIES 18
TABLE 12 TOUR-LEVEL SUMMARIES 18
TABLE 13 TRIP-LEVEL SUMMARIES 20
TABLE 14 TRANSIT RIDERSHIP - SOLANO/NAPA TRANSIT OPERATORS 23
TABLE 15 2015 LAND USE DATA 32
TABLE 16 SNABM CALIBRATION PLAN 43
TABLE 17 DAILY VALIDATION BY FACILITY TYPE 67
TABLE 18 DAILY VALIDATION BY SCREENLINE 68
TABLE 19 DAILY VALIDATION ON INTERSTATE 80 68
TABLE 20 DAILY VALIDATION IN NAPA COUNTY 69
TABLE 21 DAILY VALIDATION IN VALLEJO 70
TABLE 22 HPMS VMT 70
TABLE 23 TRANSIT RIDERSHIP BY OPERATOR 71
TABLE 24 2040 LAND USE DATA 72
TABLE 25 GROWTH IN VOLUMES BETWEEN 2015 AND 2040 BY FACILITY TYPE 74
TABLE 26 GROWTH IN VOLUMES BETWEEN 2015 AND 2040 BY SCREENLINE 74
TABLE 27 GROWTH IN VOLUMES BETWEEN 2015 AND 2040 ON I-80 75
TABLE 28 GROWTH IN VOLUMES BETWEEN 2015 AND 2040 IN NAPA COUNTY 76
TABLE 29 GROWTH IN VOLUMES BETWEEN 2015 AND 2040 IN VALLEJO 76
TABLE 30 SNABM WORKING DIRECTORY CONTENTS 79
TABLE 31 /CTRAMP DIRECTORY CONTENTS 80
TABLE 32 INPUT DIRECTORY CONTENTS 80
TABLE 33 SETPATH.BAT CONFIGURATION 81

1.0 INTRODUCTION

The Solano Transportation Authority (STA) and the Napa Valley Transportation Authority (NVTA) has been working with TJKM and RSG to update the Solano Napa Activity-Based Model (SNABM). The purpose of this project was to develop a model that is consistent with the Metropolitan Transportation Commission (MTC) regional travel model, while providing additional detail, calibration and validation within Solano and Napa counties. With this goal, the project team decided to deploy MTC's Travel Model 1.5 (TM1.5) for this update. TM1.5 is a major update to MTC's Travel Model One (TM1). It specifically includes capabilities to model Autonomous Vehicles (AVs), Transportation Networking Companies (TNCs) and taxi modes (this is optional in the model and can be turned off if not required). TM1.5 covers the entire nine-county San Francisco Bay Area with 2.8 million households and 7.5 million persons in 2015. The model belongs to the Coordinated Travel – Regional Activity-based Modeling Platform (CT-RAMP) family of models, implemented in several large metropolitan areas including Atlanta, San Diego, Miami, and Chicago.

The older version of SNABM (v1) was developed as a focused version of MTC TM1 and was not calibrated to provide public transportation ridership forecasts. The v1 model was developed using 2013 MTC Plan Bay Area (PBA) Regional Transportation Plan (RTP) land use forecasts. Moreover, the v1 model uses a 15% sample rate which can lead to significant Monte Carlo simulation error in model outputs which is particularly problematic for subarea and corridor studies. In order to reduce this error, the project team implemented an intelligent sampling feature in the SNABM initially implemented by RSG in the Marin County activity-based model. This feature initiates a set of calculations that make the model more practical for county-level analysis. These calculations reduce runtime and hardware requirements and reduce the Monte Carlo simulation (stochastic variance) for Solano and Napa counties.

The major enhancements that have been implemented during this project are as follows:

- Highway and transit networks within Solano and Napa counites were updated based on inputs from STA and NVTA.
- Traffic Analysis Zones (TAZs) were split near major transit stops and the land use was updated based on inputs from local jurisdictions. The land use and demographic data was also updated to MTC's latest 2019 RTP.
- PopulationSim was implemented to generate a synthetic population for the 2015 base year and one 2040 future year scenario.

- The synthetic population in Solano and Napa counties was oversampled three times while households in the rest of the Bay Area were sampled at a decreasing rate with respect to distance from Solano and Napa counties.
- The model was calibrated to observed data for 2015. This includes the 2012-13 California Household Travel Survey (CHTS), 2012-16 5-year American Community Survey (ACS) data, and on-board transit survey data for Bay Area. The model was validated to 2015 traffic counts and transit ridership data from transit operators.

This document first describes the model design and all the enhancements that were implemented under this effort. Next, the observed data used for calibration and validation is described followed by details of model input data preparation. The subsequent sections present model calibration and validation results. The report ends with a summary and recommendations.

2.0 MODEL DESIGN

The SNABM design is based on TM 1.5, and includes all enhancements made by MTC to TM1 as part of the TM1.5 development process. This chapter describes the general model design and refinements implemented for the SNABM update. Appendix A provides details of SNABM setup, and instructions on configuring and running SNABM.

2.1 GENERAL MODEL DESIGN

TM 1.5 belongs to the CT-RAMP family of ABMs. The CT-RAMP model, which is fully described in the following sections, has the following characteristics:

- Utilizes tours (sequences of trips beginning and ending at an anchor location such as home or work) as an organizing principle for the generation of travel and to ensure consistency across trips within a tour.
- Utilizes micro-simulation for modeling travel choices, in which a synthetic population is generated, and explicit mobility and travel choices are made for each decision-maker in the population according to contextual probability distributions.
- Addresses both household-level and person-level travel choices including intra-household interactions between household members.
- Uses an hourly temporal resolution and schedules tours into time-windows to ensure there are no overlapping travel episodes.
- Offers sensitivity to demographic and socio-economic changes observed or expected in the dynamic San Francisco Bay Area. This is ensured by the enhanced and flexible population synthesis procedures as well as by the fine level of model segmentation.
- Accounts for the full set of travel modes including non-motorized travel and transit.
- Reflects and responds to detailed demographic information including household structure, aging, changes in wealth, and other key attributes.

The following sections describe the basic conceptual framework at which the model operates.

Model Segmentation

The model system has been implemented in a micro-simulation framework. A key advantage of the micro-simulation approach is that there are essentially no computational constraints on the number of explanatory variables that can be included in a model specification. However, even with this flexibility, the model system includes some segmentation of decision-makers. Segmentation is a useful tool to both structure models (for example, each person type segment could have their own model for certain choices) and to characterize person roles within a household. Segments can be created for persons as well as households.

Treatment of Space

TM1.5 uses the 1454 TAZ system representing the entire Bay Area. The zones are fairly large for the region, which may somewhat distort the representation of transit access in mode choice. Therefore, walk market segmentation is used by the model. Walk market segments describe the proportion of the zone with respect to walk accessibility to transit. The segments are short walk (within 1/3 mile of transit), long walk (between 1/3 and 2/3 mile of transit), and no-walk (greater than 2/3 mile to the nearest transit stop). To further reduce spatial aggregation bias in the SNABM implementation, TAZs were split near major transit stops in the Napa and Solano counties. This improved the modeling of transit access and egress within the Napa and Solano counties. Table 1 compares the number of TAZs by county in TM1.5 and SNABM. The number of TAZs for Napa County increased from 27 to 224, and from 80 to 769 for Solano County. The total number of TAZs for the 9-county modeling region increased from 1,454 to 2,340. TM1.5 had a total of 21 external stations to represent the internal-external travel. Under this effort, the 4 external zones in Napa and Solano counties were split into 19 external zones, resulting in a total of 36 external zones for SNABM.

TABLE 1 MTC TM1.5 AND SNABM TAZS BY COUNTIES

COUNTY	TM1.5 TAZS	SNABM TAZS
San Francisco	190	190
San Mateo	156	156
Santa Clara	368	368
Alameda	325	325
Contra Costa	171	171
Solano	80	769
Napa	27	224
Sonoma	86	86
Marin	51	51
Total	1,454	2,340

The updated SNABM TAZ system within the Napa and Solano counties is shown on Figure 1. The TM1.5 model scripts were updated to work with the updated TAZ system.

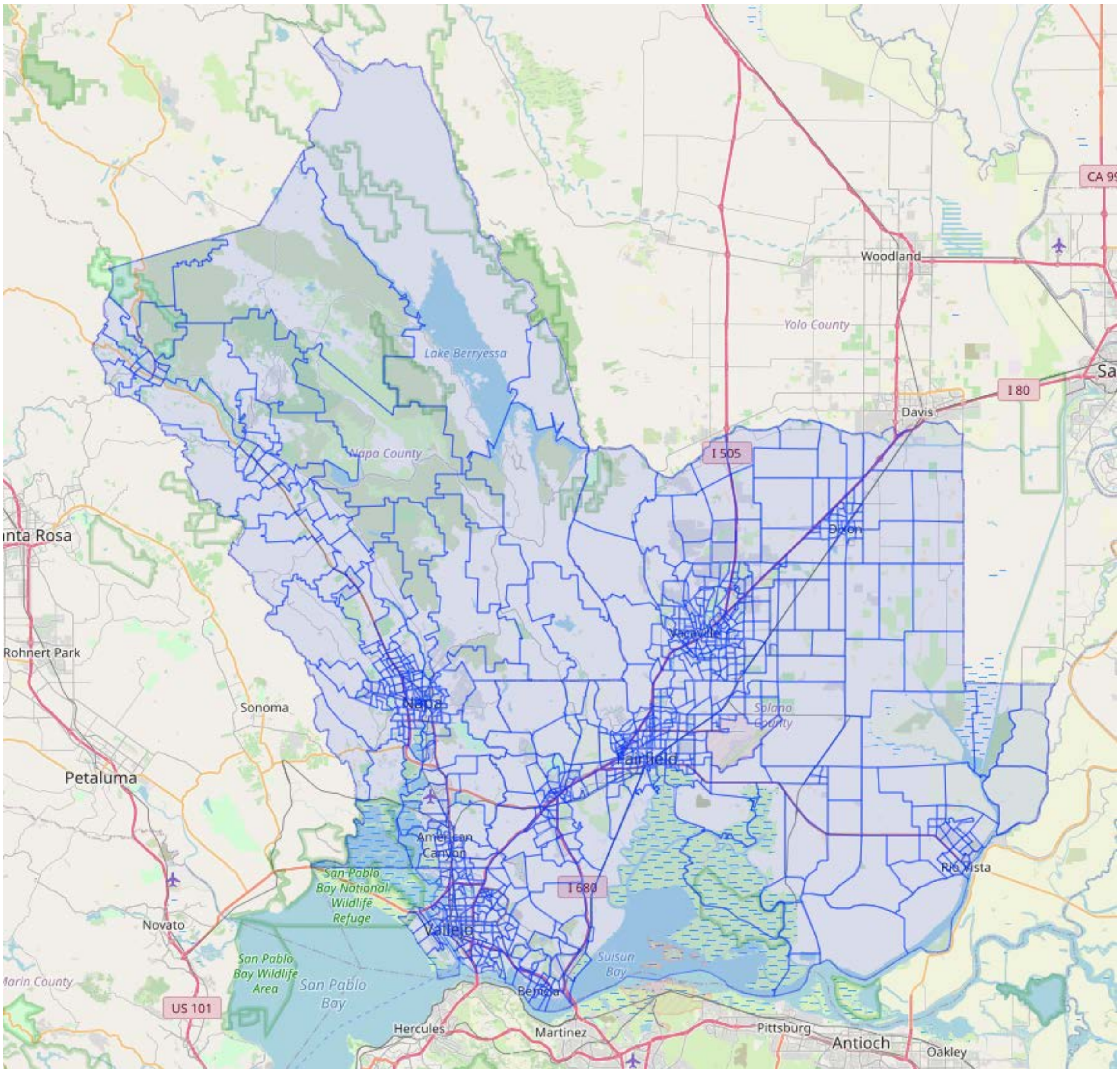


FIGURE 1 SNABM TAZ SYSTEM WITHIN SOLANO/NAPA COUNTIES

Synthetic Population

ABMs operate in a micro-simulation framework, wherein the travel choices of person and household decision-making agents are predicted by applying Monte Carlo methods to behavioral methods. This requires a dataset of persons and households representing the entire population in the modeling region. This dataset is created using a population synthesis software using census population samples and forecasted TAZ-level marginal distributions of socio-demographic variables of interest. Further details on the population synthesis software implementation for SNABM can be found in the Data section.

In order to reduce runtime and increase the stability of model predictions for Solano and Napa County residents, we implemented an intelligent sampling methodology for SNABM. This sampling methodology increases the sample rate of households in Solano and Napa counties, and reduces the sample rate for residents of other counties based on distance to Solano and Napa counties. Households in Solano and Napa counties are sampled at a rate three times higher (300 %) than their representation in the synthetic population, to reduce Monte Carlo variation. Their generated travel is then factored down by three (one third) before assigning their trips to networks. This would be equivalent to taking the average of three model runs for Napa and Solano County residents but offers reduced time for input/output and skimming/assignment procedures. For households in zones outside Napa and Solano Counties, a sampling rate is applied based on zone-to-zone distance to the nearest Napa/Solano County zone. Table 2 shows the sample rates for these zones ranging from 100 percent to 5 percent.

TABLE 2 SAMPLE RATES BASED ON DISTANCE FROM NAPA/SOLANO COUNTIES

DISTANCE TO NEAREST NAPA/SOLANO COUNTY ZONE (MILES)		SAMPLE RATE
0	3	100 %
3	5	50 %
5	10	40 %
10	15	30 %
15	20	20 %
20	40	10 %
Over 40 Miles		5 %
Within Napa/Solano County		300 %

Person Types

A total of eight segments of person-types, shown in Table 3 are used in the TM1.5 and SNABM model system. The person-types are mutually exclusive with respect to age, work status, and school status. These person types are coded in the synthetic population according to person level attributes. The same methodology is used to code person type in the household travel survey data (described in the Data section) used for model calibration.

TABLE 3 MODEL PERSON TYPES

NUMBER	PERSON-TYPE	AGE	WORK STATUS	SCHOOL STATUS
1	Full-time worker	16+	Full-time	None
2	Part-time worker	16+	Part-time	None
3	College student	16+	Any	College +
4	Non-worker & non-student	16 – 64	Unemployed	None
5	Non-working senior	65+	Unemployed	None
6	Driving age student	16-19	Any	Pre-college
7	Non-driving student	6 – 15	None	Pre-college
8	Pre-school	0-5	None	None

Further, workers are stratified by their income segments, shown in Table 4. These income groups are used to segment the destination choice size terms for work location choice.

TABLE 4 WORK LOCATION SIZE SEGMENTS

SEGMENT	INCOME RANGE
Low income segment	<$30K
Medium income segment	$30K - $60K
High income segment	$60K - $100K
Very high-income segment	>$100K

Activity Types

The activity types modeled in SNABM are shown in Table 5. The activity types are also grouped according to whether the activity is mandatory, maintenance, or discretionary, and eligibility requirements are assigned determining which person-types can be used for generating each activity type. The classification scheme of each activity type reflects the relative importance or natural hierarchy of the activity, where work and school activities are typically the most inflexible in terms of generation, scheduling and location, whereas discretionary activities are typically the most flexible on each of these dimensions. Each out-of-home location that a person travels to in the simulation is assigned one of these activity types, as shown in the table.

TABLE 5 SNABM ACTIVITY TYPES

TYPE	PURPOSE	DESCRIPTION	CLASSIFICATION
1	Work	Working at regular workplace.	Mandatory
2	University	College +	Mandatory
3	Grade/High School	Grades K-12	Mandatory
4	Escorting	Pick-up/drop-off passengers	Maintenance
5	Shopping	Shopping away from home.	Maintenance

6	Other Maintenance	Personal business/services, and medical appointments.	Maintenance
7	Social/Recreational	Recreation, visiting friends/family.	Discretionary
8	Eat Out	Eating outside of home.	Discretionary
9	Other Discretionary	Volunteer work, religious activities.	Discretionary

Treatment of Time

The TM1.5 model system operates on an hourly temporal resolution, beginning with 3 A.M. and ending with 3 A.M. the next day. Temporal integrity is ensured so that no activities are scheduled with conflicting time windows, except for short activities/tours that are completed within an hour increment. For example, a person may have a short tour that begins and ends within the 8am-9am period, as well as a second longer tour that begins within this time period but ends later in the day.

A critical aspect of the model system is the relationship between the temporal resolution used for scheduling activities, and the temporal resolution of the network simulation periods. Although each activity generated by the model system is identified with a start time and end time in one-hour increments, level-of-service matrices are only created for five aggregate time periods – early A.M., A.M., midday, P.M., and evening. The trips occurring in each time period reference the appropriate transport network depending on their trip mode and the mid-point trip time. The definition of time periods for level-of-service matrices is given in Table 4, below.

TABLE 6 INSERT TIME PERIODS

NUMBER	DESCRIPTION	BEGIN TIME	END TIME
1	Early A.M.	3:00 A.M.	5:59 A.M.
2	A.M. Peak	6:00 A.M.	8:59 A.M.
3	Midday	9:00 A.M.	2:59 P.M.
4	P.M. Peak	3:00 P.M.	6:59 P.M.
5	Evening	7:00 P.M.	2:59 A.M.

Trip Modes

There are 21 trip modes in TM1.5 and SNABM including auto by occupancy and toll/non-toll choice, walk and bike non-motorized modes, and walk and drive access to five different line haul modes. TM1.5 was updated to explicitly represent the choice of Taxi and Transportation Network Company (TNC) modes. A new Taxi\TNC nest was added with sub-alternatives for traditional taxi versus new TNC modes including single party TNC versus shared TNC. These capabilities can be used to represent different assumptions about actual and/or perceived times and costs of taxi and TNC modes. Mode share sensitivity tests can be designed by varying these assumptions. Table 7 shows the complete list of 23 trip modes in TM1.5 and SNABM.

TABLE 7 SNABM TRIP MODES

NUMBER	TRIP MODE
1	Drive Alone (Free)
2	Drive Alone (Pay)
3	Shared Ride 2 (Free)
4	Shared Ride 2 (Pay)
5	Shared Ride 3+ (Free)
6	Shared Ride 3+ (Pay)
7	Walk
8	Bike
9	Walk – Local Bus
10	Walk – Light Rail/Ferry
11	Walk – Express Bus
12	Walk – Heavy Rail
13	Walk – Commuter Rail
14	Drive – Local Bus
15	Drive – Light Rail/Ferry
16	Drive – Express Bus
17	Drive – Heavy Rail
18	Drive – Commuter Rail
19	Taxi
20	TNC – Single
21	TNC – Shared

2.2 CT-RAMP DESIGN DETAILS

The general CT-RAMP model design for TM1.5 is shown in Figure 2[1]. The green boxes in the schematic represent choices ranging from long-term decisions to stop level decisions. The automobile ownership, coordinated daily activity pattern, and joint tour frequency models cover choices that relate to the entire household or a group of household members and assume explicit modeling of intra-household interactions. The other models are applied to individuals, though they may consider household-level influences on choices.

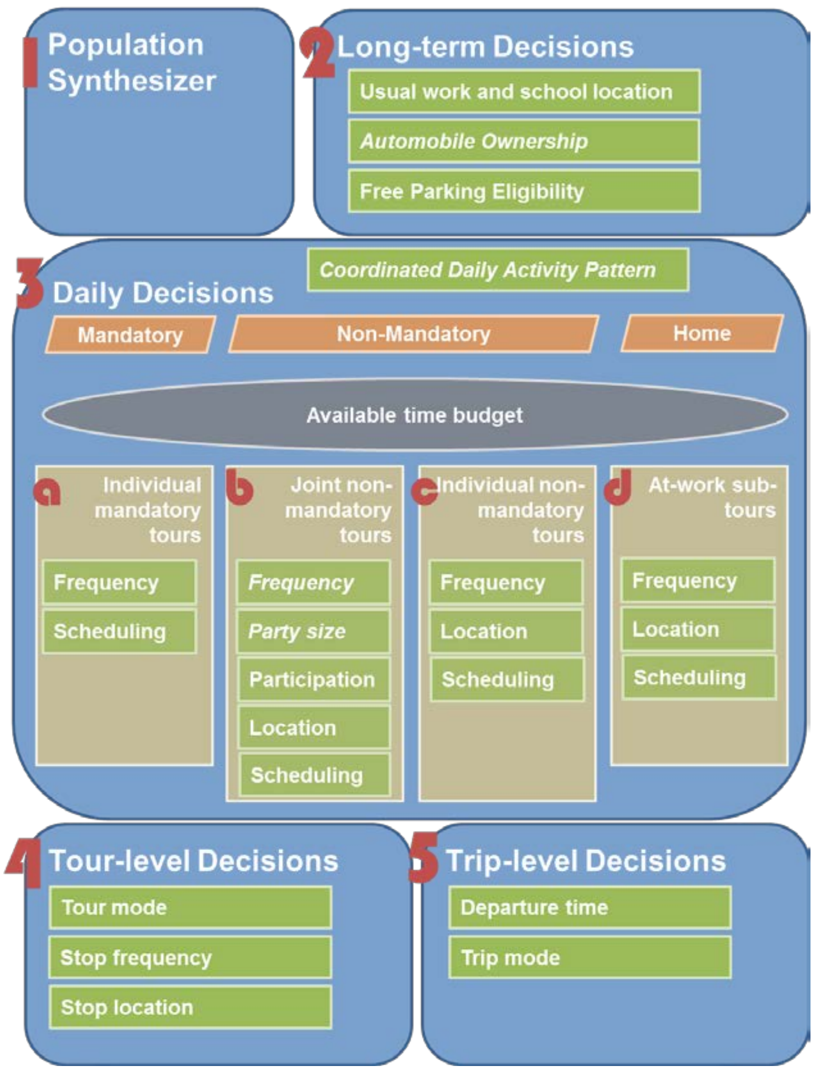


FIGURE 2 TM1.5 CT-RAMP MODEL DESIGN

[1] https://github.com/BayAreaMetro/modeling-website/wiki/ModelSchematic

The basic sequence of CT-RAMP sub-models and associated travel choices is outlined in Table 8 and discussed in detail below.

TABLE 8 TM1.5 CT-RAMP DESIGN OUTLINE

STEP	DESCRIPTION
Step 1: Population Synthesis	
1.1	**Synthetic person and household databases** are synthesized using a Population Synthesizer with Census data and land use forecasts as inputs
Step 2: Long-term Decisions	
2.1	**Mandatory activity locations** for each household member are determined (workplace/university/school)
2.2	**Household car ownership** is forecasted based on household/person attributes and household accessibilities
2.3	**Free parking eligibility** for workers
Step 3: Daily Decisions	
3.0	**Daily travel pattern** type for each household member (main activity combination, at home versus on tour) with a linkage of choices across various person categories.
3.a	**Individual mandatory activities/tours** for each household member are forecasted (note that locations of mandatory tours have already been determined in long-term choice model)
3.a.1	Frequency of mandatory tours is determined
3.a.2	Mandatory tour time of day (departure/arrival time combination) is determined
3.b	**Joint travel tours** for each household member are forecasted (conditional upon the available time window left for each person after the scheduling of mandatory activities
3.b.1	Joint tour frequency/composition, which predicts the exact number of joint tours (1 or 2), the purpose of each tour, and the composition of each tour (adults, children, or mixed)
3.b.2	Person participation in each joint tour
3.b.3	Primary destination for each joint tour
3.b.4	Joint tour time of day (departure/arrival time combination)
3.b.5	Joint tour frequency/composition, which predicts the exact number of joint tours (1 or 2), the purpose of each tour, and the composition of each tour (adults, children, or mixed)
3.c	**Individual non-mandatory tours** for each household member are forecasted (conditional upon the available time window left for each person after the scheduling of mandatory and joint non-mandatory activities)
3.c.1	Individual non-mandatory tour frequency, applied for each person

STEP	DESCRIPTION
3.c.2	Individual non-mandatory tour primary destination
3.c.3	Individual non-mandatory tour departure/arrival time
3.d	**At-work sub-tours** for each household member are forecasted (conditional upon the available time window within the work tour duration)
3.d.1	At-work sub-tour frequency, applied for each work tour
3.d.2	At-work sub-tour primary destination
3.d.3	At-work sub-tour departure/arrival time
Step 4: Tour-level Decisions	
4.1	**Tour mode choice**
4.2	**Intermediate stop frequency and purpose**
4.3	**Intermediate stop location**
Step 5: Trip-level Decisions	
5.1	**Intermediate stop departure time choice**
5.2	**Trip mode choice** conditional upon the tour mode

Population Synthesis

The starting point for CT-RAMP ABMs is a synthetic population representing the entire population of the modeling region. The synthetic population for both TM1.5 and SNABM is generated using PopulationSim[2]. The synthetic population needs to be generated only once unless there is a change in the inputs to the population synthesizer such as a different land-use scenario or model year. The CT-RAMP software loads the input database of person and households created by PopulationSim and runs CT-RAMP sub-models for all household and person objects.

Long Term Decisions

The model starts with long-term choices that relate to the usual workplace/university/school location (sub-model 2.1) for each worker and student. Next, the auto ownership model (sub-model 2.2) is run to estimate number of autos for each household. It is followed by free parking eligibility model (sub-model 2.3) for each worker.

Daily Decisions

The daily activity pattern type of each household member (sub-model 3.0) is the first travel-related sub-model in the modeling hierarchy. This model classifies daily patterns by three types: 1)

[2] https://activitysim.github.io/populationsim/

mandatory (that includes at least one out-of-home mandatory activity), 2) non-mandatory (that includes at least one out-of-home non-mandatory activity, but does not include out-of-home mandatory activities), and 3) home (that does not include any out-of-home activity and travel).

The pattern choice set contains a non-travel option in which the person can be engaged in an in-home activity only (purposely or because of being sick) or can be out of town. In the model system application, a person who chooses a non-travel pattern is not considered further in the modeling stream, except that they can make an internal external trip. Daily pattern-type choices of the household members are linked in such a way that decisions made by some members are reflected in the decisions made by the other members.

The next set of sub-models (3.a) define the frequency, and time-of-day for each mandatory tour. The scheduling of mandatory activities is generally considered a higher priority decision than any decision regarding non-mandatory activities for either the same person or for the other household members. “Residual time windows,” or periods of time with no person-level activity, are calculated as the time remaining after mandatory tours have been scheduled. The temporal overlap of residual time windows among household members are estimated after mandatory tours have been generated and scheduled. Time window overlaps, which are left in the daily schedule after the mandatory commitment of the household members has been made, affect the frequency of joint and individual non-mandatory tours, and the probability of participation in joint tours. At-work sub-tours are modeled next, taking into account the time window constraints imposed by their parent work tours (sub-models 3.d).

The next major model component relates to joint household travel. Joint tours are tours taken together by two or more members of the same household. This component predicts the exact number of joint tours by travel purpose and party composition (adults only, children only, or mixed) for the entire household (sub model 3.b.1), and then defines the participation of each household member in each joint household tour (sub-model 3.b.2). It is followed by choice of destination (sub-model 3.b.3), and time-of-day (sub model 3.b.4).

The next stage relates to individual maintenance (escort, shopping, and other household-related errands) and discretionary (eating out, social/recreation, and other discretionary) tours. All these tours are generated by person in sub-model 3.c.1. Their destination, and time of day are chosen next (sub models 3.c.2, and 3.c.3).

Tour-level Decisions

The next set of sub-models relate to tour level decisions. They include the choice of tour mode (sub-model 4.1), frequency of stops in each direction and purpose of each stop (sub-model 4.2), and location of each stop (sub-model 4.3).

Trip-level Decisions

The tour level sub-models are followed by trip level sub-models that add details at trip or stop level. They include the choice of intermediate stop departure (sub-model 5.1) and trip mode choice (sub-model 5.2). Finally, the trips are assigned to highway and transit networks depending on trip mode and time-period.

3.0 DATA

This section describes the observed datasets used for calibration and validation of SNABM. The base year for SNABM is 2015, and therefore, datasets closest to 2015 were selected for model calibration and validation. Table 9 presents the list of all datasets used for SNABM calibration and validation followed by details of each dataset.

TABLE 9 SNABM CALIBRATION AND VALIDATION DATASETS

DATASET	YEAR	SOURCE	PURPOSE
California Household Travel Survey (CHTS)	2012-13	California Department of Transportation (Caltrans)	CT-RAMP Calibration
Transit On-Board Surveys (OBS)	2012-17	MTC On-Board Survey Program	CT-RAMP Mode Choice Calibration
American Community Survey (ACS)	2012-16	US Census Bureau	CT-RAMP Calibration (auto ownership, work from home and county-to-county worker flow)
Traffic Count Data	2015		Highway Validation
Transit Ridership Data	2015-16	MTC Statistical Summary of Bay Area Transit Operators, Ridership Reports of Bay Area Transit Operators	Transit Validation

3.1 CALIFORNIA HOUSEHOLD TRAVEL SURVEY

The California Household Travel Survey (CHTS) is the primary dataset utilized to calibrate most of the resident models (CT-RAMP sub-models) in SNABM. The calibration targets were based on the Bay Area sample of CHTS, conducted by the California Department of Transportation (Caltrans) in 2012-13. The Bay Area sample consists of 9,719 households. Households that reported travel on weekend days and during non-school days were omitted from the model calibration target tabulations. Thus, the calibration targets are based on a sample of 4,305 households who made weekday trips when school was in session. The survey expansion factors were adjusted to account for this reduction in the base household sample.

CHTS details the activity and travel information for all household members during a specific 24-hour period. The travel data was processed into tours and trips records consistent with the CT-RAMP framework. The format of the processed CHTS dataset is same as the CT-RAMP outputs.

SNABM Calibration Targets Preparation

The CHTS dataset has been previously used to calibrate TM1.5. The R scripts used for creating the calibration targets for TM1.5 were used as the starting point for SNABM. We updated the R scripts to geocode the CHTS dataset to the SNABM TAZ system. SNABM produces outputs for the entire 9-county Bay Area region, but the calibration and validation effort focused mainly on Solano and Napa County. We developed separate Napa and Solano County specific calibration targets along with the 9-county version. To aid in the calibration and validation process, we deployed a visualization and diagnostics tool designed for CT-RAMP based ABMs (CT-RAMP Visualizer Tool[3]). The Visualizer Tool creates a HTML dashboard of summary comparisons of various models in the CT-RAMP framework. The tool can compare model performance against a household survey as part of a validation exercise or compare two model runs for sensitivity testing. Two dashboards are generated, one for the entire 9-county region and a second one only for the Napa/Solano County residents for a more focused analysis. Figure 3 shows the snapshot of the *Overview* page of the dashboard.

FIGURE 3 CT-RAMP VISUALIZER TOOL

The summaries and charts in the dashboard have been grouped based on their order of implementation within the CT-RAMP modeling framework. The tab names on the navigation bar bears the name of these groups – Overview, Long Term, Tour Level, and Trip Level.

[3] https://github.com/BayAreaMetro/travel-model-two/tree/master/visualizer

The subsequent sections discuss these summaries in more detail. Note that the discussion in these sections is limited to form of summaries, their definition/purpose, and if any special processing is performed. These sections do not present any numerical targets. The numerical targets are included in the next chapter (Model Calibration) which discusses the performance of the model outputs compared to the observed travel data.

Overview Summaries

The overview summaries compare the aggregate travel behavior between the model and the observed data. For the region wide vehicle-miles travelled (VMT) calculation, the person trips are converted to vehicle trips using an occupancy factor of 1, 2, and 3.33 for drive-alone, shared-ride 2, and shared-ride 3+, respectively. Aggregate VMT is computed by summing vehicle trips by their respective distance skim. Table 10 provides a complete list of statistics and summaries available on the overview page.

TABLE 10 OVERVIEW SUMMARIES

SUMMARY	DEFINITION
Population	Total number of persons in the San Diego County
Households	Total number of households
Total Tours	Total number of tours
Total Trips	Total number of trips
Total Stops	Total number of stops made on tours
Total VMT	Total vehicle miles travelled (in miles)
Tours Per Person	Total tours divided by total persons
Trips Per Person	Total trips divided by total persons
Stops Per Person	Total stops divided by total persons
Trips Per Households	Total trips divided by total households
Person Type Distribution	Number of persons by person type
Household Size Distribution	Number of households by household size

Long Term

These summaries (see Table 11) are prepared to examine the long-term choices in CT-RAMP. The long-term choices include, household auto ownership and mandatory (work, university, and school) tour destination choice. The commuter flow summaries examine the movement (flow) of workers from residence county to the work county. The flow of workers is a useful summary in understanding the distribution of mandatory travel in the region. The summaries also include average tour lengths for the three mandatory purposes by home county. While most of the targets were computed from CHTS dataset, auto ownership and worker flow distribution came from the American Community Survey (ACS) database.

TABLE 11 LONG-TERM SUMMARIES

SUMMARY	DEFINITION
Long Term Models	
Auto Ownership	Number of vehicles per household.
Mandatory TLFD	
Work	Distribution of workers by distance between home and usual workplace
University	Distribution of university going students by distance between home and university
School	Distribution of school going students by distance between home and school
Flows & Tour Lengths	
District-District Flows of Workers	Number of workers by home district and usual workplace district
Average Mandatory Tour Lengths	Average distance between home and mandatory (work, university, and school) location by county
Employment vs Workers	Scatter plot between number of jobs in an TAZ vs number of workers assigned to that TAZ

Tour Level

Table 12 presents the list of tour-level summaries. These include individual tour frequency, joint tour frequency, tour destination, tour time-of-day, and tour mode summaries.

TABLE 12 TOUR-LEVEL SUMMARIES

SUMMARY	DEFINITION
Tour Summaries	
Daily Activity Pattern (DAP)	Percentage of persons by their daily activity pattern (M: mandatory, N:non-mandatory, and H:at-home). Available by person type.
Mandatory Tour Frequency	Percentage of persons by mandatory activity type (1 work, 2 work, 1 school, 2 school, and >1 work & >1 school). Only persons with DAP as M are included. Available by person type.
Total Tour Rate (only active Persons)	Tours per person by person type. Includes joint tours as well.
Persons by Individual Non-Mandatory Tours (by person type)	Percentage of persons with number of non-mandatory tours (0, 1, 2, 3+). Available by person type.

SUMMARY	DEFINITION
Joint Tours	
Joint Tour Frequency	Frequency (percentage) of households by number of joint tours
Joint Tour Composition	Frequency (percentage) of tours by composition (adults only, children only, and mixed)
Joint Tours by Number of Household Members	Frequency (percentage) of joint tours by the number of household member participating
Joint Tours by Household Size	Frequency (percentage) of households by household size and the number of joint tours per household
Party Size Distribution by Joint Tour Composition	Frequency (percentage) of joint tours by party size and tour composition.
Destination	
Non-Mandatory Tour Length Distribution	Distribution of tours by distance between tour origin and destination for each non-mandatory purpose
Average Non-Mandatory Tour Lengths (Miles)	Average tour lengths between origin and destination by non-mandatory tour purpose
TOD	
Tour Departure-Arrival Profile	Distribution of tours by departure time (time leaving home or work) and arrival time (time arriving back at home or work) in 30-mins time bins. Available by tour purpose. Distribution of tour duration is also available.
Tour Aggregate Departure-Arrival Profile	Distribution of tours by departure time and arrival time in five model time periods
Tour Mode	
Tour Mode Choice	Distribution of tours by tour mode and HH auto sufficiency (0-autos, autos>=adults, and autos <adults). Available by tour purpose.

The tour summaries provide details of the overall tour participation of the individuals. At a higher level, a distribution of individuals' daily activity pattern in three aggregated categories (mandatory, non-mandatory, and at-home) describes the primary type of daily travel made by the individuals. The average number of tours by an active person measures the activity level of the individuals who travel on a given day.

Further, the joint tours summaries examine joint travel by members of the same household. The joint tour summaries explore joint tour participation in detail by looking at joint tour frequency, party size, and party composition (adults only, children only, or mixed).

The destination summaries explore the spatial aspect of the non-mandatory travel in the region. Note that mandatory travel is already examined in the long-term summaries. The spatial distribution of non-mandatory travel destinations is examined by summarizing the average tour lengths for each non mandatory tour purpose. The tour lengths are calculated as the distance between the origin and the primary destination of a tour. The tour lengths are summarized as tour length frequency distributions and also average tour lengths. These two summaries together provide a general understanding of individuals' behavior of choosing locations for non-mandatory activity participation.

The temporal behavior is inspected in the time of day (TOD) summaries. The TOD summaries include frequency of tours by departure from tour origin and arrival back at the tour origin in two temporal details: one-hour time-of-day bins and five model time periods (Early AM, AM Peak Period, Mid-Day, PM Peak period, and Evening). The entire day is divided into one-hour bins. The first bin covers the period from 3 am to 6 am, while the last bin goes from 11 pm to 3 am, resulting in a total of 18 bins.

The tour mode summaries are indicators of individual's preference of mode (vehicle) of travel from origin to primary destination and back. The tour mode summaries examine the mode preference by tour purpose and by availability of auto(s) for the travel. The availability of auto(s) is characterized in terms of auto sufficiency of the individual's household. Three classes are created based on the number of cars and the number of adults in the household to define the auto sufficiency of a household as 0-autos, autos<adults, and autos>=adults. The household travel survey generally underestimates the transit ridership in a modeling region. The transit onboard surveys (OBS) are more reliable in representing the transit ridership of a region. The tour mode choice summaries from the CHTS data were augmented with transit tour mode choice summaries from OBS. More details on preparation of mode choice targets are presented in the section on Development of Mode Choice Targets.

Trip Level

The Trip Level Summaries group are created to examine trip generation in CT-RAMP. The trip level summaries include stop frequency, stop location, time of day, and trip mode. These examine the magnitude, spatial distribution, temporal distribution, and mode preference of travel at a trip or activity level. Table 13 presents a list of all summaries available in this category.

TABLE 13 TRIP-LEVEL SUMMARIES

SUMMARY	DEFINITION
Stop Frequency	
Stop Frequency - Directional	Percent of tours by number of stops on the tour and tour direction (inbound/outbound). Available by tour purpose and total.
Stop Frequency - Total	Percent of tours by number of total stops (inbound + outbound) on the tour

SUMMARY	DEFINITION
Stop Purpose by Tour Purpose	Percent of intermediate stops by stop purpose and tour purpose
Location	
Stop Location - Out of Direction Distance	Distribution of intermediate stops by out of direction distance and tour purpose
Average Out of Direction Distance (Miles)	Average out of direction distance by tour purpose
TOD	
Stop & Trip Departure	Stop - Distribution of stops in 30-mins departure time bins. Trip - Distribution of all trips in 30-mins departure time bins. Trips include trips to/from intermediate stops and the tour primary destination. The summaries are also available by tour purpose.
Aggregate Stop & Trip Departure	Frequency of stops and trips in five model time periods. The summaries are also available by tour purpose.
Trip Mode	
Trip Mode Choice	Distribution of trips by trip mode and tour mode. Tour mode constraints the availability of each trip mode and influences the utility of each available trip mode). The summaries are also available by tour purpose.

The stop location summaries include an analysis of out of direction distance on a tour. The out of direction distance is defined as the additional distance to be travelled to pursue an activity at a stop location. For stops in the outbound direction, it is based on the distance between the last known location (the tour origin or previous outbound stop) and the tour primary destination. For stops in the inbound direction, it is based on the distance between the last known location (the tour primary destination or previous inbound stop) and the tour origin. The analysis of the out of direction distance is by tour purpose and in two forms: frequency distribution of the distance and average distance.

The time of day summaries include frequency of stops and trips by their respective departure times. Like the time of day summaries at tour level, the stop and trip frequencies are created in two temporal details: one-hour time-of-day bins and five model time periods (Early AM, AM Peak Period, Mid-Day, PM Peak period, and Evening). The entire day is divided into one-hour bins. The first bin covers the period from 3 am to 6 am, while the last bin goes from 11 pm to 3 am, resulting in a total of 18 bins.

Lastly, the trip mode summaries describe mode preference at trip level. The trip mode preference is summarized as frequency of trips by trip mode and tour mode, which constrains the availability of each trip mode. The frequencies are also examined by tour purpose. The trip mode choice summaries are also augmented by the more reliable transit trip mode choice summaries from the OBS. The creation of mode choice calibration targets is discussed in further detail below.

3.2 TRANSIT ON-BOARD SURVEY

MTC began a coordinated regional surveying effort in 2012 with a goal of surveying all the Bay Area operators over five years. The BART customer survey was completed in spring 2015 and most recently, the SF Muni and Santa Clara VTA customer surveys were completed in fall 2017.

MTC merged the on-board surveys conducted by for different operators into a single dataset. Tour mode and detailed tour purpose were defined prior to the development of the calibration targets. For the records with missing data, automobile sufficiency, access and egress modes are determined using Monte Carlo simulation, with probabilities that are based on the observed distribution by operator. The output of this data processing is a standard dataset with tour purpose, tour mode, access mode and automobile sufficiency coded consistent with TM1.5 definitions.

Since the OBS data was collected between 2012 and 2017, MTC initially expanded it to match the 2015 ridership targets by operator as reported in the MTC's Statistical Summary of Bay Area Transit Operators[4]. However, the initial expansion process revealed inconsistent transfer rates between surveys conducted for each major operator (local buses, express buses, ferry, BART, and Caltrain). In most cases, the local bus surveys underestimate transfers to premium modes such as heavy rail and commuter rail. A secondary survey expansion process was implemented using PopulationSim to match operator level boarding targets and resolve inconsistencies in the reported transfers between operators.

The final OBS dataset was used to develop transit calibration targets for TM1.5. Several small, typically non-federal claiming operators have not been surveyed. Their contributions to the calibration targets were estimated based on 2014-15 ridership and access mode shares from a representative surveyed agency.

3.3 TRANSIT RIDERSHIP DATA

Table 14 presents the transit operator-level transit ridership data obtained from the weighted MTC transit on-board survey, MTC Statistical Summary of Bay Area Transit Operators and ridership reports of Bay Area transit operators for the purposes of transit ridership validation. While TM1.5 predicts transit ridership for all operators in Bay Area, transit ridership validation was performed only for the major operators in Solano and Napa counties.

[4] https://mtc.ca.gov/sites/default/files/StatSumBook2016-11-2-2017.pdf

TABLE 14 TRANSIT RIDERSHIP - SOLANO/NAPA TRANSIT OPERATORS

OPERATOR	RIDERSHIP	YEAR	SOURCE
Napa Vine	3,865	2015	MTC OBS dataset
SolTrans	5,047	2015	MTC OBS dataset
Vallejo Ferry	2,686	2015	MTC Statistical summary of Bay Area operators (https://mtc.ca.gov/sites/default/files/Statistical_Summary_2015.pdf)
Farfield Bus	3,373	2015	MTC OBS dataset
Vacaville Bus	1,765	2015	MTC OBS dataset
Total	16,736		

Development of Mode Choice Targets

There are two mode choice models in CT-RAMP; a tour mode (also referred to as a preferred mode) choice model, and a trip mode (also referred to as a mode switching) model. The tour mode choice model conditions and constrains the modes allowed for the trips on a tour. The trip mode choice model reflects the fact that travelers can and do switch modes on a tour. For example, a traveler may leave home with a child, drop that child off at school, then drive to work. The first trip is a shared-ride 2 trip, the second trip mode is drive-alone. Since there are two models, we need two sets of mode choice targets.

The household travel survey (CHTS in this case) is the primary source for tour level information. However, household travel survey typically do not accurately represent the transit usage in the region. On the other hand, the transit on-board survey is carefully designed to represent transit usage more precisely. The OBS data gives detailed information about transit trips, but little information about transit tours. Therefore, both the household travel survey and the OBS dataset are used to inform tour and trip mode choice targets. RSG processed the CHTS data to produce tour mode choice summaries by tour purpose and automobile sufficiency category. Cross-tabulations of trips by trip mode and tour mode for each tour purpose were also developed from the CHTS data. Transit trips by tour purpose, tour mode and automobile sufficiency category were summarized from the OBS dataset (see Transit On-Board Survey).

There are certain restrictions to tour purpose, tour mode and/or market segment combinations that are included in some model specifications even though these combinations may be observed and/or reported in the travel surveys. For example, in some models (including SNABM), persons from zero car households are not allowed to drive alone, though in reality they may borrow or rent a car and do so. These restrictions are imposed to simplify the model specification without unduly sacrificing its realism. We have applied such restrictions to the cross-tabulations from the survey data. Additionally, since the on-board surveys do not provide any information on whether a tour is individual, partially joint or fully joint, the individual and joint tour splits for non-mandatory tours were developed from the household travel survey data and applied to on-board survey data.

Next, the average number of trips per tour is derived for each transit tour mode and tour purpose using the household survey. This was used to estimate the number of transit tours that corresponds to the number of transit trips. The minimum number of average trips per tour in the CHTS dataset is set to 2.0, to reflect at least one outbound and one return trip. The derived transit tour totals were split into different automobile sufficiency segments based on the observed percentage of transit trips by tour purpose, automobile sufficiency, and access mode as implied by the OBS.

The number of transit tours by transit access mode and automobile sufficiency category was held constant while total tours for other modes were scaled to match the total number of tours estimated by the model for each tour purpose and automobile sufficiency segment. This step needs to be repeated after each calibration run because the total number of tours will change. Holding the transit tour target constant is a necessary requirement for the calibrated model to predict transit trips and transit boardings that are consistent with the OBS boardings. Finally, the transit trips by tour access mode and technology from the on-board survey were swapped with the corresponding values in the CHTS tour mode by trip mode table.

Please note that TM1.5 has been calibrated using the mode choice targets developed as per this methodology. For SNABM calibration, we deployed the same methodology to prepare mode choice targets specific to Solano and Napa counties. The OBS dataset was filtered to only include Napa and Solano county residents.

3.4 TRAFFIC COUNT DATA

Traffic counts were compiled from several sources including conventional traffic counts prepared for the previous model development effort, the Caltrans Performance Measurement System (PeMS) data[5] and Caltrans Traffic Count Database[6]. Figure 4 shows locations for daily counts and figure 5 shows those for peak hour counts in both Napa and Solano.

[5] PeMS, http://pems.dot.ca.gov/
[6] Caltrans Traffic Count Database, http://traffic-counts.dot.ca.gov/

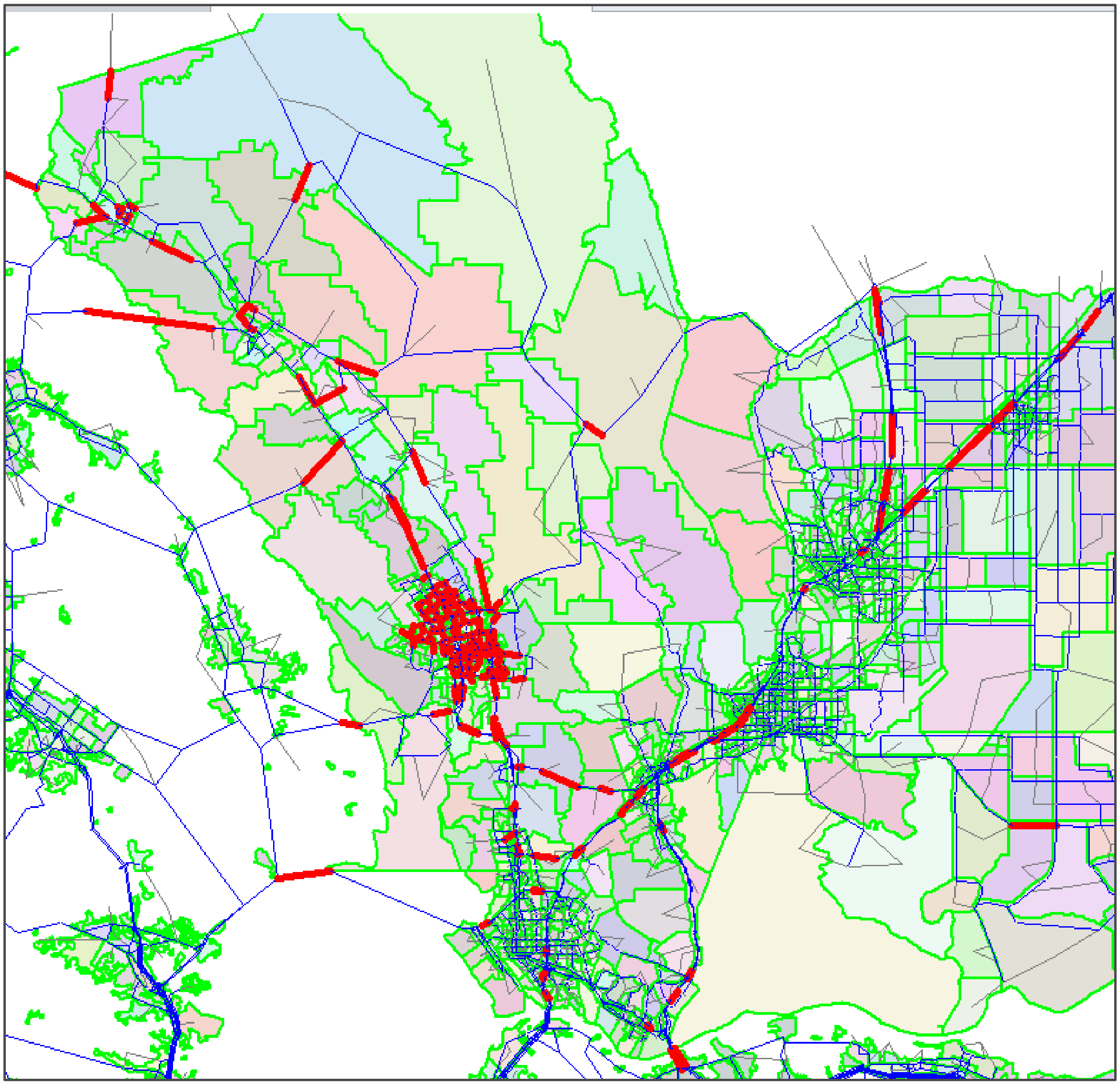


FIGURE 4 – COUNT LOCATIONS FOR DAILY VALIDATION

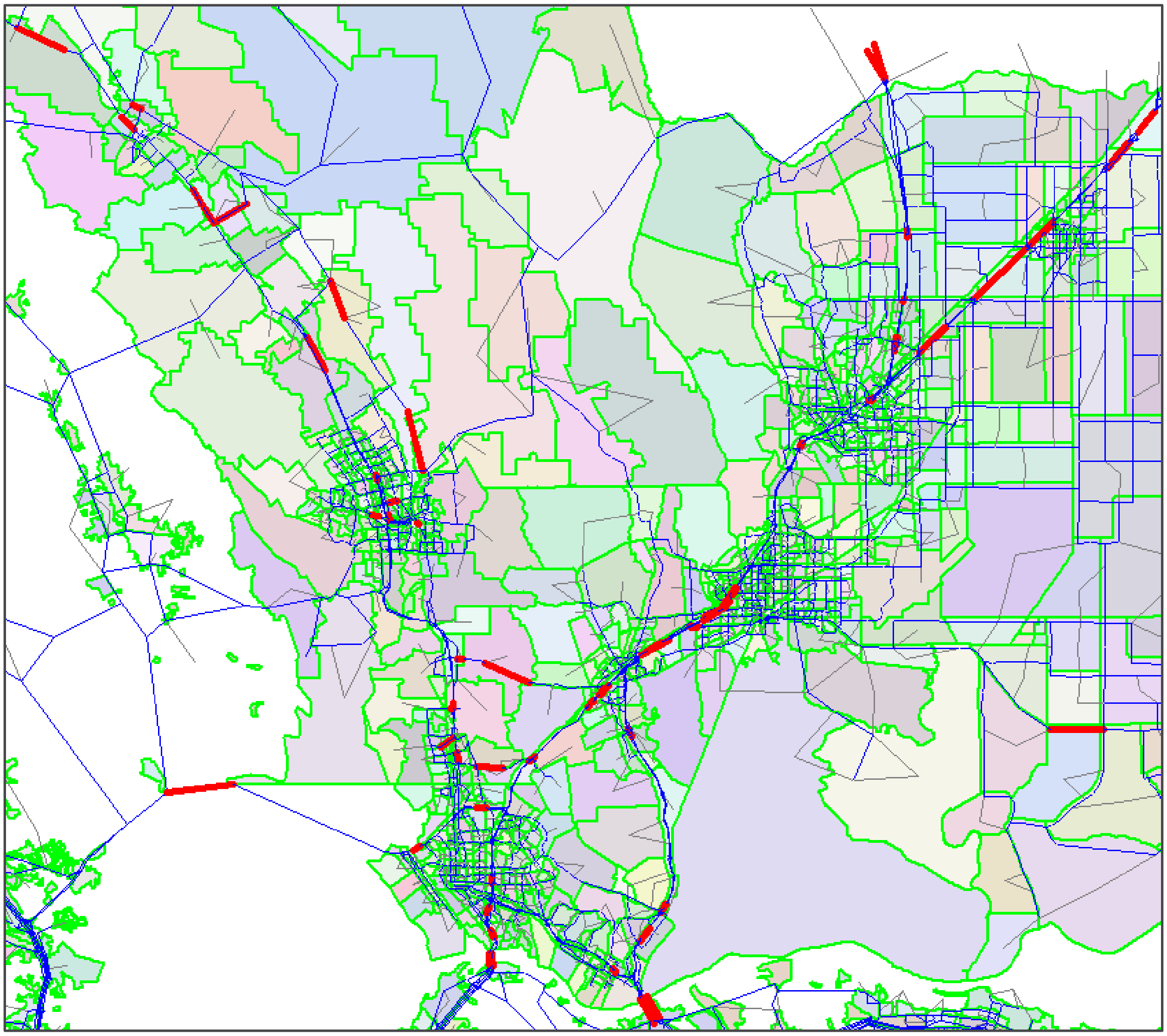


FIGURE 5 – COUNT LOCATIONS FOR PEAK HOUR VALIDATION

4.0 MODEL INPUTS

This chapter presents the details of SNABM input data. Besides model parameters and settings, the main inputs to SNABM are the synthetic population, zonal data, and network data. We implemented PopulationSim to generate synthetic population for both base and future year SNABM runs. The TAZs within Solano and Napa counties were split around major transit stops to improve the representation of transit access and egress. As a result, the land use data had to be updated for Napa and Solano counties. We reviewed the Solano and Napa county network and made appropriate refinements. The details of all these updates and refinements are presented in the following sub-sections.

4.1 POPULATIONSIM IMPLEMENTATION

The TM1.5 uses PopulationSim to generate synthetic population for both base and future years. RSG implemented PopulationSim for SNABM to generate synthetic populations for the base year (2015) and one future year (2040). PopulationSim is state-of-the-art population synthesizer software initially developed for the Oregon Department of Transportation (ODOT) and its partner agencies. PopulationSim makes several advancements over many existing population synthesizers. Implemented in the ActivitySim[7] framework, PopulationSim uses an entropy maximization-based list balancing approach to speed and improve convergence to marginal controls, allows the user to specify controls at multiple levels of geography, and uses linear programming-based techniques which eliminate stochasticity in output.

We used the MTC setup[8] for PopulationSim as the starting point for SNABM implementation. The Python-based setup includes scripts to create all inputs for a PopulationSim run, which includes a seed sample, marginal controls, and a geographic crosswalk file. The MTC setup generates residential and group quarters synthetic populations simultaneously. We used the 2013-17 5-year ACS PUMS as the seed sample. We updated the MTC setup to work with the SNABM geographic system, which has more TAZs within Napa and Solano counties.

All marginal controls are specified at the TAZ level and are as follows:

- Total number of households
- Number of persons in the household: 1, 2, 3, 4+
- Household income: 0-$30K, $30-$60K, $60-$100K, $100K+
- Number of workers in the household: 0, 1, 2, 3+

[7] https://activitysim.github.io/
[8] https://github.com/BayAreaMetro/populationsim

- Person age groups: 0-4, 5-19, 20-44, 45-64, 65+ years
- Group quarter unit type: university, military, other non-institutional

Marginal controls for all TAZs outside Napa and Solano counties were not changed. We generated the marginal controls for new SNABM TAZs within Napa and Solano counties by applying the distribution from the parent TM1.5 TAZ. The controls for TAZs within Napa and Solano counties were scaled to match the total households, persons, or group quarters population as per the updated SNABM land use data.

We also generated a future year synthetic population using the same set of controls. The controls were scaled to match the number of households and population for each TAZ from the 2040 land-use data file. We deployed the household size sub-model used by MTC to estimate the forecast year controls for SF-MTC Population Synthesizer[9]. The household size sub-model is a proportional allocation model used to convert the average household size to the number of households grouped by size. The average household size was computed from the future year number of households and population; and used to estimate the future household size distribution. The household size sub-model is graphically shown in Figure 6.

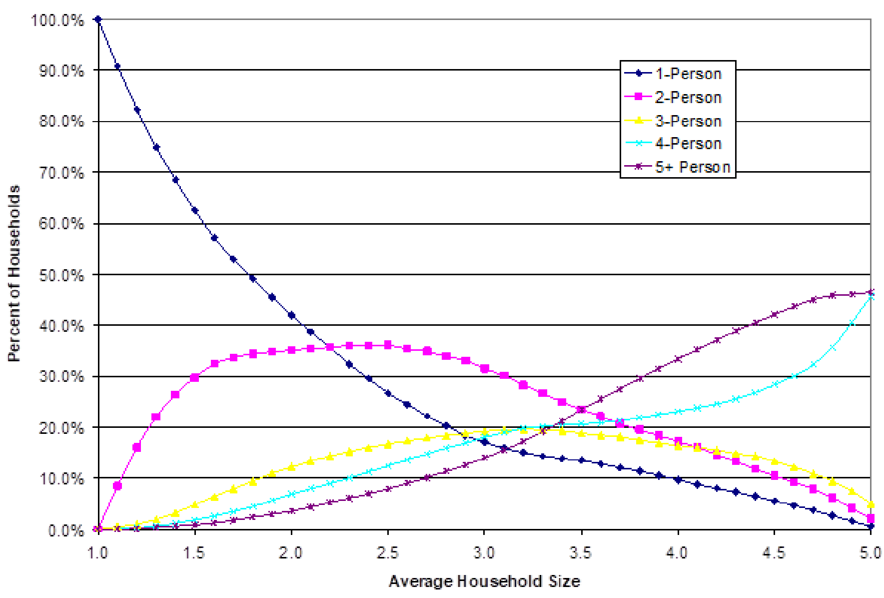


FIGURE 6 SF-MTC POPULATION SYNTHESIZER - HOUSEHOLD SIZE SUB-MODEL

[9] San Francisco County Transportation Authority (SFCTA). 2007. SF-MTC Popualtion Synthesizer Final Documentation. Excerpt from the update of the San Francisco chained activity modeling process (SF-CHAMP)

PopulationSim Validation

RSG validated the synthetic population to ensure that it matches marginal controls well. We developed validation scripts in R to report summary statistics and validation plots in addition to the reports generated by PopulationSim. The R script computes the following three statistics:

- the average percentage difference between the control totals and the synthesized totals,
- the standard deviation of the percentage difference – this measure informs us of how much dispersion from the average exists, and
- the percentage root mean square error (RMSE) - an indicator of the proximity of synthesized and control totals.

A chart visualizing these statistics is also produced. A form of dot and whisker plot is generated for each control where the dots are the mean percentage differences and horizontal bars are twice the STDEV or RMSE centered around zero.

PopulationSim validation was performed for the entire 9-county region and separately for Napa and Solano counties. Figure 7 presents the base year 2015 validation results for the TAZs within Napa and Solano counties. The validation results indicate that PopulationSim performs reasonably well overall for Napa and Solano counties. The mean percentage difference between each control and the generated synthetic population is close to zero across all controls. The total number of households is matched perfectly across all TAZs as required. There is a slightly higher standard deviation and a bit underestimation of population, as can be seen from the performance of population by age group controls. This might be a result of inconsistency between the household size distribution and the total population.

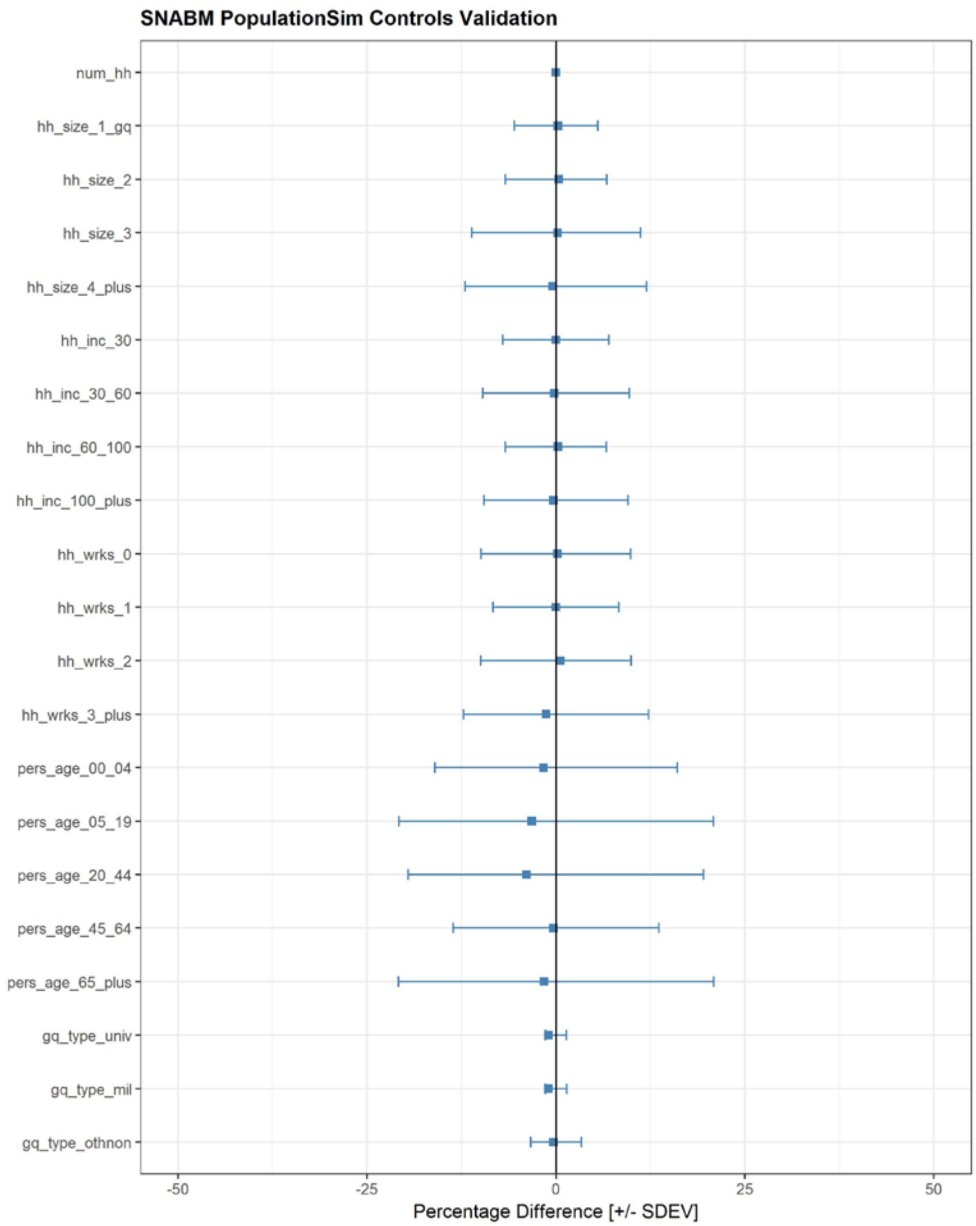


FIGURE 7 BASE YEAR (2015) POPULATIONSIM VALIDATION -NAPA/SOLANO COUNTIES

Figure 8 presents the Napa/Solano validation results for the year 2040. These results are very similar to the base year results with close to zero percentage difference across all the controls. The standard deviations are also low except for population and group quarters controls.

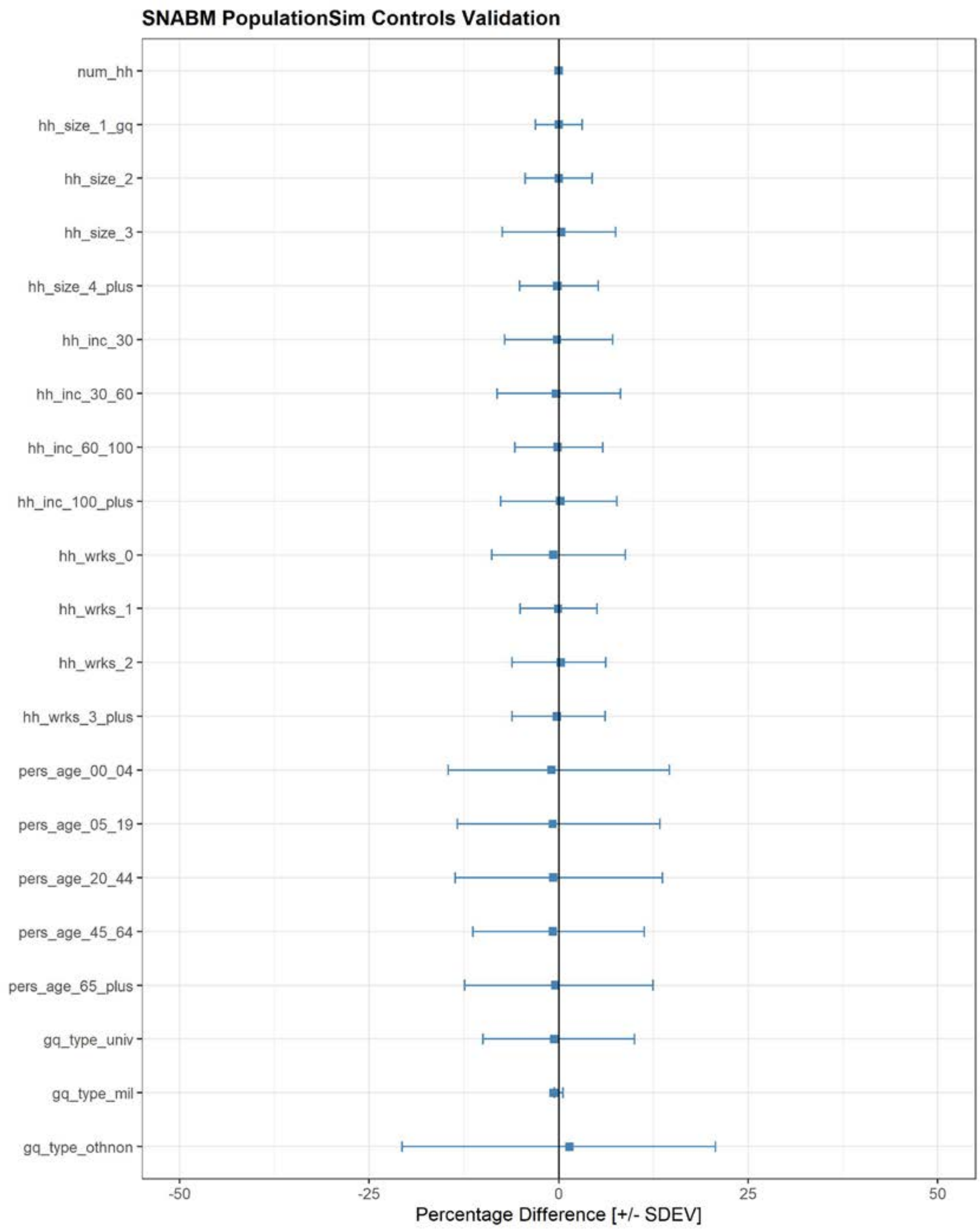


FIGURE 8 2040 POPULATIONSIM VALIDATION - NAPA/SOLANO COUNTIES

4.2 ZONAL DATA

Land use data is one of the primary inputs to every model and is a key component for trip generation.

MTC along with ABAG projects land use every 4 years for their Regional Transportation Plan. They use a variety of forecasting methods and consult with local jurisdictions to develop future year land use for 5 year increments. Once finalized and approved, the base year land use data is distributed for use travel demand models across the nine county bay area.

At the time of this study, the approved data was the Plan Bay Area 40 (PBA40) data which was used in the model. Data from MTC's 1454 TAZs were disaggregated to SNABM's 2340 TAZs using census block data distribution.

The land use data contains residential and employment data by TAZ. Residential data includes total households, population living in households, total population, employment residents, single family and multi family dwelling units. Employment data includes total employment by TAZ, retail trade employment, financial / professional services employment, health / educational / recreational service employment, agricultural / natural resources employment, manufacturing / wholesale trade / transportation employment, and other types of employment.

The table below summarizes the 2015 land use data used in SNABM.

Table 15– 2015 LAND USE DATA

County	Total Population	Total Households	Total Employment
Solano County	407,734	141,454	130,626
Napa County	137,993	49,212	70,789
Grand Total	545,727	190,666	201,415

RSG reviewed the base year land use data by creating maps and performing visual inspections against Google Maps Imagery. TAZ-level maps were created for total households, employment, and school enrollment. The distribution of households generally seemed reasonable as households are concentrated in residential areas. We compared the zones with no employment against Google Map Imagery and investigated to determine whether Google Maps showed any evidence of commercial land use. RSG found 10 zones with no employment for which Google Maps showed evidence for commercial land use. Figure 7 shows an example of such a zone which houses a medical center. TJKM performed a secondary review of these zones. Please note that this analysis was done using the latest version of Google Maps, however, the land use data represents year 2015 employment. For some of the reported zones, the commercial development shown in Google Maps happened after 2015, and therefore, was not included in the land use data. In other cases, the problem was fixed by moving the employment from neighboring zones to appropriately represent the employment in the problem zones.

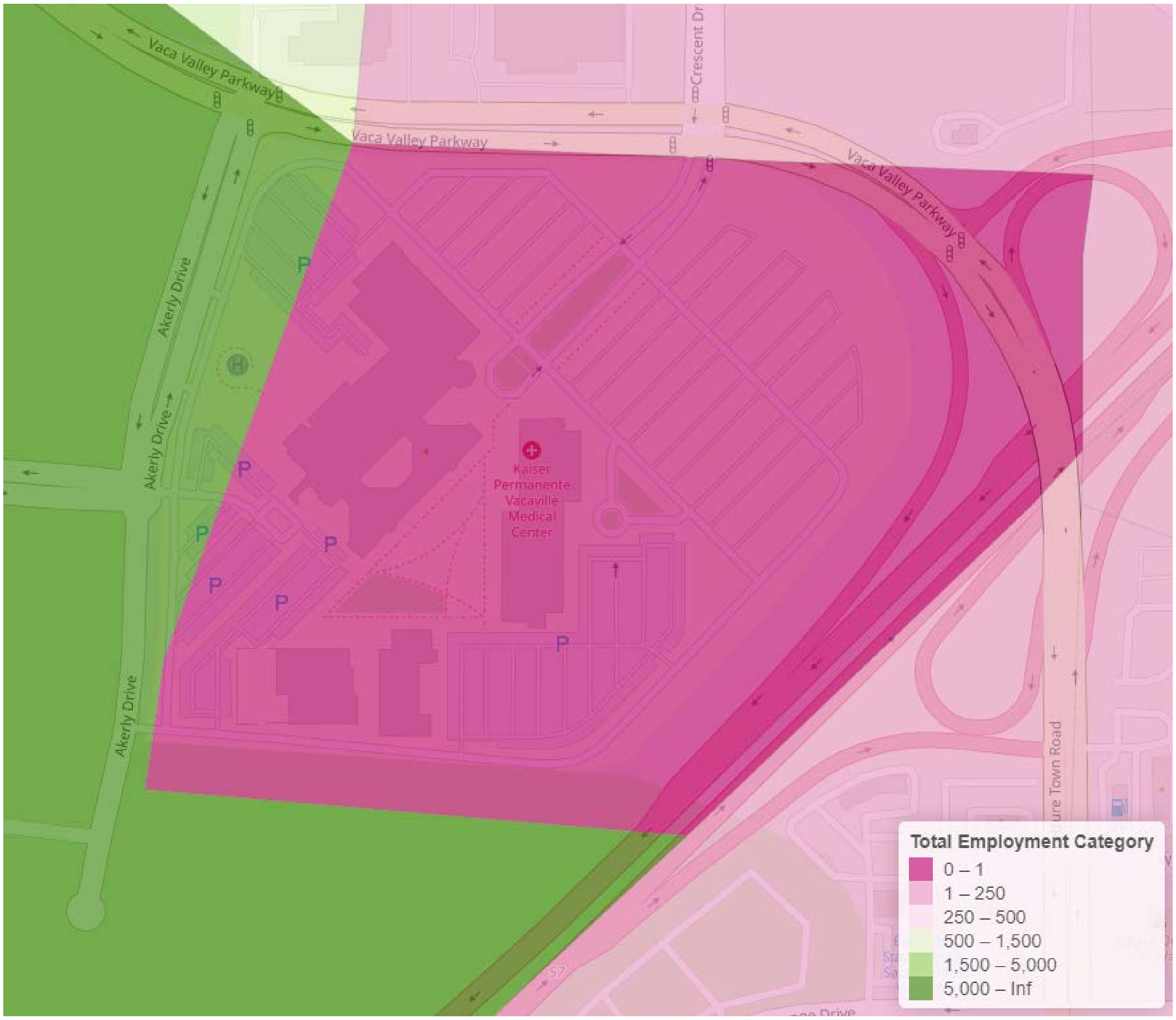


FIGURE 9 ZERO EMPLOYMENT ZONE (TAZ 852) INVESTIGATION

RSG created zone level maps for students enrolled in high school to ensure that students were distributed reasonably around each high school in Napa and Solano counties. Distribution of high school students around Rodriguez high school, Jesse Bethel high school, American Canyon high school, and Benecia high school indicated some problems in the data. Figure 10 shows the uneven distribution of student population around Vallejo high school. TJKM updated the high school student population data after comparing against the data from local jurisdictions.

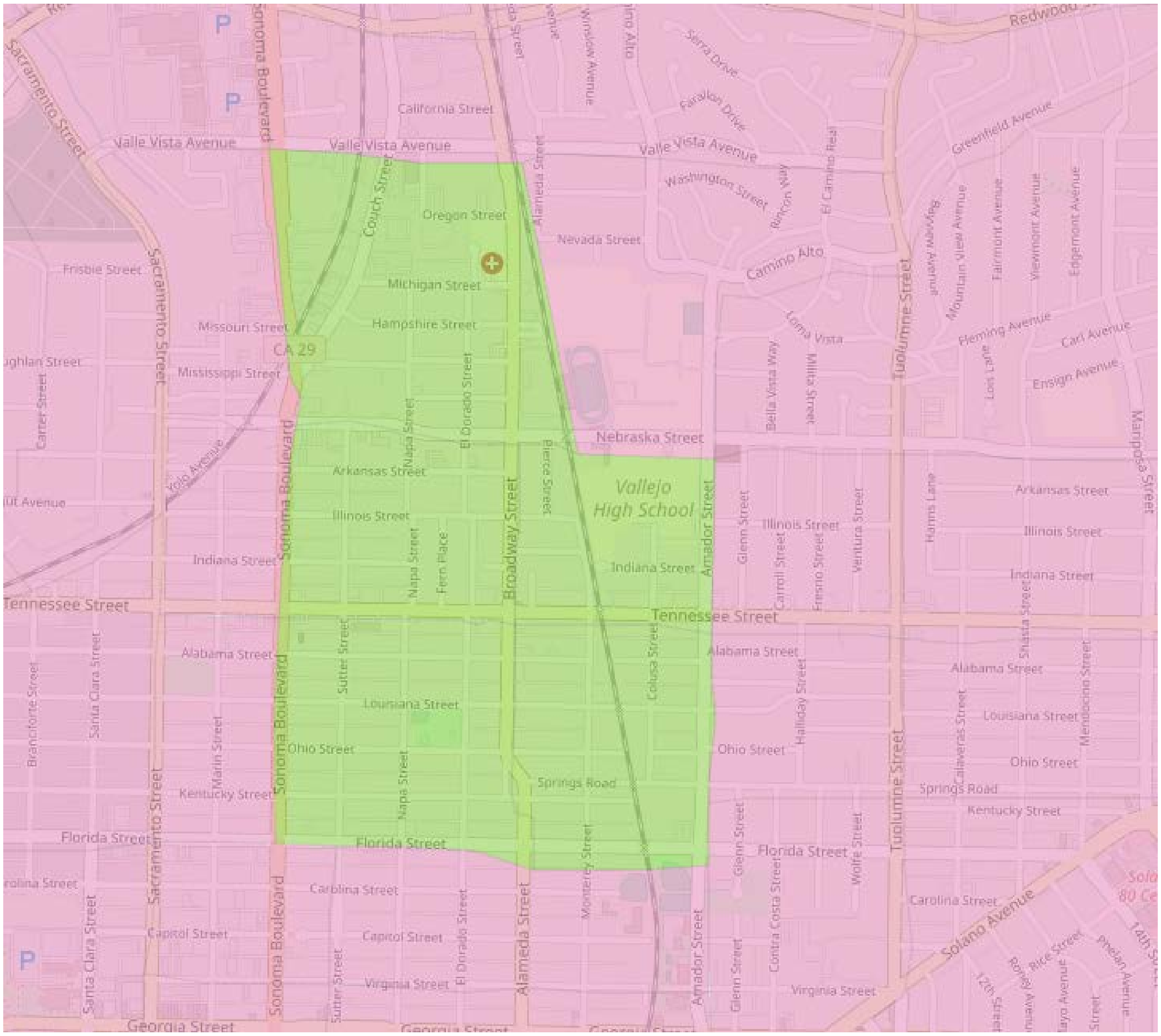


FIGURE 10 VALLEJO HIGH SCHOOL - STUDENT POPULATION DISTRIBUTION

RSG and TJKM worked together to remove discrepancies in the data set.

4.3 HIGHWAY NETWORK

TJKM used MTC's 2015 network as the starting point for this model development effort. The Solano and Napa county networks in the previous version of the SNABM model had no correspondence to the original MTC network nodes. They were prepared from scratch and included a lot more collectors and local streets. However, the rest of the 9-county area was identical to the MTC network. To update the network. relevant fields in the rest of Bay Area were updated using the latest MTC highway network and node correspondence. After comparison to the MTC network, missing links were added. For Napa and Solano counties, major roads (Freeways, Expressways, and Major Arterials) were compared against the MTC network and free-flow-speed, number of lanes, capacity, and HOV/Express lane identifiers were updated. Missing public transportation station nodes and park-and-ride (PNR) nodes were also added.

Once the base year network was created, TJKM sent out maps to the local jurisdictions for comments and updates. The local jurisdictions sent back comments regarding updates to number of lanes and speed limits within city limits. These changes were incorporated into the final network file used for the SNABM base year model run.

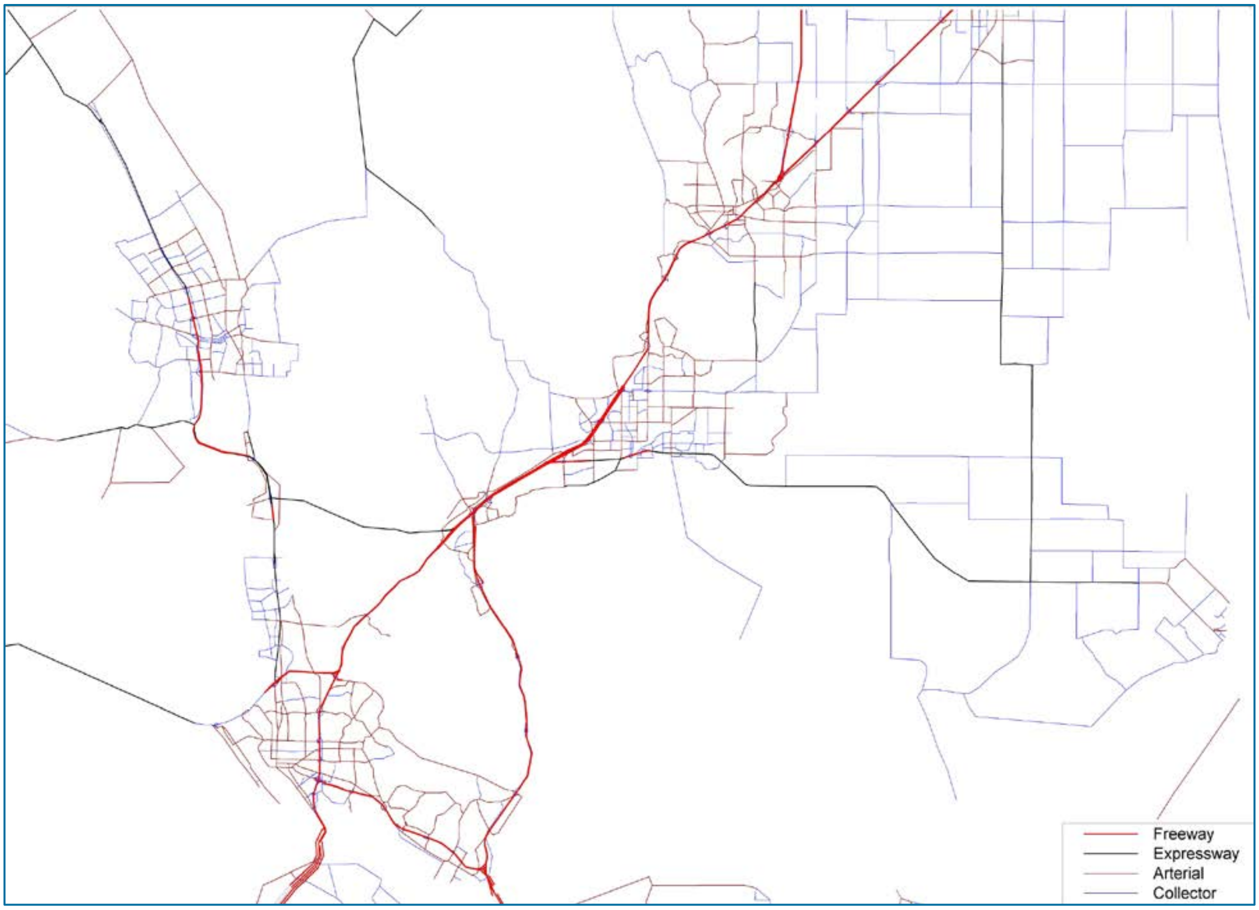


FIGURE 11 SNABM 2015 BASE YEAR NETWORK

RSG compared the highway network prepared by TJKM against Google Maps Imagery. The number of lanes and free-flow speed were plotted and reviewed for accuracy and reasonableness. In some cases, the coded number of lanes and speeds were compared against the road geometry and posted speed limits from Google Maps. There were only a few cases where the coded speed was significantly different from the posted speeds. TJKM reviewed the flagged links and performed appropriate updates. Generally, the coded number of lanes in Napa/Solano counties represent the observed number of lanes as per Google Maps Imagery. The facility type coding is also reasonable for Napa/Solano counties.

4.4 TRANSIT NETWORK

TJKM lead the development of transit network for SNABM. The Napa/Solano routes were recoded based on the latest MTC lines and the new SNABM highway network. Node IDs were updated for MTC lines that use Alameda and Contra Costa road network. No changes were made to rest of the Bay Area transit lines. The access, transfer and walk access supplement files were also updated.

After the initial transit network preparation, RSG compared the transit routes in the Napa/Solano network against the MTC network. The coded routes and headways were compared against online maps and schedules for Solano and Napa County operators. RSG accessed the cached[10] versions of the Solano/Napa transit operators to get the 2015 version of the route schedules and maps. Overall, the transit network for the rest of Bay Area is consistent with the MTC network. Based on this review, TJKM realigned few routes, added missing PNR nodes, and updated headways for some lines. The following figure is a map of the various transit lines in Solano and Napa counties.

[10] https://web.archive.org/

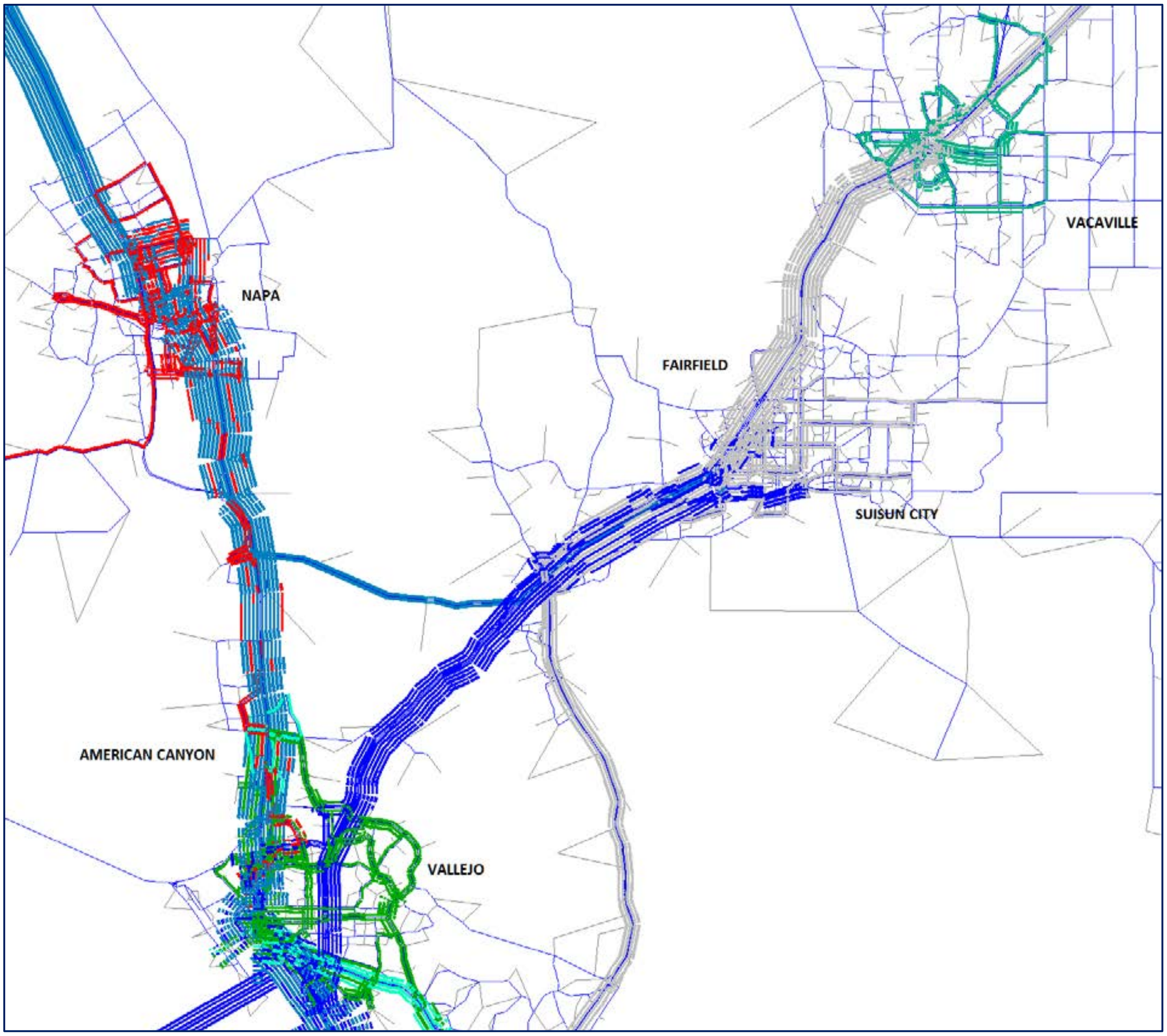


FIGURE 12 SNABM TRANSIT LINES

4.5 WALK ACCESS BUFFERS

As part of the development work, RSG updated the walk access buffers for Napa and Solano counties. The distance threshold for the short and long walk access buffers are 1/3rd of a mile and 2/3rd of a mile. The process starts with a node shape file containing points representing the transit stops nodes and transit stations in Napa/Solano counties. Next, transit walksheds are determined, which are polygons composed of circular areas around transit stop nodes. This polygon layer of circular buffers around transit nodes is overlaid over the zone polygon layer. The overlay between buffers and TAZs is then used to compute percent walk to transit value for each zone. These steps are performed for both distance thresholds to compute short and long walk access percent for each zone. To compute exclusive short and long access buffers (as per the TM1.5 specification), the short walk access percent is subtracted from the long walk access percent. Figure 9 and Figure 10 plots the short and long walk access percentages for all TAZs within Napa and Solano counties. The plots show a reasonable distribution of short and long transit walk percentages with higher percentages in zones with high transit stop density.

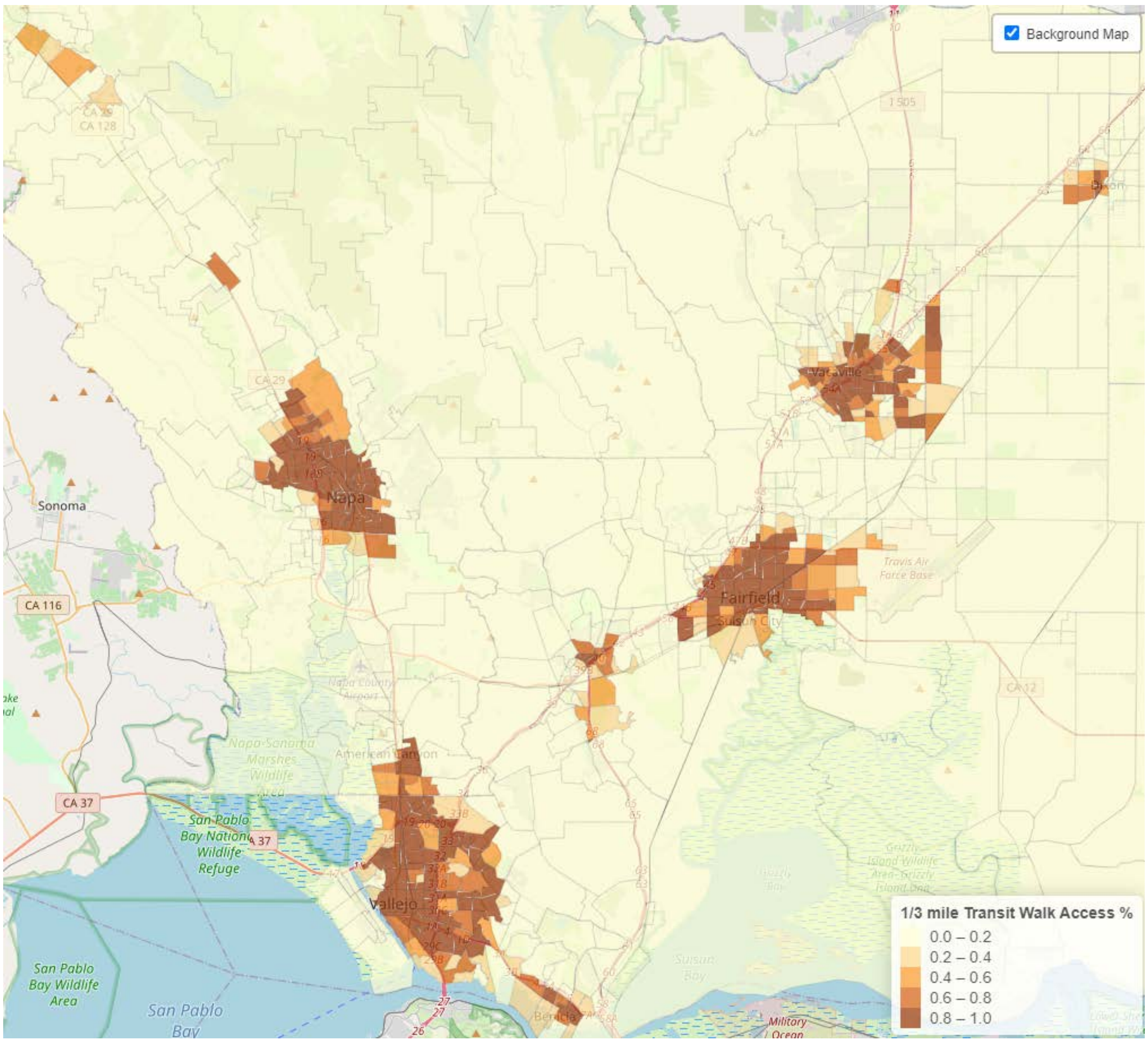


FIGURE 13 SHORT TRANSIT WALK ACCESS PERCENT

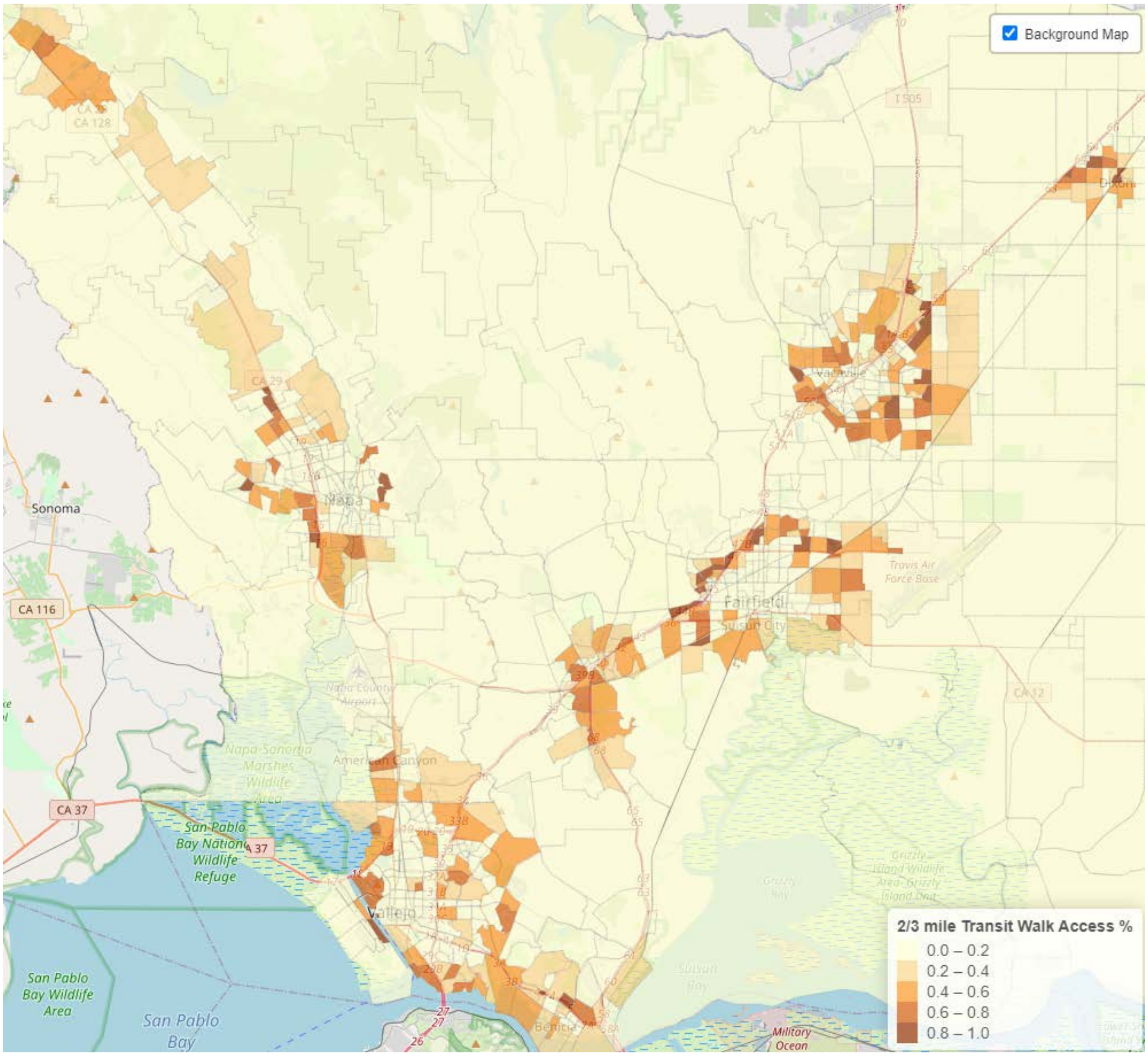


FIGURE 14 LONG TRANSIT WALK ACCESS PERCENT

4.6 NONRESIDENTIAL DEMAND

There are three nonresidential demand components in TM1.5 – airport, internal-external and trucks. TM1.5 uses fixed internal/external and air passenger demand for each forecast year. For SNABM implementation, the project team decided to use the fixed airport and internal-external demand matrices from the v1 model. The airport passenger trip tables sum to ~118K daily auto trips. The internal/external trip tables sum to ~658K daily auto trips. TM1.5 uses a simple 3-step (generation, distribution, and assignment) truck model to generate estimates of four types of commercial vehicles. The four vehicle types are very small (two-axle, four-tires), small (two-axle, six-tire), medium (3-axle), and large or combination trucks (four-or-more-axle).

During validation, the external trip tables were updated to match traffic counts at gateways. Also, local truck movement within Napa and Solano were increased to increase the % of truck traffic as compared to total traffic.

5.0 MODEL CALIBRATION

In model calibration, we iteratively adjust each component of the travel model until the model outputs reasonably fits the travel patterns revealed in the observed data. Each component of SNABM represents a travel related choice made by a household or a person for a long term or a daily decision. Most decisions in the CT-RAMP framework are modeled as a multinomial logit (MNL) model. Calibrating an MNL model involves updating the alternative-specific constants (ASCs) and other model parameters to move the aggregate model predictions in the desired direction. The mode choice models are nested logit models. These models are also calibrated by adjusting alternative-specific constants and/or other parameters.

Generally, the following steps are followed to calibrate a model within the CT-RAMP framework:

- Compare observed distributions against the predicted outputs. The Visualizer Tool and spreadsheets are used for comparative analysis. The following steps are implemented if the model distributions do not match the target distributions.
- ASC adjustments are calculated in a spreadsheet as follows for each alternative:

$$New\ ASC\ =\ old\ ASC\ +\ ln(\frac{Target\ Proportion}{Model\ Proportion})$$

- The adjustments computed in the previous step are entered into the appropriate model Utility Expression Calculator (UEC).
- The core CT-RAMP procedure reads in the model parameters from UEC spreadsheets.
- The model is run with the updated values.

The following sections summarizes the details of SNABM calibration and then present the final calibration results.

5.1 CALIBRATION OVERVIEW

Following refinements to land use, highway network, transit network, and nonresidential demand, we completed an initial model run with intelligent household sampling for Napa/Solano counties. TM1.5 has been calibrated by MTC for the entire Bay Area to match the observed distributions from CHTS. As a result, the model performed reasonably well at the regional level. However, TM1.5 was not calibrated at the county-level, therefore, there were some differences between the model outputs and CHTS data for Napa/Solano counties. Specifically, the tour and stop lengths were shorter compared to CHTS, transit share was being underestimated, and worker county flow was off for some county-county pairs compared to 2015 ACS flow of workers. The project team decided to calibrate SNABM in successive rounds of calibration. In each round of calibration, we summarized the entire model system using the HTML dashboard and developed validation summaries. Each round of calibration focused on different set of models. The models from the previous rounds of calibration were re-visited in a subsequent round for tighter calibration. We performed three rounds of calibration. Table 16 lists the sub-models that were adjusted during each round of calibration.

TABLE 16 SNABM CALIBRATION PLAN

CALIBRATION ROUND	SNABM MODEL COMPONENT
Round 1	o Mandatory destination choice – tour lengths o Mandatory destination choice – county-county flow of workers o Tour mode choice
Round 2	o Tighter mandatory destination choice o Non-mandatory destination choice o Internal-external demand o Individual non-mandatory tour frequency o Stop frequency o Stop location
Round 3	o Tour destination choice – county-county trip flow o Tour mode choice o Trip mode choice o Commercial Vehicle Model

The findings from each round of calibration were used to formulate the steps for the next round. The following sub-sections presents the details and key takeaways from each round of calibration.

Round 1

In the first round of calibration, we focused on issues identified from the analysis of the initial SNABM run. The mandatory tour destination choice was adjusted to match the CHTS tour lengths for Napa and Solano counties. We introduced county-level constants in the work location choice model to match the ACS county-to-county flow of workers. The tour mode choice model was adjusted to match the mode choice targets for Napa/Solano counties. After these adjustments, we performed a full model run and reviewed the highway validation results. Specifically, we reviewed the observed traffic counts validation at Napa/Solano County boundaries and important corridors such as I-80, Highway 29, and Highway 37. The countywide VMT from the model was compared against the VMT reported in Highway Performance Monitoring System (HPMS). Based on the analysis of the assignment results, we found that SNABM was underestimating the traffic movement on Carquinez and Benecia bridges. This was because of the underestimation of traffic flow to/from Contra Costa and Alameda counties to Solano and Napa counties. We found that this underprediction continued along I-80 all the way to Fairfield. Upon further investigation, we found that low volumes on bridges were exacerbated due to a few network errors that were missed during the initial network investigation, which we fixed during model calibration.

- the *TOLLCLASS* value for Carquinez bridge was missing.
- the network preparation procedures were not excluding the future express lane links from the MTC master network.

In addition to the above, the intra-county trips in Napa and Solano counties were being underestimated. The Solano County VMT matched the HPMS VMT closely, while the Napa County VMT was slightly underestimated.

Round 2

In this round of calibration, we focused on addressing the issues identified in the first round. To increase the traffic flow on bridges, we introduced county-level constants in tour destination choice model to increase traffic flow to Napa and Solano counties from the rest of Bay Area. The destination choice models were also adjusted to improve intra-county traffic flows in Napa and Solano counties. We also adjusted the stop frequency model to improve intra-county flows. At the end of Round 1 calibration, the countywide VMTs were not very far from the HPMS VMT target, while the average stop out-of-direction distances were being slightly overestimated. Therefore, we calibrated down the stop out-of-direction distance to balance the stop frequency increase. To investigate the underestimation of traffic volume on I-80, we reviewed the fixed internal/external demand matrix. The model was using the fixed internal/external demand from the v1 model. We compared this with the TM1.5 version of the fixed internal/external demand and found that the trips going towards Sacramento were much higher in the TM1.5 version. We updated the SNABM fixed internal/external trip table to match the TM1.5 number of trips going towards Sacramento. This helped in increasing the flow on I-80. With these adjustments in place, we performed a full model run and reviewed the highway validation results again. The second round of calibration adjustments resolved most of the inter-county flow issues. The countywide VMT was also matching reasonably with the HPMS VMT. However, the model was still underpredicting intra-county flows, especially within Napa County.

Next, the project team investigated the underprediction of intra-county flow in Napa county. We noted that the synthetic population for Napa/Solano counties has a lower share of fulltime workers compared to CHTS (see Figure 15 below).

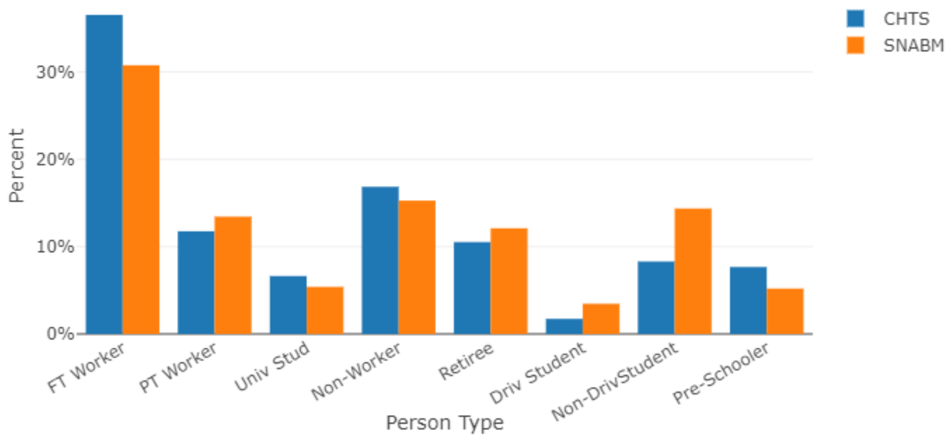


FIGURE 15 PERSON TYPE DISTRIBUTION - NAPA/SOLANO

A possible reason for this could be the underestimation of undocumented immigrant workers in Napa County. As per Public Policy Institute of California (PPIC) there are up to 16,000 undocumented migrant workers in Napa County[11]. Increasing the number of fulltime workers in the synthetic population was beyond the scope of this project. However, most of the migrant workers in Napa County are generally employed in agricultural sector, and their trips should be accounted by the Commercial Vehicle Model (CVM or Truck Model). Federal research indicates that freight trips account for 4-25% of total VMT[12]. In case of Napa county, the CVM VMT was <12% of the countywide VMT. The project team decided to increase the share of CVM VMT for Napa County. In this round, we also created volume-to-capacity maps to review highway assignment performance. We observed that traffic volumes on various downtown Napa County links were being underestimated. Since downtown Napa has a lot of restaurants and small shops, a significant portion of downtown trips would be contributed by visitors. The project team decided to boost the very small commercial vehicle trip rates for downtown Napa TAZs to account for missing visitor trips. In addition, our investigation of Napa County network revealed network loading issues related to centroid connector placement, which we fixed. The project team also cleaned the count database and discarded inconsistent counts.

---

[11] https://www.ppic.org/content/pubs/jtf/JTF_UndocumentedImmigrantsJTF.pdf
[12] Final Report: Accounting for Commercial Vehicles in Urban Transportation Models, prepared for Federal Highway Administration by Cambridge Systematics, Inc., March 2004, p. 5-1.

Round 3

The focus of this final round of calibration was to address the issues identified in the previous rounds and fine-tune the CTRAMP models to improve the model goodness-of-fit. We adjusted the CVM model to increase the share of CVM VMT for Napa County. We also boosted the very small commercial vehicle trips rates for downtown Napa TAZs. In addition, we fine-tuned the tour destination choice models until all volumes and VMTs were within acceptable ranges. After these adjustments, a full model run was completed. We compared the estimated boardings against observed ridership for all transit operators in Napa/Solano counties. The tour and trip mode choice models were adjusted to reasonably match the operator level transit ridership.

The next section presents the final calibration results. Following that the final model validation results are presented.

5.2 FINAL CALIBRATION RESULTS

This section presents the summaries from the HTML dashboard comparing the SNABM outputs to the observed data (mainly CHTS). The dashboard summaries label the observed data as "CHTS" and the model output as "SNABM". As mentioned earlier, we created two versions of dashboard, one for the 9-county Bay Area, and a second subset dashboard summarizing only the Napa/Solano County residents. Please note that the CHTS data has a very small sample of Napa and Solano County residents. Therefore, the CHTS summaries for the Napa and Solano County residents often turn out to be very lumpy. The Bay Area version of summaries are presented for models that were not adjusted for Napa/Solano counties.

Overview

Figure 16 compares the person type distribution between CHTS and SNABM synthetic population. Figure 17 compares the person type distribution only for Napa/Solano County residents. While at the regional level the person type distribution matches reasonably well, there are some difference for Napa/Solano counties. Specifically, there are fewer fulltime workers in Napa/Solano counties compared to CHTS. As mentioned earlier, this might be because of underrepresentation of undocumented migrant workers in the census data. Figure 18 compares the household size distribution and shows a reasonable match at the regional level.

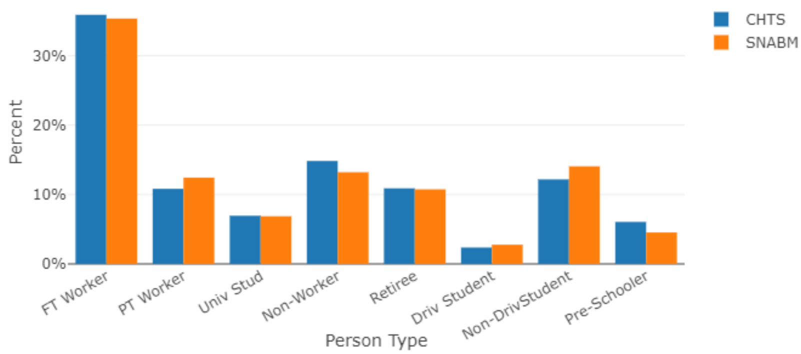


FIGURE 16 PERSON TYPE DISTRIBUTION - BAY AREA

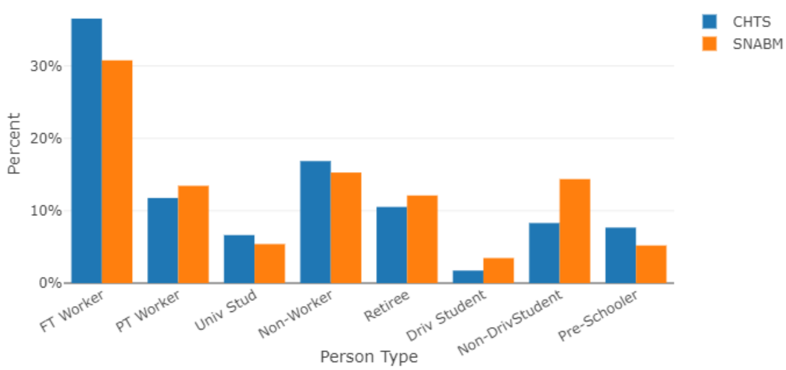


FIGURE 17 PERSON TYPE DISTRIBUTION - NAPA/SOLANO

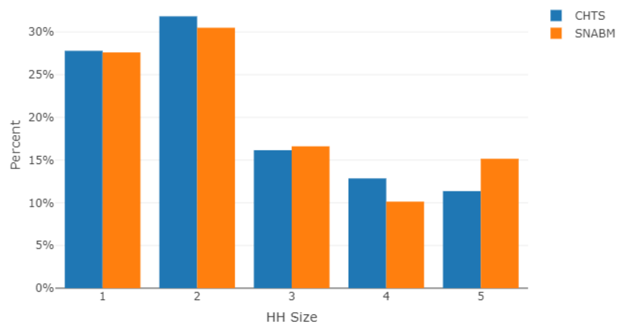


FIGURE 18 HOUSEHOLD SIZE DISTRIBUTION - BAY AREA

Figure 19 compares the aggregate travel and demographic statistics between CHTS and SNABM for Napa/Solano counties. SNABM has higher population compared to CHTS for Napa/Solano counties. Please note that CHTS was conducted in 2012-13 while the SNABM's base year is 2015.

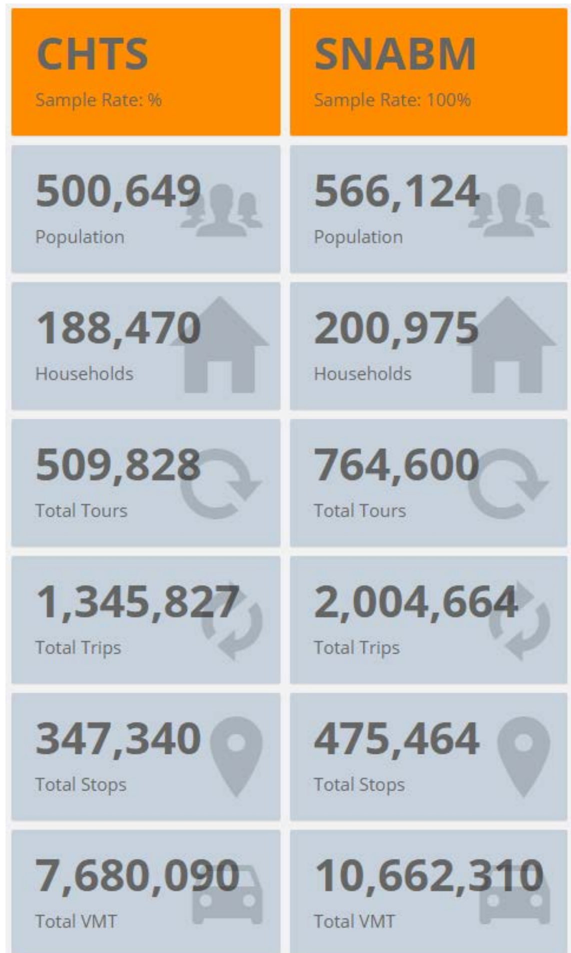


FIGURE 19 OVERVIEW AGGREGATE SUMMARIES - NAPA/SOLANO

Aggregate travel rates in Napa/Solano counties are presented in Figure 20. The CHTS data estimates that, on average, a Napa/Solano County resident makes 2.69 trips on a weekday (result of 1.02 tours and 0.69 stops). On the other hand, SNABM predicts that, on average, a Napa/Solano County resident makes 3.54 trips on a weekday (result of 1.35 tours and 0.84 stops). CHTS' underreporting of travel rates is a well-known issue, and therefore we did not calibrate the model to match the CHTS travel rates closely.

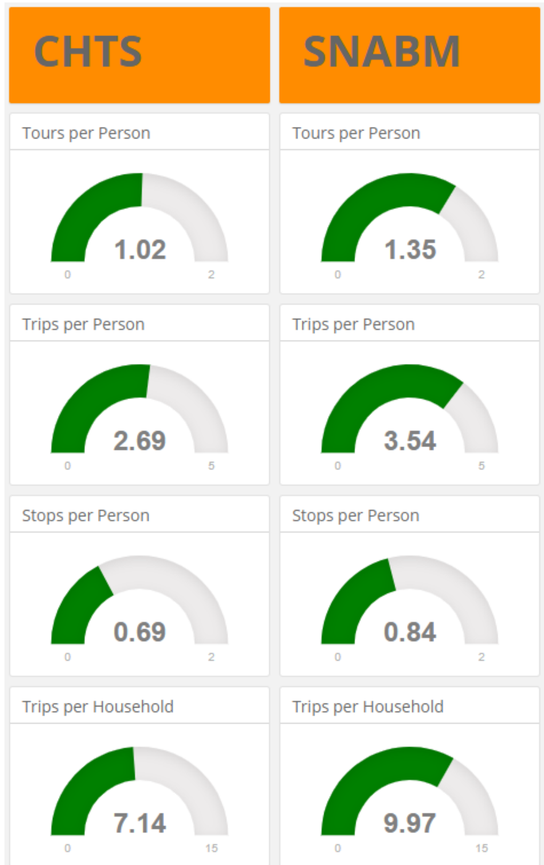


FIGURE 20 TRAVEL RATES - NAPA/SOLANO

Long-Term Travel Choices

An individual's daily travel is generally organized around long-term travel choices. Owning a car (auto ownership) and choosing a location for work and school (work and school location choice) are the primary long-term choices.

Auto ownership

The auto ownership model was calibrated to match the 5-year ACS 2013-17 sample. It can be observed on Figure 21 that the auto ownership model output matches the ACS distribution closely.

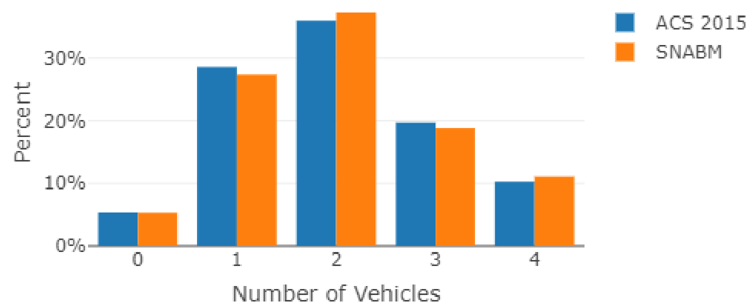


FIGURE 21 AUTO OWNERSHIP DISTRIBUTION - NAPA/SOLANO

Mandatory tour lengths

The tour length frequency distributions for work, university and school purposes are shown on Figure 22, Figure 23, and Figure 24, respectively. The CHTS distributions are generally lumpy due to lack of data in some of the distance bins, especially for the university purpose. The SNABM distributions are relatively smoother and generally follow the CHTS tour length frequency distribution profiles for the Bay Area.

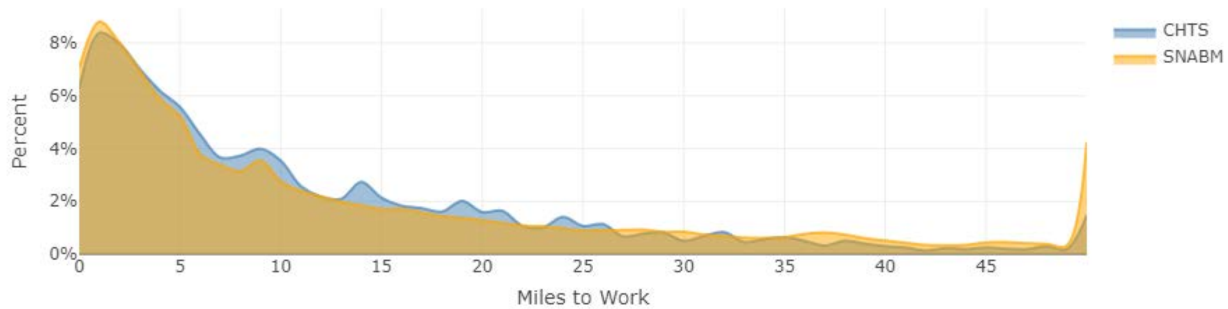


FIGURE 22 WORK TOUR LENGTH FREQUENCY DISTRIBUTION - BAY AREA

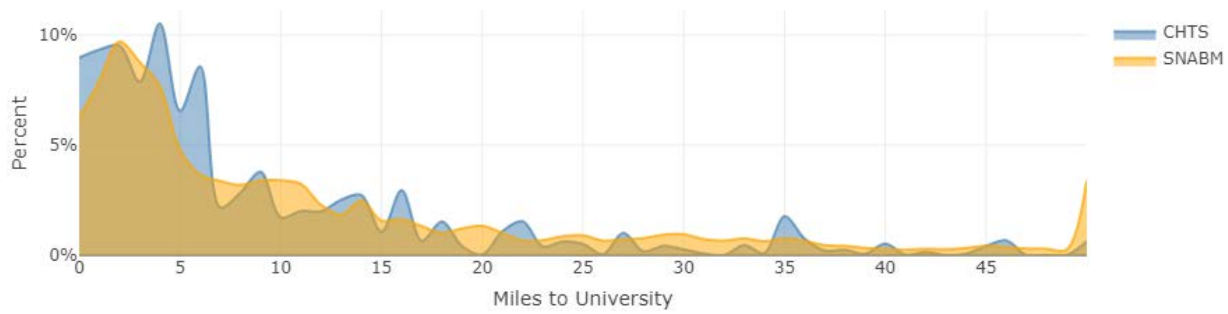


FIGURE 23 UNIVERSITY TOUR LENGTH FREQUENCY DISTRIBUTION - BAY AREA

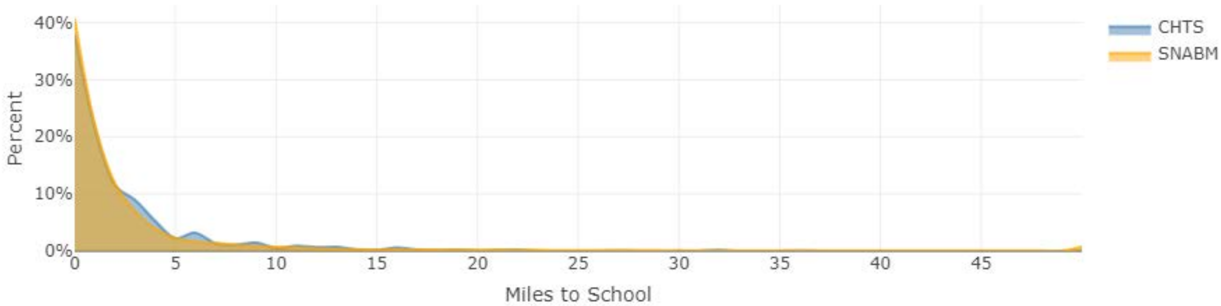


FIGURE 24 SCHOOL TOUR LENGTH FREQUENCY DISTRIBUTION - BAY AREA

Figure 25 presents a summary of average mandatory activity tour lengths by county and for the Bay Area. Bay Area workers on average travel 12.14 miles daily to get to work. Napa County workers have an average work tour length of 11.17 miles. Solano County workers have the longest average work tour length of 16.84 miles. SNABM replicates the work tour lengths at the regional level and matches the county-level results reasonably.

Students generally attend schools close to their home location and as per the CHTS data, on average travel 2.94 miles to attend school. University students usually must travel longer distance to attend university. However, most travel surveys under-sample university students, resulting in unreliable estimates for their travel patterns. Regardless, the model matches the target-CHTS school and university tour lengths reasonably.

	CHTS				SNABM		
Home District	Work	University	School	Home District	Work	University	School
Alameda	11.86	6.82	2.48	Alameda	12.26	8.00	2.79
Contra Costa	15.63	12.61	3.34	Contra Costa	15.88	14.48	2.98
Marin	12.57	14.46	2.27	Marin	13.05	15.29	3.62
Napa	11.17	7.57	2.87	Napa	11.07	8.86	2.70
San Francisco	9.11	5.66	2.71	San Francisco	8.65	3.04	2.00
San Mateo	10.84	10.28	2.62	San Mateo	11.99	8.65	3.12
Santa Clara	11.22	8.23	3.13	Santa Clara	11.10	7.53	2.89
Solano	16.84	16.61	4.25	Solano	16.69	16.94	4.14
Sonoma	12.87	13.80	3.20	Sonoma	10.99	11.49	3.47
Total	12.14	9.38	2.94	Total	12.11	8.90	2.98

FIGURE 25 AVERAGE MANDATORY TOUR LENGTHS

Commuter flows

Figure 26 and Figure 27 present the commuter flow between home county and workplace county. SNABM commuter flows are being compared to worker flow data from the 5-year ACS 2013-17 sample. Please note that ACS commuter flows do not include workers with workplace outside the 9-county region, resulting in fewer workers compared to SNABM. Except for few county pairs, the SNABM flows generally match the ACS flows reasonably.

ACS 2013-17

X	Alameda	Contra.Costa	Marin	Napa	San.Francisco	San.Mateo	Santa.Clara	Solano	Sonoma	Total
Alameda	476,304	39,864	4,593	289	104,454	38,154	74,676	1,923	852	741,109
Contra Costa	101,187	298,315	9,527	1,625	63,544	12,240	14,806	8,522	1,106	510,872
Marin	4,273	2,236	79,201	523	30,025	2,667	1,170	479	4,659	125,233
Napa	1,277	1,875	1,447	53,019	1,880	600	247	4,791	3,159	68,295
San Francisco	23,756	4,450	6,877	294	375,039	50,752	28,697	601	990	491,456
San Mateo	13,545	2,131	1,097	91	85,335	228,127	62,689	244	306	393,565
Santa Clara	40,007	2,898	329	73	19,334	47,323	810,443	277	519	921,203
Solano	10,768	20,676	5,146	12,530	10,178	2,479	2,468	113,121	3,053	180,419
Sonoma	2,048	965	16,321	4,575	7,606	1,159	1,029	759	207,142	241,604
Total	673,165	373,410	124,538	73,019	697,395	383,501	996,225	130,717	221,786	3,673,756

FIGURE 26 DISTRICT - DISTRICT FLOW OF WORKERS (ACS)

SNABM

X	Alameda	Contra.Costa	Marin	Napa	San.Francisco	San.Mateo	Santa.Clara	Solano	Sonoma	Total
Alameda	541,588	66,988	6,470	319	80,451	43,734	86,617	2,186	223	828,575
Contra Costa	125,444	295,685	10,646	2,454	48,817	12,947	16,659	19,485	619	532,757
Marin	7,832	6,218	77,894	269	29,843	3,236	264	75	5,483	131,114
Napa	1,223	2,580	999	57,239	1,551	793	37	4,206	2,478	71,107
San Francisco	30,142	6,254	9,687	429	368,252	56,150	36,265	1,482	242	508,903
San Mateo	22,957	3,290	3,967	184	72,901	227,806	77,378	942	217	409,642
Santa Clara	80,643	4,092	298	63	20,564	58,263	819,116	895	39	983,974
Solano	14,052	18,340	6,863	13,579	10,640	3,272	2,234	127,750	1,957	198,688
Sonoma	1,002	1,557	16,886	2,666	2,542	338	8	570	226,883	252,453
Total	824,884	405,004	133,710	77,202	635,562	406,538	1,038,579	157,592	238,141	3,917,212

FIGURE 27 DISTRICT - DISTRICT FLOW OF WORKERS - SNABM

Tour Level Choices

After the long-term choices, the model predicts the daily travel choices. The daily choice starts with daily activity pattern and tour generation. Only after the tour generation are intermediate stops along the tour determined and individual trips generated. Therefore, ascertaining a reasonable match of tour generation is necessary before looking into trip-level comparisons.

Daily activity pattern

Figure 28 compares the DAP distribution across all person types between CHTS and SNABM. CHTS has a lower share of persons with mandatory DAP which is in line with the lower estimated tour rate in the CHTS dataset. Since CHTS underestimates the tour rate, we did not adjust the DAP pattern under this effort.

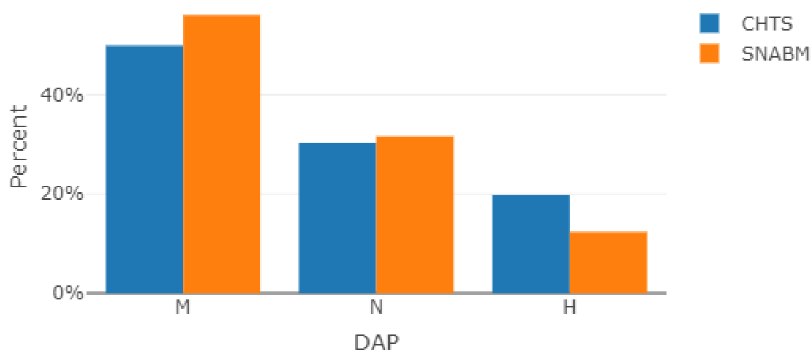


FIGURE 28 DAILY ACTIVITY PATTERN - TOTAL [BAY AREA]

Mandatory tour frequency

Figure 29 compares the results of the mandatory tour frequency model against the CHTS distribution. This model was also not calibrated under this calibration effort. However, the model matches the CHTS distribution very well for Napa/Solano County residents.

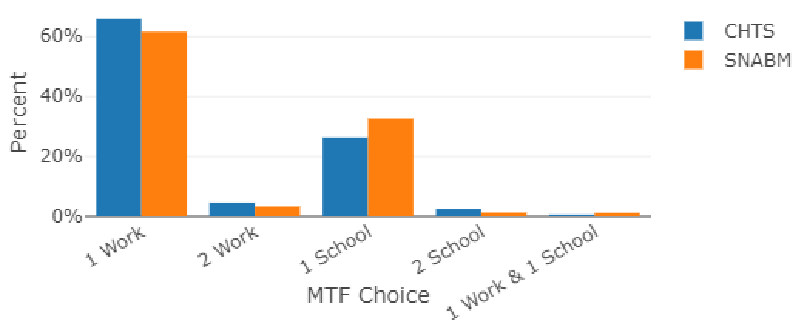


FIGURE 29 MANDATORY TOUR FREQUENCY - TOTAL [NAPA/SOLANO]

Individual non-mandatory tour frequency

Figure 30 compares the results of the individual non-mandatory tour frequency model against the CHTS distribution. It should be noted that joint tour participation was counted as a non-mandatory tour for each participant. Since CHTS underestimates tour rates, we did not adjust this model to match the CHTS distribution.

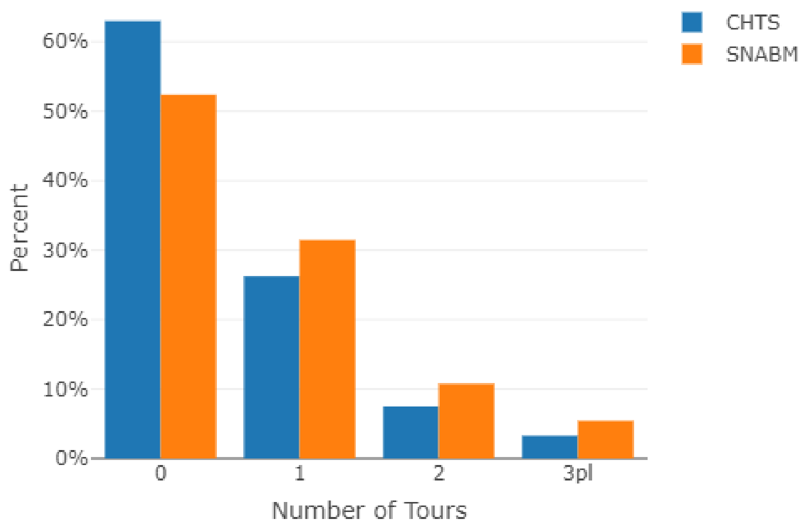


FIGURE 30 PERSONS BY INDIVIDUAL NON-MANDATORY TOURS - TOTAL [NAPA/SOLANO]

Tour rates (active persons)

Active persons are defined as persons who performed at least one out-of-home activity on a given day (activity pattern is either mandatory or non-mandatory). As shown on Figure 31, the model tour rates for active persons are generally close to the CHTS tour rates. SNABM tour rates for active retirees and driving age student is higher than CHTS rates, but we did not adjust the tour rates due to overall underestimation of tour rates in CHTS.

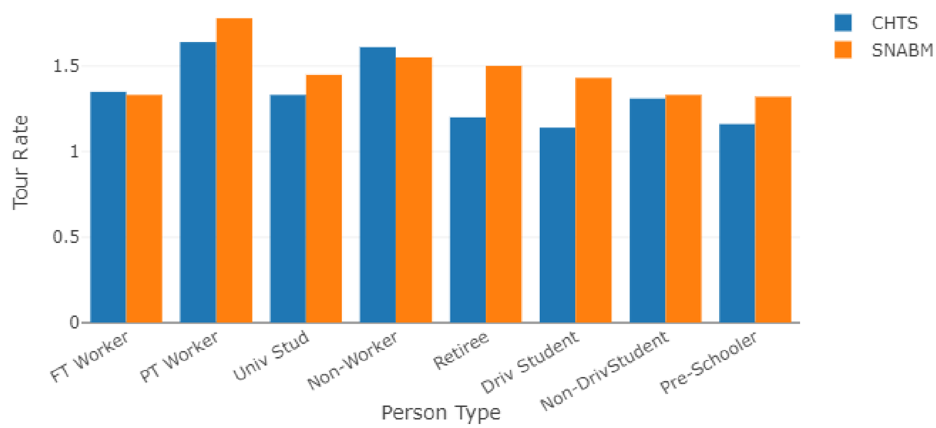


FIGURE 31 TOTAL TOUR RATE (ACTIVE PERSONS) - NAPA/SOLANO

Joint tour frequency

Figure 32 compares the SNABM joint tour frequency distribution with the CHTS distribution. The joint tour frequency model was also not calibrated under this effort. The model generally follows the CHTS distribution for households with single joint tours but underestimates multiple joint tours.

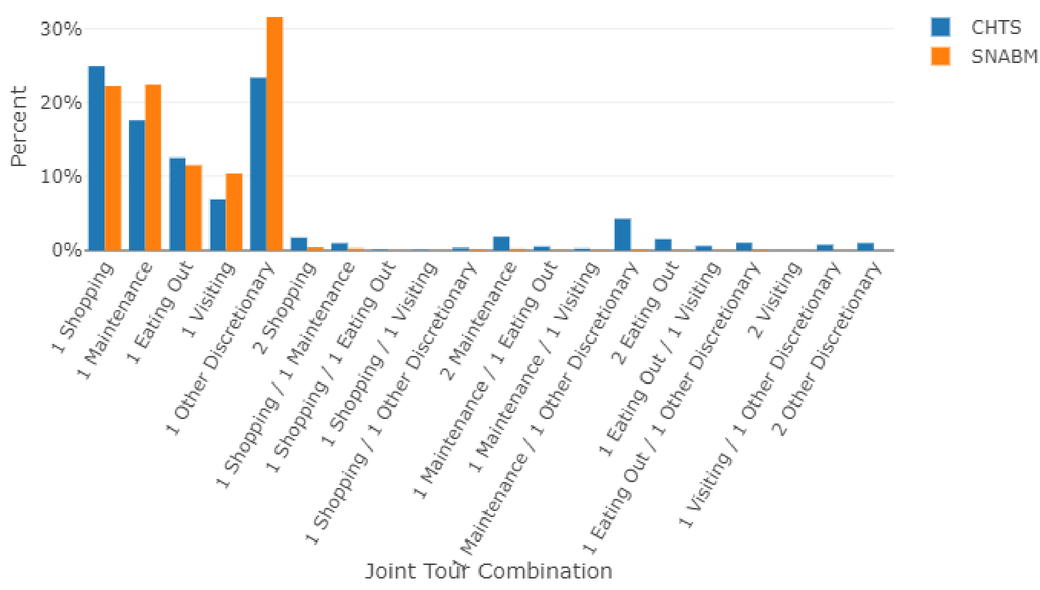


FIGURE 32 HOUSEHOLDS BY JOINT TOUR COMBINATION - BAY AREA

Non-mandatory tour destination

Comparison of tour lengths is a good way to verify if people are travelling far enough to participate in activities. Overall, the tour length frequency distribution of non-mandatory tours in SNABM matches well with CHTS data as shown on Figure 33.

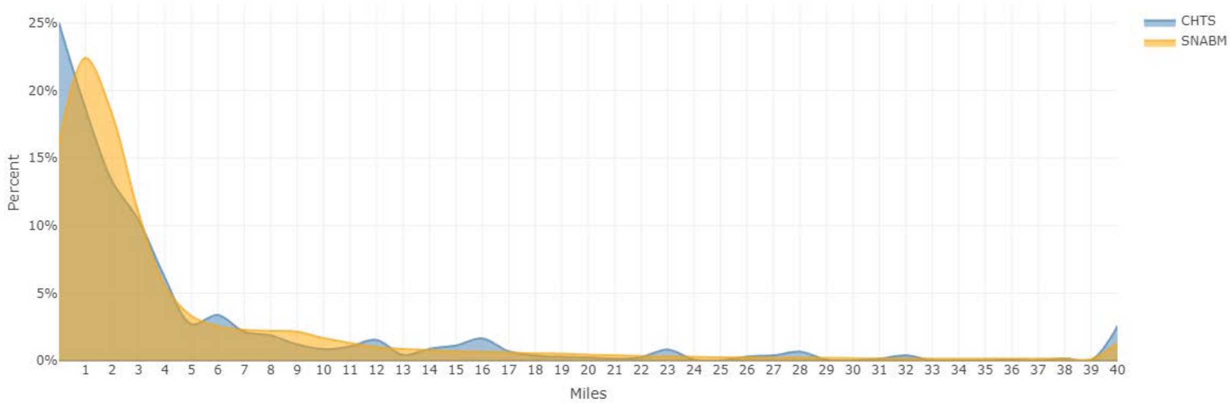


FIGURE 33 NON-MANDATORY TOUR LENGTH DISTRIBUTION - NAPA/SOLANO

The average non-mandatory tour length in CHTS is around 5.85 miles for Napa/Solano County residents. The average non-mandatory tour length in SNABM is around 5.46 miles (see Figure 34). Due to small sample, the CHTS tour length distribution by purpose is very lumpy for Napa/Solano counties. Therefore, we did not attempt to match the observed tour lengths by purpose.

Purpose	CHTS	SNABM
Escorting	2.46	3.25
Indi-Maintenance	5.96	5.67
Indi-Discretionary	8.24	7.07
Joint-Maintenance	11.04	6.19
Joint-Discretionary	3.78	7.78
At-Work	5.07	3.51
Total	5.85	5.46

FIGURE 34 AVERAGE NON-MANDATORY TOUR LENGTHS (MILES) - NAPA/SOLANO

Tour time-of-day

Figure 35 and Figure 36 compare the tour departure and arrival profiles between SNABM and CHTS for Napa/Solano counites. In general, SNABM and CHTS arrival and departure profiles match reasonably well. Figure 37 compares the SNABM and CHTS tour duration profiles. Again, SNABM and CHTS duration profiles match reasonably well.

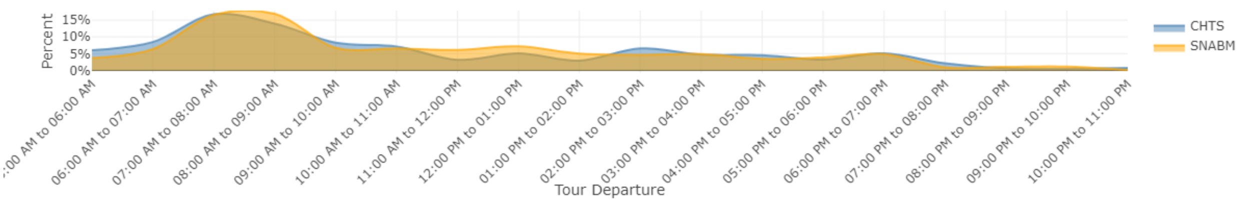


FIGURE 35 TOUR DEPARTURE PROFILE - NAPA/SOLANO

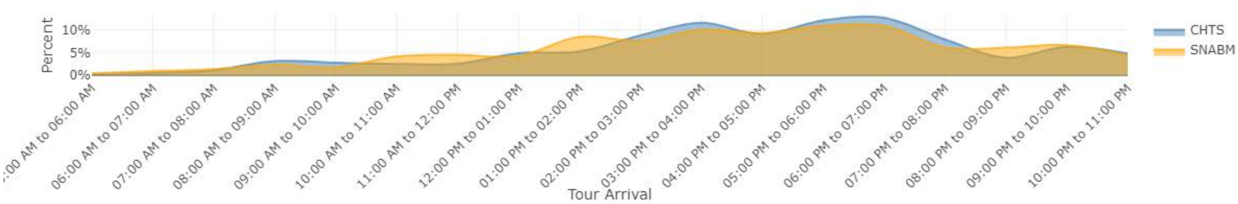


FIGURE 36 TOUR ARRIVAL PROFILE - NAPA/SOLANO

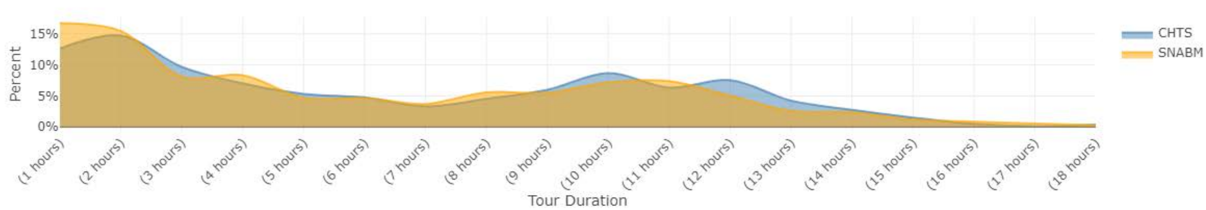


FIGURE 37 TOUR DURATION PROFILE - NAPA/SOLANO

Figure 38 and Figure 39 show the aggregate tour departure and arrival distributions for Napa/Solano counties. It can be observed that SNABM generally matches the CHTS distributions reasonably well by aggregate time periods as well.

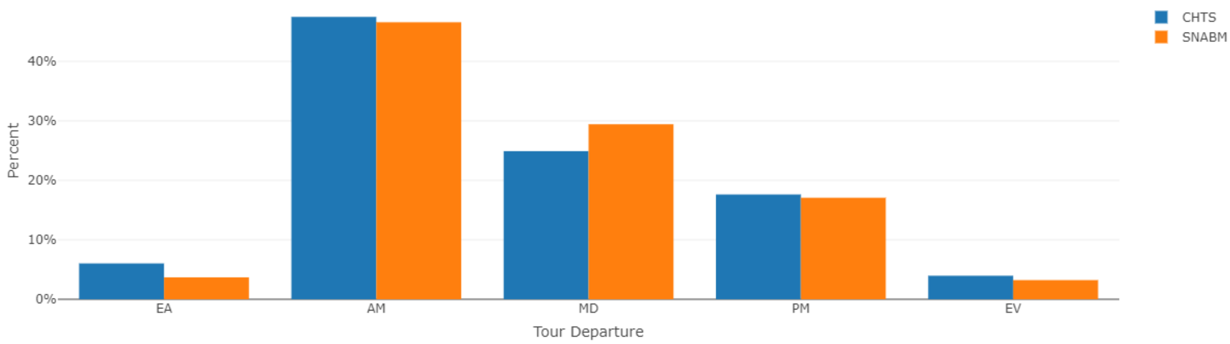


FIGURE 38 TOUR AGGREGATE DEPARTURE - NAPA/SOLANO

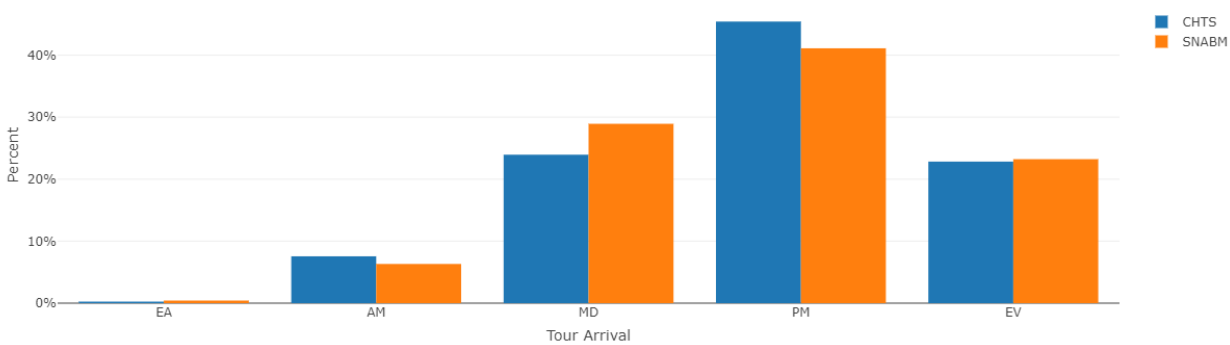


FIGURE 39 TOUR AGGREGATE ARRIVAL - NAPA/SOLANO

Tour mode

Tour mode is defined as the main mode of travel used to get from the origin to the primary destination of the tour and back. As mentioned earlier, the CHTS mode choice targets are supplemented with the OBS data. The augmented mode choice targets need to be scaled again based on the number of tours produced by the model. Scaling is done to produce the same number of transit trips irrespective of the number of person tours produced in the model. Generally, the tour mode choice calibration aims to adjust the mode choice model so that the distribution of tours by mode matches the observed share. Therefore, tour mode choice adjustments are made based on mode shares. As transit tour targets are calculated directly from the transit trips reported in OBS, the model needs to be calibrated to the absolute number of transit tours and trips. To achieve this, the survey targets are further adjusted by keeping the transit tours constant but scaling other modes to match the total tours in the model.

The summaries presented in this section shows the scaled calibration targets from the final calibration run. Figure 40, Figure 41, Figure 42 and Figure 43 present the tour mode choice distribution comparison by auto sufficiency groups for households in Bay Area. SNABM matches the observed shares closely at regional level. We adjusted the SNABM tour mode choice model to match the ridership targets of Napa/Solano County transit operators. Figure 44 presents the tour mode choice comparisons for all households in Napa and Solano counties. Transit ridership comparison for Napa/Solano County transit operators is presented in the Transit Validation section.

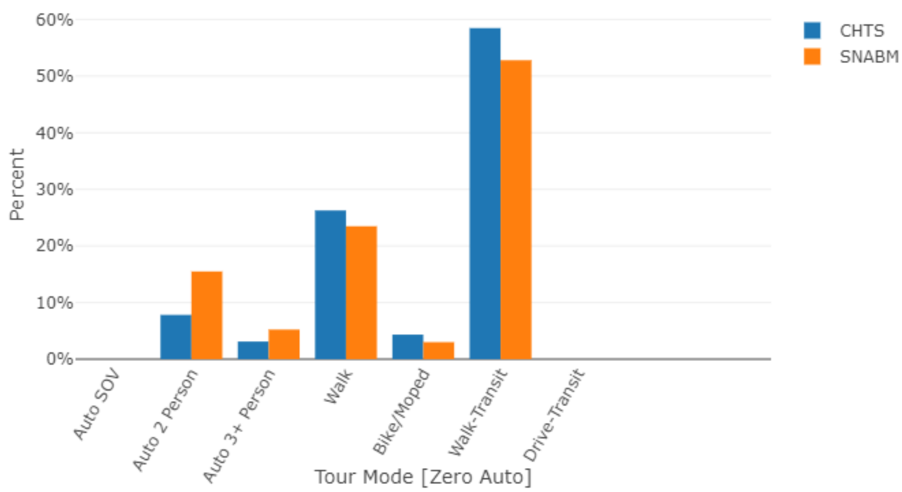


FIGURE 40 TOUR MODE [ZERO AUTO] - BAY AREA

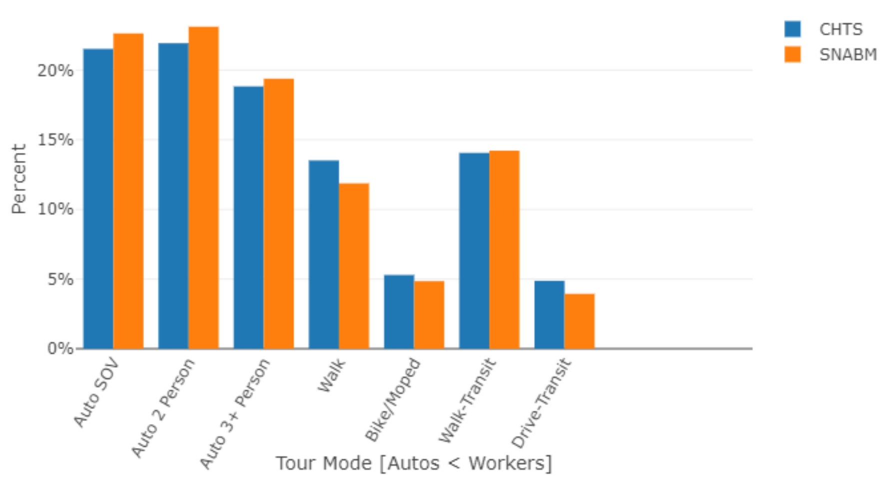


FIGURE 41 TOUR MODE [AUTO DEFICIENT] - BAY AREA

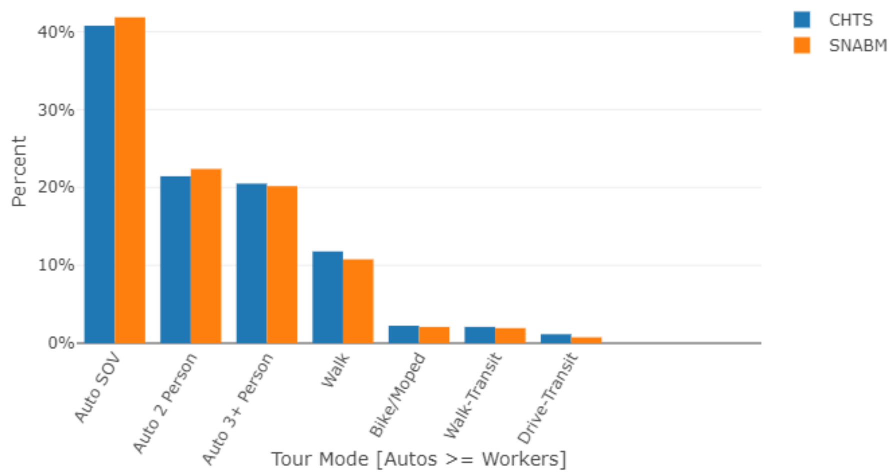


FIGURE 42 TOUR MODE [AUTO SUFFICIENT] - BAY AREA

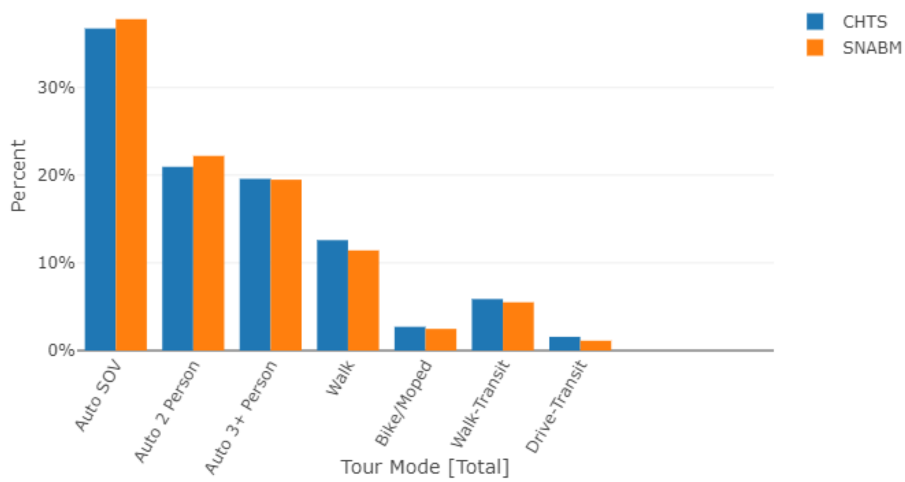


FIGURE 43 TOUR MODE [TOTAL] - BAY AREA

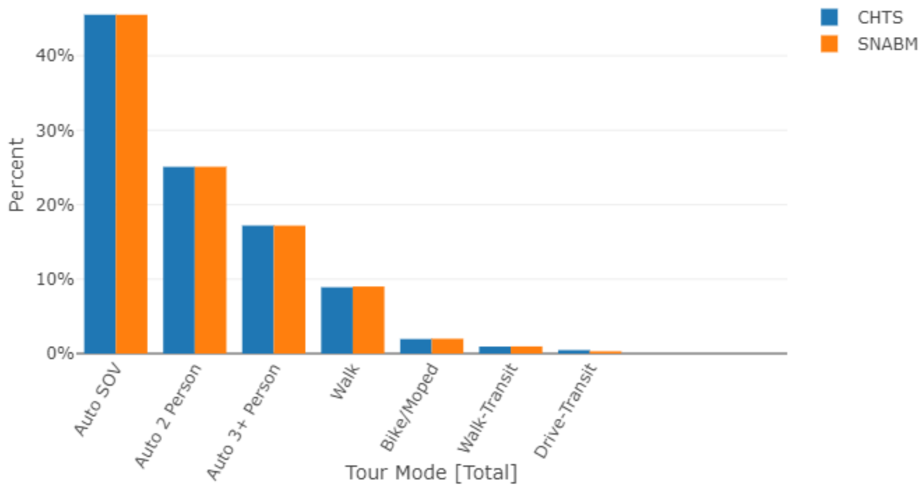


FIGURE 44 TOUR MODE [TOTAL] - NAPA/SOLANO

Trip Level Choices

This section examines trip generation in SNABM. The number of trips is a function of tours and the intermediate stops made along those tours. The four sets of summaries include stop frequency, stop location, time of day, and trip mode to examine the magnitude, spatial distribution, temporal distribution, and mode preference of travel at an activity level, respectively.

Stop frequency

Figure 45 presents the percentage of tours by frequency of intermediate stops. The model was not adjusted to match the CHTS distribution due to underestimation of stops in CHTS.

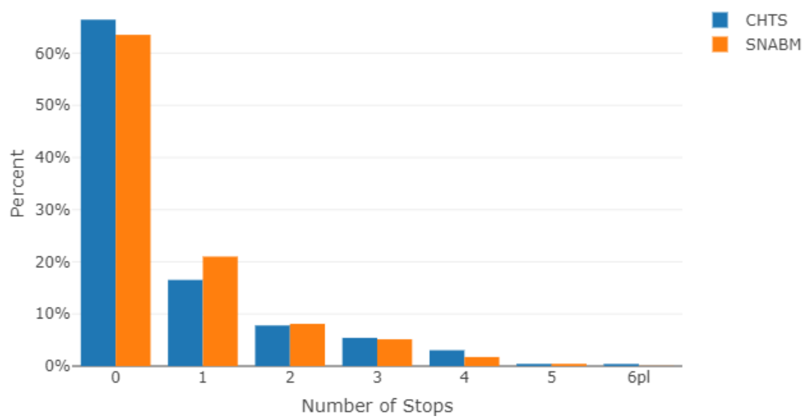


FIGURE 45 STOP FREQUENCY - NAPA/SOLANO

Stop location

The stop location model is compared against the CHTS data by comparing the out of direction distance of stop locations. Figure 46 shows the distribution comparison across all stops for Napa/Solano County residents while Figure 47 shows the average out of direction distances by tour purpose. In general, the model matches the total observed distribution and averages. There are some differences at purpose level, but the observed distribution is lumpy at purpose level due to smaller CHTS sample size for Napa/Solano counties.

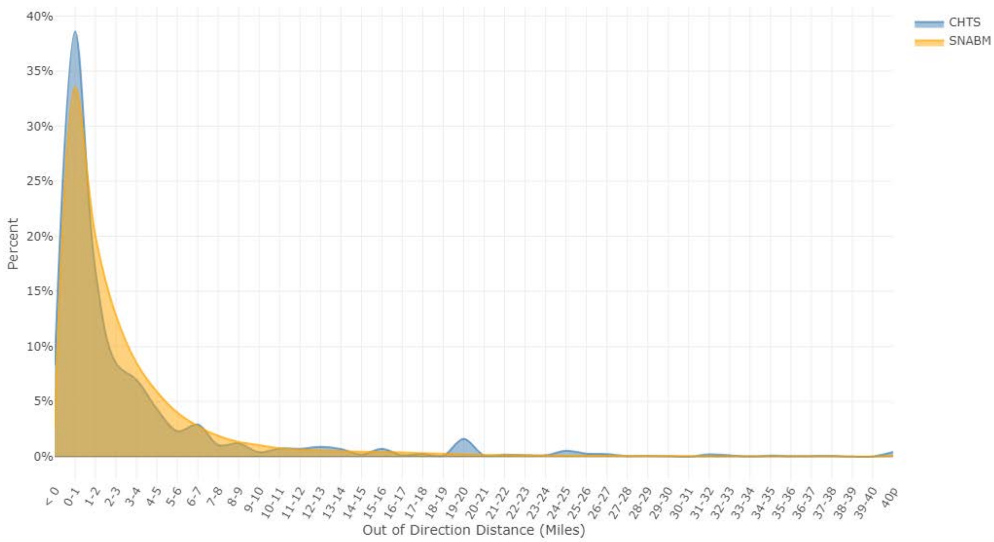


FIGURE 46 STOP LOCATION OUT OF DIRECTION DISTANCE (MILES) - NAPA/SOLANO

Tour_Purpose	CHTS	SNABM
Work	3.45	3.15
University	0.69	3.87
School	4.20	3.01
Escorting	2.35	2.57
Indi-Maintenance	3.24	2.96
Indi-Discretionary	3.11	2.72
Joint-Maintenance	3.06	2.84
Joint-Discretionary	3.83	3.01
At-Work	3.77	2.38
Total	2.98	2.96

FIGURE 47 AVERAGE OUT OF DIRECTION DISTANCE (MILES) - NAPA/SOLANO

Time-of-day

TM1.5 does not have a stop departure model and instead uses a lookup table to determine the stop departure time. Figure 48 and Figure 49 compare the stop and trip departure profiles between SNABM and the CHTS data. As shown, the model overestimates trip departures in the evening period while slightly underestimates trip departures in the PM period.

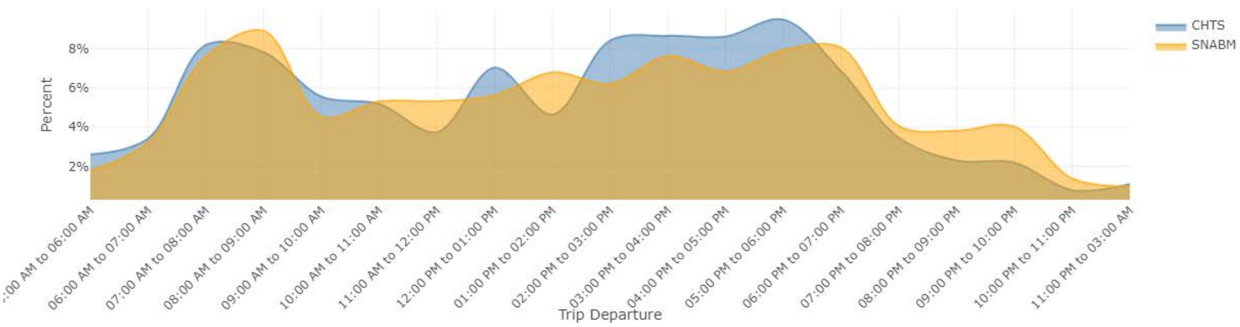


FIGURE 48 TRIP DEPARTURE PROFILE - NAPA/SOLANO

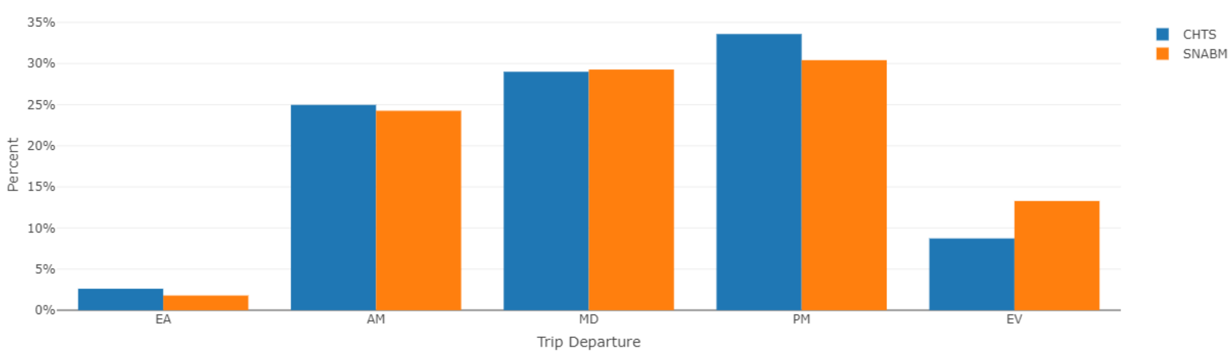


FIGURE 49 AGGREGATE TRIP DEPARTURE - NAPA/SOLANO

Trip mode

The trip mode choice targets were also augmented by the transit trip mode choice targets from the OBS. The final trip mode choice targets were scaled after each run to calibrate the trip mode choice model to the absolute number of transit trips from OBS. Please note that the trip mode choice model was calibrated only for the Napa and Solano County residents. Figure 50 compares SNABM trip mode choice distribution against the final target trip mode choice distribution targets. It can be observed that the model replicates the target distribution closely.

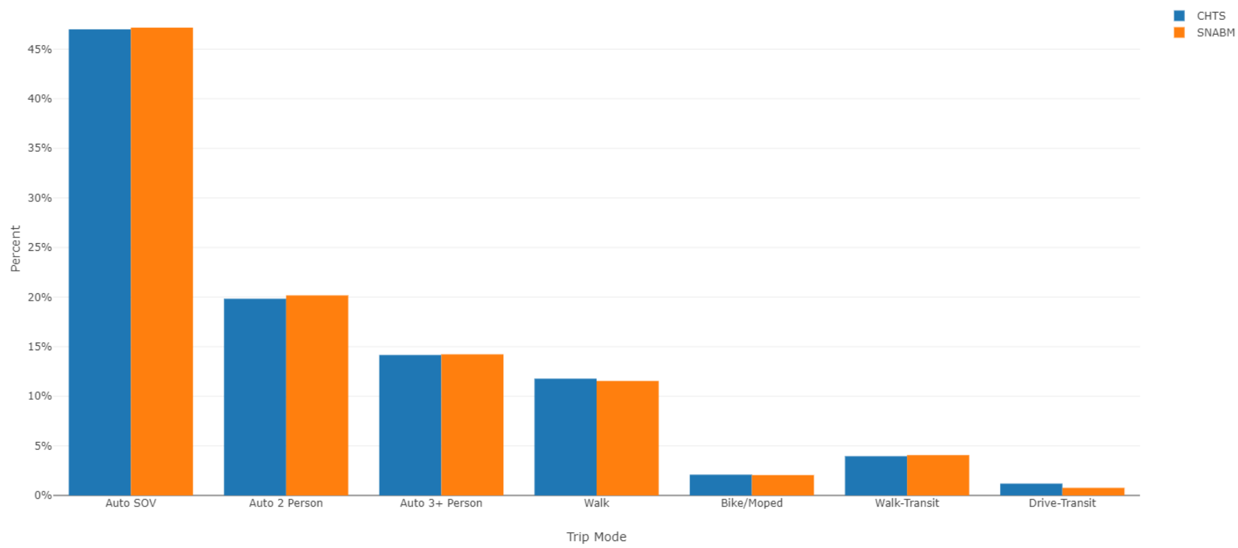


FIGURE 50 TRIP MODE - BAY AREA

Figure 51 presents trip mode percentages for all Napa/Solano County-generated trips. As shown, SNABM closely matches the observed shares. The transit shares were adjusted to match the transit ridership targets of the Napa/Solano County transit operators.

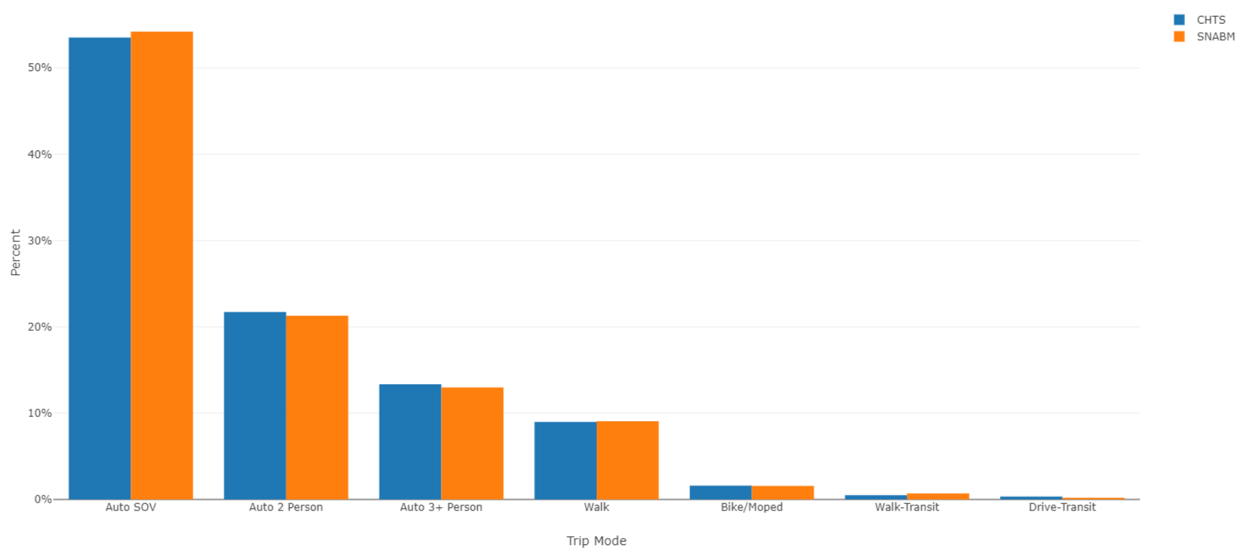


FIGURE 51 TRIP MODE - NAPA/SOLANO

6.0 MODEL VALIDATION

6.1 HIGHWAY VALIDATION

Travel demand models must replicate actual ground conditions before being considered fit to be used in travel forecasting studies. Traffic volumes for each roadway facility type are compared to traffic counts and various statistics are calculated to see if the model meets the conditions published in the FHWA "Travel Model Validation and Reasonableness Checking Manual"[13]. The following aggregate validation standards were used: Percent error, Percent root mean squared error (percent RMSE) and R squared (R^2). Each of these measures is described below:

- Percent Error: This is simply the difference between observed and model values. This is an aggregate validation statistic and is typically used to show how well the model is performing across different roadway classes. It is a simple measure but is often misleading because of it can cancel out over and under prediction of values. Formula is listed below:

$$percent\ error = \frac{\sum_{i=1}^{N} Volume_i - \sum_{i=1}^{N} Count_i}{\sum_{i=1}^{N} Count_i} \times 100$$

- Percent Root Mean Squared Error: This is a more precise measure that overcomes the shortcomings of % error by squaring differences. % RMSE measures how many of the individual links match the observed data as opposed to average in the previous measure. A % RMSE of less than 40% is the target for most regional models, however it is harder for smaller models to meet it.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}[(Count_i - Model_i)^2]}{N}}$$

and

$$\%RMSE = \frac{RMSE}{\left(\frac{\sum_{i=1}^{N} Count_i}{N}\right)} \times 100$$

[13] Federal Highway Administration, Travel Model Improvement Program, Model Validation and Reasonableness Checking Manual, 2010, https://www.fhwa.dot.gov/planning/tmip/publications/other_reports/validation_and_reasonableness_2010/fhwahep10042.pdf.

- **R^2** - Observed and Model volumes are plotted in a graph to see how close the values are to the fitted regression line. The coefficient of determination or R^2 is calculated from the plot and shows the variation in model volumes and traffic counts.

Daily Highway Validation

Daily validation was conducted using all the available counts. Model Volumes were compared to traffic counts on a variety of statistics. Aggregate statistics like validation by facility type and area type give an overall indication of the quality of the model. Typically, Screen line analysis is performed where traffic volumes along North-South and East-West corridors are compared to traffic counts to ensure accuracy of trip distribution. This needed counts on several geographic locations on key facilities which was not available. Instead, validation tables were developed on several locations on key freeways such as I-80, Hwy 29, SR 37 etc. and all the county-county border crossings. The following tables show that the model volumes match traffic counts within the acceptable error. Figure 52 shows the scatter plot of daily volumes versus traffic counts and the R^2 which shows that it is a good fit.

In addition to the daily assignment, AM and PM peak hour assignments were also performed. Number of counts available for peak hours was limited but adequate for the analysis. Similar validation tests were conducted on these time periods and the model was found to be performing well. Tables X to Y shows the peak hour validation.

Table 17: Daily Validation by Facility Type

Facility Type	Observed Traffic Counts	2015 Estimated Volumes	Daily Validation	
	Sum of Counts	Sum of Model Volumes	Percent RMSE	Percent Error
Freeways	3,288,180	3,178,518	24%	-3%
Expressways	606,692	645,598	26%	6%
Arterials	1,206,174	1,047,429	58%	-13%
All	**5,101,046**	**4,871,545**	**41%**	**-4%**

Table 18: Daily Validation by Screenline

Screenline	Observed Traffic Counts	2015 Estimated Volumes	Daily Validation	
	Sum of Counts	Sum of Model Volumes	Difference	Percent Error
Napa - Solano Border (EB)	95,333	106,268	10,936	11%
Solano - Napa Border (WB)	95,333	104,016	8,683	9%
Napa - Sonoma (NW)	8,761	10,374	1,613	18%
Sonoma - Napa (SE)	8,761	10,727	1,967	22%
Sonoma - Napa/Solano (EB)	31,995	32,069	75	0%
Napa/Solano-Sonoma (WB)	31,995	32,069	75	0%
Solano Contra Costa (SB)	125,001	126,889	1,888	2%
Contra Costa - Solano (NB)	125,001	117,659	-7,342	-6%
Solano-Yolo/Sacramento NB	84,151	83,150	-1,001	-1%
Yolo/Sacramento-Solano SB	84,120	83,150	-970	-1%
Grand Total	**522,177**	**540,072**	**17,895**	**3%**

Table 19: Daily Validation on Interstate 80

Screenline	Observed Traffic Counts	2015 Estimated Volumes	Daily Validation	
	Sum of Counts	Sum of Model Volumes	Difference	Percent Error
I80 Carqinez Br	125,000	116,197	-8,803	-7%
I80 before 780	125,000	120,330	-4,670	-4%
I80 after 780	152,000	138,269	-13,731	-9%
I80 AmCanyon Rd	139,000	108,840	-30,160	-22%
I80 Red Top Rd	136,000	116,701	-19,299	-14%

I80 before Hwy12	171,395	118,335	-53,060	-31%
I80 after 680	190,231	194,685	4,454	2%
I80 after 12E	175,318	176,577	1,259	1%
I80 W Texas Rd	151,382	168,339	16,957	11%
I80 Travis Blvd	164,375	154,377	-9,998	-6%
I80 Pleasant Valley Rd	167,226	180,552	13,326	8%
I80 Elmira Rd	169,000	189,872	20,872	12%
I80 Vaca Valley Pkwy	136,000	143,331	7,331	5%
I80 Dixon Ave	132,000	144,761	12,761	10%
I80 Stratford Ave	131,000	132,214	1,214	1%
I80 Tremont Rd	135,000	142,107	7,107	5%
I80 Sol-Sac Border	140,000	142,282	2,282	2%
Grand Total	**2,539,927**	**2,487,770**	**(52,157)**	**-2%**

Table 20: Daily Validation in Napa County

Screenline	**Observed Traffic Counts**	**2015 Estimated Volumes**	**Daily Validation**	
	Sum of Counts	Sum of Model Volumes	Difference	Percent Error
After American Canyon Rd	52,881	66,163	13,282	25%
After Eucalyptus Dr	53,413	62,342	8,929	17%
Before 221	64,000	84,171	20,171	32%
After 221	47,500	53,761	6,261	13%
Before 121	47,500	52,480	4,980	10%
Redwood Rd	44,500	51,669	7,169	16%
Yountville Cross Rd	38,070	34,605	-3,465	-9%
Larkmead	12,195	9,818	-2,377	-19%
Calistoga	13,083	16,337	3,254	25%
Grand Total	**373,142**	**431,345**	**58,203**	**16%**

Table 21: Daily Validation in Vallejo

Screenline	Observed Traffic Counts	2015 Estimated Volumes	Daily Validation	
	Sum of Counts	Sum of Model Volumes	Difference	Percent Error
Sears Point Rd	33,800	37,949	4,149	12%
Mare Island	36,700	39,749	3,049	8%
Six Flags I80	67,863	67,606	-257	0%
Grand Total	**138,363**	**145,305**	**6,942**	**5%**

Figure 52: Scatterplot of Daily Volumes

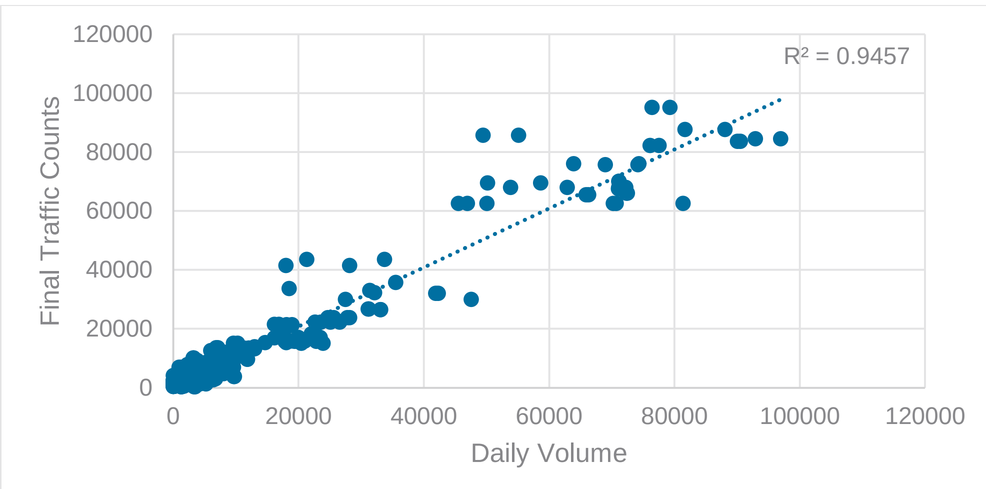


Table 22: HPMS VMT

County	HPMS VMT	Model VMT	% Difference	% Difference Target
Solano County	13,153,967	13,666,792	4%	**+/- 5%**
Napa County	3,329,608	3,312,423	-1%	
All	**16,483,575**	**16,979,215**	**3%**	

6.2 TRANSIT VALIDATION

Transit validation was performed for the daily time period. Total boarding's for each operator were compared to observed numbers of the same and were found to match well. Table 23 shows the transit ridership by operator. Observed ridership was obtained from MTC's Statistical Summary of Bay Area Operators.

Table 23: Transit Ridership by Operator

Operator	Observed Ridership	Model Ridership
Napa Vine	3,865	3,678
Soltrans	5,047	5,773
Vallejo Ferry	2,686	3,224
Fairfield Bus	3,373	3,532
Vacaville Bus	1,765	1,585
Total	**16,736**	**17,792**

6.3 FUTURE YEAR RUN

The calibrated and validated travel model was used to develop future year forecasts. Inputs to the future year model included:

- Future Year Transportation Networks – Both Highway and transit networks were developed based on the regional MTC model's Plan Bay Area 40 networks. Approved projects from the MTC network were coded into the more detailed SNABM model networks.
- External trip tables – Trips that start or end or pass through the study area were derived from the MTC travel model. Growth rates from the MTC model were applied to updated 2015 external trips to develop the 2040 trip tables.
- Airport Passenger and High Speed Rail Trip Tables for 2040 – These were also taken directly from the MTC model and disaggregated to the SNABM zone system.
- 2040 Land use Projects – The starting point for 2040 demographics and land use data was MTC's Plan Bay Area 40. Procedures similar to 2015 were used to disaggregate the 1454 MTC TAZ data to 2340 SNABM TAZ data. Below is the summary table for the 2040 land use data used in SNABM. Growth in Solano and Napa has been moderate. Solano population increased at an annual growth rate of 0.8% and jobs increased at 0.6%. Absolute growth in population in Solano from year 2015 to 2040 is 22% while jobs grew 15% in the 25-year time frame. Napa population increased at an annual growth rate of 0.5% and jobs increased at 0.7%. Absolute growth in population in Napa from year 2015 to 2040 is 15% while jobs grew 18% in the 25-year time frame. Figures 53 and 54 show the growth in population and employment from 2015 to 2020.

Table 24: 2040 LAND USE DATA

County	Total Population	Total Households	Total Employment
Solano County	498,850	169,288	150,975
Napa County	158,038	54,691	83,361
Grand Total	656,888	223,979	234,336

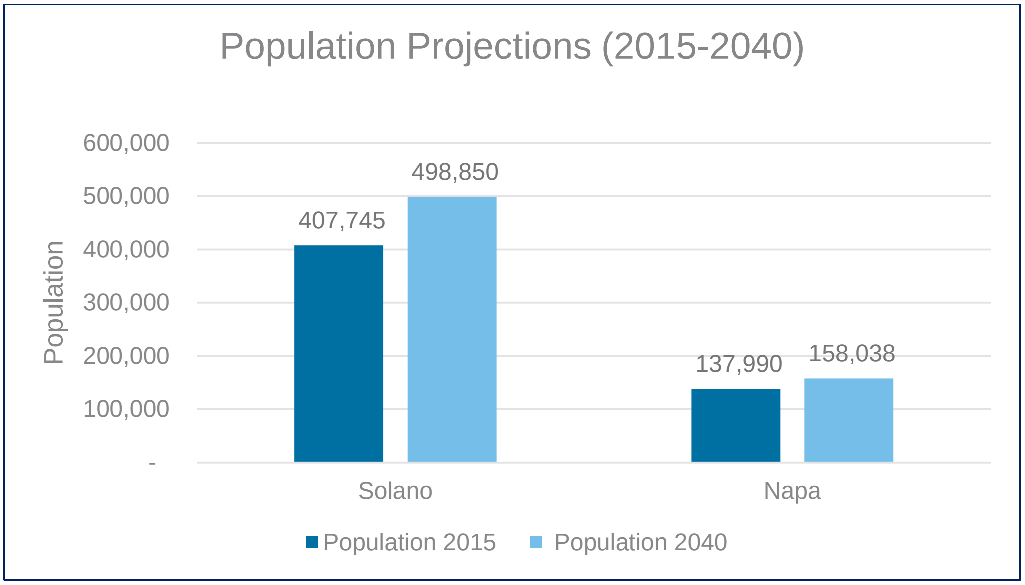


FIGURE 53 – POPULATION PROJECTIONS (2015-2040)

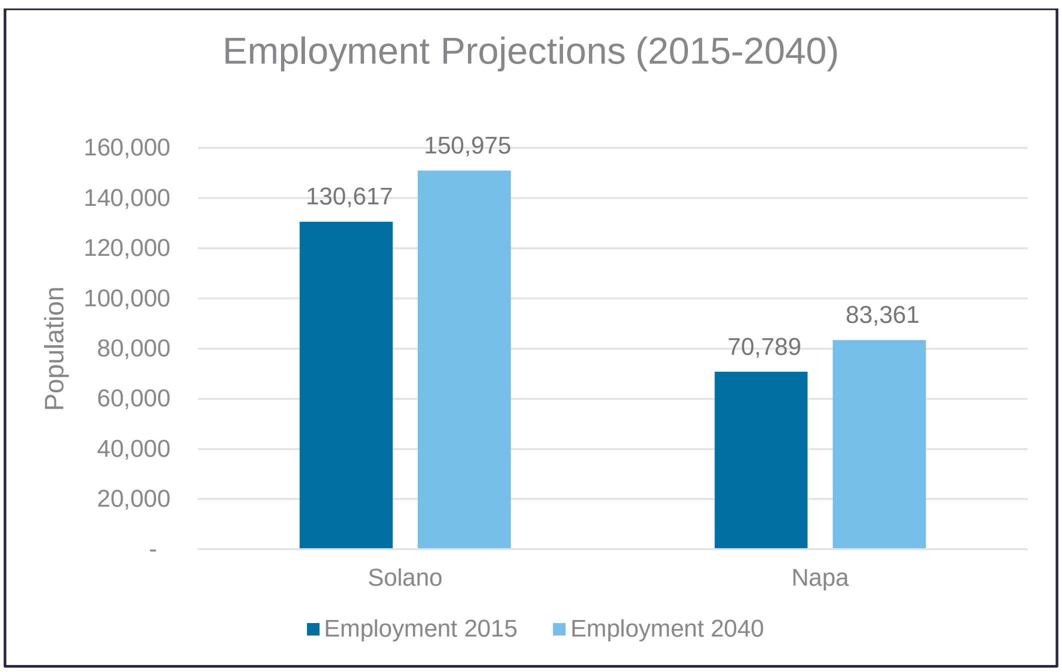


FIGURE 54 – EMPLOYMENT PROJECTIONS (2015-2040)

The travel model was run for year 2040 after updating all the inputs. Following tables show the growth in highway volumes from 2015 to 2015. The following table show the growth in volumes on the same links with traffic counts that were used in validation. Over all growth in traffic for links of all facility types increased by 52%.

Table 25: Growth in Volumes between 2015 and 2040 by Facility Type

Facility Type	2015 Estimated Volumes	2040 Estimated Volumes	Growth %
Freeways	3,277,137	4,930,771	50%
Expressways	628,414	1,117,581	78%
Arterials	1,013,654	1,447,146	43%
All	**4,919,205**	**7,495,498**	**52%**

The following tables show growth in traffic for specific screen lines in both Solano and Napa counties. Individual roadways see a growth between 20% to over 100% in some cases depending on the growth in land uses in their vicinity.

Table 26: Growth in Volumes between 2015 and 2040 by Screenline

Screenline	2015 Estimated Volumes	2040 Estimated Volumes	Growth %
Napa - Solano Border (EB)	106,306	190,134	79%
Solano - Napa Border (WB)	103,731	197,366	90%
Napa - Sonoma (NW)	9,840	30,668	212%
Sonoma - Napa (SE)	10,025	25,719	157%
Sonoma - Napa/Solano (EB)	34,265	62,355	82%
Napa/Solano - Sonoma (WB)	34,265	62,355	82%
Solano Contra Costa (SB)	134,964	262,536	95%
Contra Costa - Solano (NB)	125,408	262,550	109%
Solano-Yolo/Sacramento NB	83,150	102,774	24%
Yolo/Sacramento-Solano SB	83,150	103,210	24%
Total	**725,105**	**1,093,683**	**51%**

Table 27: Growth in Volumes between 2015 and 2040 on I-80

Screenline	2015 Estimated Volumes	2040 Estimated Volumes	Growth %
I80 Carqinez Br	125,000	266,081	113%
I80 before 780	125,000	219,882	76%
I80 after 780	152,000	224,098	47%
I80 AmCanyon Rd	139,000	205,480	48%
I80 Red Top Rd	136,000	221,417	63%
I80 before Hwy12	171,395	216,444	26%
I80 after 680	190,231	317,038	67%
I80 after 12E	175,318	269,853	54%
I80 W Texas Rd	151,382	257,612	70%
I80 Travis Blvd	164,375	236,842	44%
I80 Pleasant Valley Rd	167,226	238,265	42%
I80 Elmira Rd	169,000	216,893	28%
I80 Vaca Valley Pkwy	136,000	169,060	24%
I80 Dixon Ave	132,000	169,803	29%
I80 Stratford Ave	131,000	154,523	18%
I80 Tremont Rd	135,000	168,511	25%
I80 Sol-Sac Border	140,000	168,715	21%
Grand Total	**2,539,927**	**3,720,517**	**46%**

Table 28: Growth in Volumes between 2015 and 2040 in Napa County

Screenline	2015 Estimated Volumes	2040 Estimated Volumes	Growth %
After Amcanyon RD	52,881	109,527	107%
After Eucalyptus Dr	53,413	113,574	113%
Before 221	64,000	136,515	113%
After 221	47,500	109,290	130%
Before 121	47,500	92,114	94%
Redwood Rd	44,500	71,979	62%
Yountville Cross Rd	38,070	53,064	39%
Larkmead	12,195	20,051	64%
Calistoga	13,083	20,550	57%
Grand Total	**373,142**	**726,664**	**95%**

Table 29: Growth in Volumes between 2015 and 2040 in Vallejo

Screenline	2015 Estimated Volumes	2040 Estimated Volumes	Growth %
Sears Point Rd	33,800	71,699	112%
Mare Island	36,700	60,826	66%
Six Flags I80	67,863	97,305	43%
Grand Total	**138,363**	**229,830**	**66%**

7.0 SUMMARY

SNABM is based on TM1.5 and covers the entire 9-county Bay Area region. TJKM and RSG implemented additional features to improve the model prediction accuracy for Napa and Solano counties. We developed a finer geographic system and associated land use data for Napa and Solano counties. We also implemented intelligent household sampling to reduce Monte Carlo variation in Solano/Napa counties. PopulationSim was used to generate the synthetic population for both the base (2015) and future year (2040). The model was calibrated and validated to the local observed data for Napa/Solano counties. As discussed in Chapter 6, the base year SNABM validation generally meets all the validation criteria developed by Caltrans for the highway network within the Napa and Solano counties. Based on model calibration and validation, we believe the model is suitable for the development of alternative traffic forecasts on the highway and transit networks. However, we recommend that for project-level applications, more localized subarea validation effort must be undertaken to improve model accuracy in the study area.

We believe the next steps for SNABM include conducting sensitivity tests and subarea analysis to better understand the model behavior. As noted in Chapter 5, we found that there are fewer fulltime workers in Napa/Solano counties compared to CHTS. This should be addressed by generating synthetic population using PopulationSim controls that accurately represents the number of workers in Napa/Solano counties. As part of future work, methods should be developed to account for visitor travel, especially in downtown Napa County. Both STA and NVTA are interested in development of a VMT generating tool to fulfill the requirements of SB 743[14]. SNABM can be used for performing subarea analysis to study the impact of development activity in a TAZ. For this purpose, the 're-populate' feature in PopulationSim can be used that can add to or replace households in just one zone according to simple controls. However, since the SNABM is a simulation, reducing the simulation variance due to the use of random number draws can be challenging. The simulation variance is reduced by oversampling the households. The intelligent sampling feature that we implemented in SNABM oversamples the households in Napa/Solano counties by a factor of three, and then assigns only 1/3rd of household trips. However, oversampling by a factor of 3 may not sufficiently reduce the sample variance for very small changes in inputs. Therefore, we recommend additional testing to understand how the model behaves with respect to different level of changes in the inputs, and what sample rate to use. The sample rate selection is not automated in the current implementation, but this can be implemented under a future development effort.

[14] Caltrans. Senate Bill (SB) 743 Implementation. https://dot.ca.gov/programs/transportation-planning/office-of-smart-mobility-climate-change/sb-743

8.0 APPENDIX A: SNABM USER GUIDE

This appendix provides details of SNABM setup and instructions on configuring and performing SNABM model runs. SNABM is based on TM1.5 and has the same setup as TM1.5. MTC has created a comprehensive online user guide for TM1.5[15]. This guide covers all the details briefly and points the user to appropriate sections on MTC's online user guide. The following sections describe the software requirements, model setup, configuration details, and instructions on running the model.

8.1 SOFTWARE REQUIREMENTS

The various software required for running SNABM are briefly described in the following subsections.

Citilabs Cube Voyager

SNABM relies on Cube Voyager for skimming, network manipulation, matrix calculation, and performing highway and transit assignment. SNABM was developed using Cube 6.4.2.

Citilabs Cube Cluster

The Cube Cluster software allows for the Cube scripts to be multi-threaded. In the current approach, the travel model uses 48 computing nodes. Cube Cluster is not strictly necessary, as the Cube scripts can be modified to use only a single computing node. Such an approach would dramatically increase run times.

Java

SNABM uses the CT-RAMP modeling framework which was developed in Java programming language[16]. The 64-bit Java Development Kit (version 1.8) must be installed on each machine running the CT-RAMP software. The Java Development Kit includes the Java Runtime Environment used by Java processes.

GAWK

Certain text file manipulations are handled in SNABM using the free GAWK software. GAWK can be installed from scratch or the installation folder from another machine can be copied. The SNABM model setup deliverable includes a full install of GAWK software.

15 https://github.com/BayAreaMetro/modeling-website/wiki/UsersGuide
16 https://java.com/en/

Python

Python 2.7 (64-bit) is used to execute a variety of utility scripts. After installing Python, following packages must be installed: Shapely, numpy, pandas, SimpleParse, xlrd, xlwt, xlutils, dbfpy, pywin32, and rpy2.

Network Wrangler

The transit line files are manipulated using Python scripts in the Network Wrangler package. The Network Wrangler repository (https://github.com/sfcta/NetworkWrangler) must be cloned to a local directory on the machine where the model is being run.

R

There are some R scripts used to generate output summaries. R version 3.4 or higher must be installed. Following packages are required for running the post processing scripts: dplyr, foreign, scales, reshape2, and tidyr.

Please note that some MTC utility scripts use Tableau for creating output summaries, but these scripts are not required for the full model run.

8.2 SETUP DETAILS

The SNABM setup consists of two directories, CTRAMP and INPUT and one MS DOS batch file, RunModel.bat. These files must be copied to a working directory (see Table 30) on the computer to be used for running the model.

TABLE 30 SNABM WORKING DIRECTORY CONTENTS

File	Description
/CTRAMP	Directory containing all the model configuration files, Java instructions, and Cube scripts required to run SNABM
/INPUT	Directory containing all the Input files required to run a specific SNABM scenario
RunModel.bat	Batch file for running SNABM

The RunModel.bat contains a list of MS-DOS instructions that control model flow.

/CTRAMP Directory

Table 31 describes the contents of the */CTRAMP* directory.

TABLE 31 /CTRAMP DIRECTORY CONTENTS

File	Description
/model	Contains all the Utility Expression Calculator files that specify the choice models in CTRAMP
/runtime	Contains all the Java configuration and JAR (Java executable) files, as well as the files necessary for Java to communicate with Cube
/scripts	Contains all the Cube, Python, and other helper scripts. The scripts are grouped into the following sub-directories: ▪ */assign*: contains scripts used for performing assignment ▪ */block*: contains Cube block files for highway and transit assignment ▪ */core*_summaries: contains scripts to create model output summaries ▪ */database*: contains script to create skim database ▪ *empfac:* contains scripts relating to empfac2011 ▪ */feedback:* contains scripts to prep feedback iteration ▪ */nonres:* contains files relating to nonresidential model components ▪ */preprocess:* contains scripts used to preprocess model inputs ▪ */skims:* contains scripts used for creating skims

/INPUT Directory

Table 32 presents the contents of the */INPUT* directory.

TABLE 32 INPUT DIRECTORY CONTENTS

File	Description
/hwy	Contains the input free flow highway network, which is named, by convention, freeflow.net
/trn	Contains all the input transit network files: transit line file, transit fare file, and other supporting files
/landuse	Contains the socio-economic input land use file, walk access buffer file, and sample rate file used for intelligent sampling of synthetic population
/nonres	Contains the fixed, internal/external trip tables, the fixed, air passenger trip tables, and files used to support the commercial vehicle model
/popsyn	Contains the synthetic population files in the format required for SNABM

8.3 CONFIGURATION

Configuring SNABM involves setting paths to installed software and properties in various batch files that control the model run. Some of these paths are common across various batch files. To simplify the process of updating the paths, the common set of paths are specified in the *ctramp/runtime/config/SetPath.bat* batch file. The batch file is called at the beginning of the model run and sets the paths to the installed software which are used by the subsequent processes. The properties file controlling the CTRAMP run and JPPF services must be configured. The two batch files (JavaOnly_runMain.cmd and JavaOnly_runNode0.cmd) that executes the assigned Java tasks on a single machine also need to be configured. In addition, a few variables in the RunModel.bat file need to be configured. In order to edit a batch file, right click on the file and then select a text editor to open the file. Alternatively, the batch file can be dragged and dropped into an open text editor window. Please do not double-click the batch file to open it for editing. Double-clicking a batch file will cause it to run. Running the batch file without configuring the settings would likely result in an error.

The following sections summarize the variables that need to be configured in each of these files.

ctramp/runtime/SetPath.bat

Table 33 present the list of variables set in SetPath.bat file.

TABLE 33 SETPATH.BAT CONFIGURATION

Variable	Description
COMMPATH	Working directory
JAVA_PATH	The location of the 64-bit java development kit
GAWK_PATH	The location of the GAWK binary executable files
R_HOME	Contains the fixed, internal/external trip tables, the fixed, air passenger trip tables, and files used to support the commercial vehicle model
R_LIB	The location of R installation
TPP_PATH	The location of R libraries
PYTHON_PATH	The location of python executable
RUNTIME	The location of the MTC.JAR file

RunModel.bat

The RunModel.bat file orchestrates the SNABM run by calling various processes to execute the travel model run. A detailed description of all steps can be found here: https://github.com/BayAreaMetro/modeling-website/wiki/RunModelBatch

The main variables to be configured in this file are the host IP address, maximum number of feedback iterations, and *SELECT_COUNTY* option. The example setting for setting the host IP address is shown below:

if %computername%==PORMDLPPW02   set HOST_IP_ADDRESS=172.24.0.102

The *SELECT_COUNTY* options enables the intelligent sampling if it is set to a value greater than zero. In addition, the user must set the *PYTHONPATH* variable in Step 4.5 to the location of the cloned Network Wrangler repository.

RunIteration.bat

This file runs the core travel model processes in each iteration which includes skimming, demand models, and assignment. A full description of various steps executed by this batch file can be found here: https://github.com/BayAreaMetro/modeling-website/wiki/RunIterationBatch. The main configurable variables in this script relates to Java process. Under normal execution, the variables in this batch file are not required to be updated.

ctramp/runtime/JavaOnly_runMain.cmd

The *HOST_IP* variable must be set to the IP address of the computer being used for running the model. This script calls the java processes to launch the main java process, household manager and the matrix manager. Further information on household manager and matrix manager can be found here: https://github.com/BayAreaMetro/modeling-website/wiki/SystemDesign#household-manager-and-matrix-manager. Both the Household Manager and Matrix Manager have substantial memory footprints. A maximum and minimum amount of Random-Access-Memory (RAM) can be assigned to these Java processes by setting the *Xmx* and *Xms* arguments on the call to these Java processes. The memory must be allocated in accordance to the total available RAM on the machine.

ctramp/runtime/JavaOnly_runNode0.cmd

This script calls the Java node which connects to the main Java process. In this script too, the *HOST_IP* variable must be set to the IP address of the computer being used for the model run. Appropriate amount of memory must be allocated based on the total available RAM.

ctramp/runtime/config/ jppf-driver.properties & jppf-node0.properties

In these properties file, the *HOST_IP* variable must be set to the IP address of the computer being used for the model run. The allocated memory can be adjusted using the *other.jvm.options* settings.

ctramp/runtime/mtcTourBased.properties

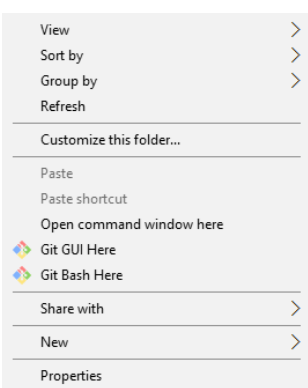
The core CTRAMP software that executes the demand portion of the SNABM model is controlled by the mtcTourBased.properties file. The main settings to be configured when setting up a new scenario are the location of the project directory and names of input synthetic population files. A detailed description of all CTRAMP properties exposed via this properties file can be found here: https://github.com/BayAreaMetro/modeling-website/wiki/PropertiesFile.

RuntimeConfiguration.py

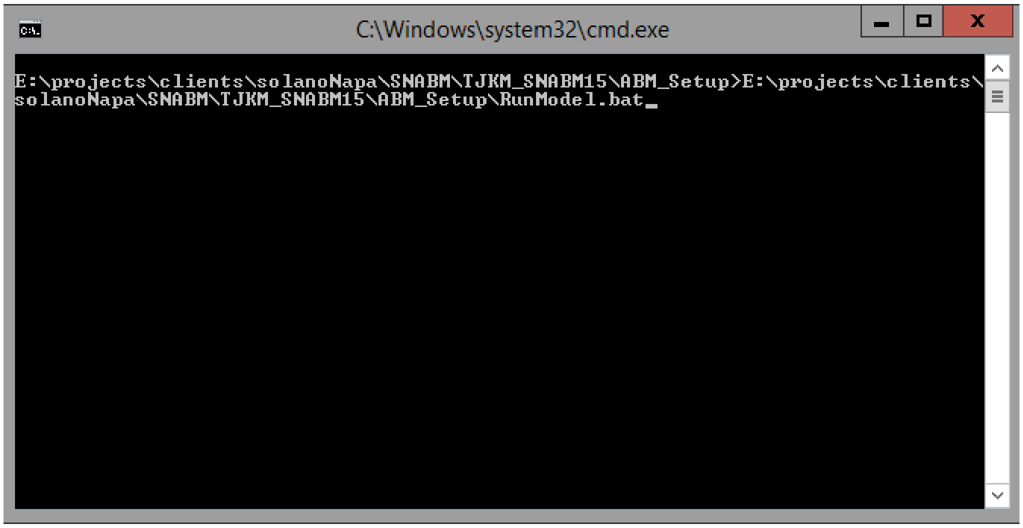
RuntimeConfiguration.py is meant to do the remainder of the runtime-setup automatically, so that all the configuration is in one place and done automatically. See the script for details here: https://github.com/BayAreaMetro/travel-model-one/blob/master/model-files/scripts/preprocess/RuntimeConfiguration.py.

8.4 RUNNING SNABM

Before launching a run, the user must ensure that all configuration settings have been specified appropriately. The user can launch the SNABM run by double-clicking the RunModel.bat file. Alternatively, the run can be launched by opening a command prompt window in the working directory and calling the batch file. The command prompt window can be opened by holding shift and right-clicking inside the working directory and selecting the "Open command window here" option as shown below:



Next, drag and drop the batch file into the open CMD window or simply type the batch file name. Press Enter to launch the SNABM run.



A complete SNABM model run takes approximately 12 to 16 hours to complete, depending on the machine configuration of the end user.

RSG
the science of insight

55 Railroad Row
White River Junction, VT 05001
802.295.4999
www.rsginc.com

White River Junction & Burlington, VT | **Arlington, VA** | **Chicago, IL** | **Evansville, IN** | **Portland, OR** | **Salt Lake City, UT** | **San Diego, CA**