FRESNO COUNCIL OF GOVERNMENTS

FRESNO ACTIVITY-BASED MODEL UPDATE

August 30, 2018

PREPARED FOR:
FRESNO COUNCIL OF GOVERNMENTS

SUBMITTED BY:
RSG

600 B Street, Suite 2202
San Diego, CA 92101
619.501.0559
www.rsginc.com

IN COOPERATION WITH:
BOWMAN RESEARCH AND CONSULTING

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION

This report describes updates to the Fresno Council of Governments (Fresno COG) activity-based travel demand model (ABM) to meet existing and evolving transportation planning needs. The new model system is capable of addressing policies such as compact and mixed-use development, active transportation, transit, and pricing. The model is credible for forecasting demand for highway alternatives such as new river crossings and corridor improvements, and appropriately sensitive to land-use changes such as new planned developments and provide useful information for traffic impact studies.

The report describes the new ABM system in detail and is also aimed to serve as a user's guide for the model system. The ABM is integrated with Cube software, which is primarily used for network models (skimming and assignment) and auxiliary demand models (truck model and external truck demand). The ABM base-year model (2014) is calibrated using 2012 California Household Travel Survey, 2010 National Household Travel Survey, and 2014 transit on-board survey. The model system is also validated against observed data for traffic counts, transit ridership, and bike and pedestrian counts.

This project undertook the following key tasks to update the Fresno activity-based model:

- Developed a synthetic population (PopSynIII)
- Implemented latest version of DaySim
- Implemented oversampling of households by geography
- Implemented non-motorized skimming and assignment
- Generated input external auto trip tables from the California Statewide model
- Implemented zone-based transit fare
- Analyzed Streetlight origin-destination travel data
- Developed output reporting procedures
- Calibrated and validated the ABM
- Examined model sensitivities to land-use and network changes
- Set up a GitHub repository for model versioning[1]

The rest of this report refers to the new ABM system as "Fresno ABM" and is organized as follows. The next chapter, Chapter 2, describes features of the new ABM system. Then Chapter 3 presents various summaries of outputs from the calibrated and validated

[1] Fresno ABM GitHub repository: https://github.com/RSGInc/FresnoABM

model. Chapter 4 discusses the tests examining model sensitivities to land-use and network changes. Chapter 5 provides instructions to run the model system. Chapter 6 describes the model run directory and its contents. At the end of the report, several appendices (Appendix A-D) provide supplemental information about various tasks performed during this project.

2.0 MODEL DESIGN

The Fresno ABM is implemented based on the VMIP2 model for the Fresno County. The most significant modification includes replacing the resident travel generation (trip generation, trip distribution, and mode choice) with an activity-based model (DaySim). The other important updates include internal-external demand and non-motorized (bike and ped) skimming/assignment. Auxiliary demand (truck model and external truck travel) and network models (skimming and assignment) are retained from the VMIP2. The Fresno ABM is versioned on a GitHub repository[2]: https://github.com/RSGInc/FresnoABM.

2.1 OVERVIEW

Figure 1 presents structure of the new Fresno activity-based model (ABM) system showing relationship among different sub-model components.

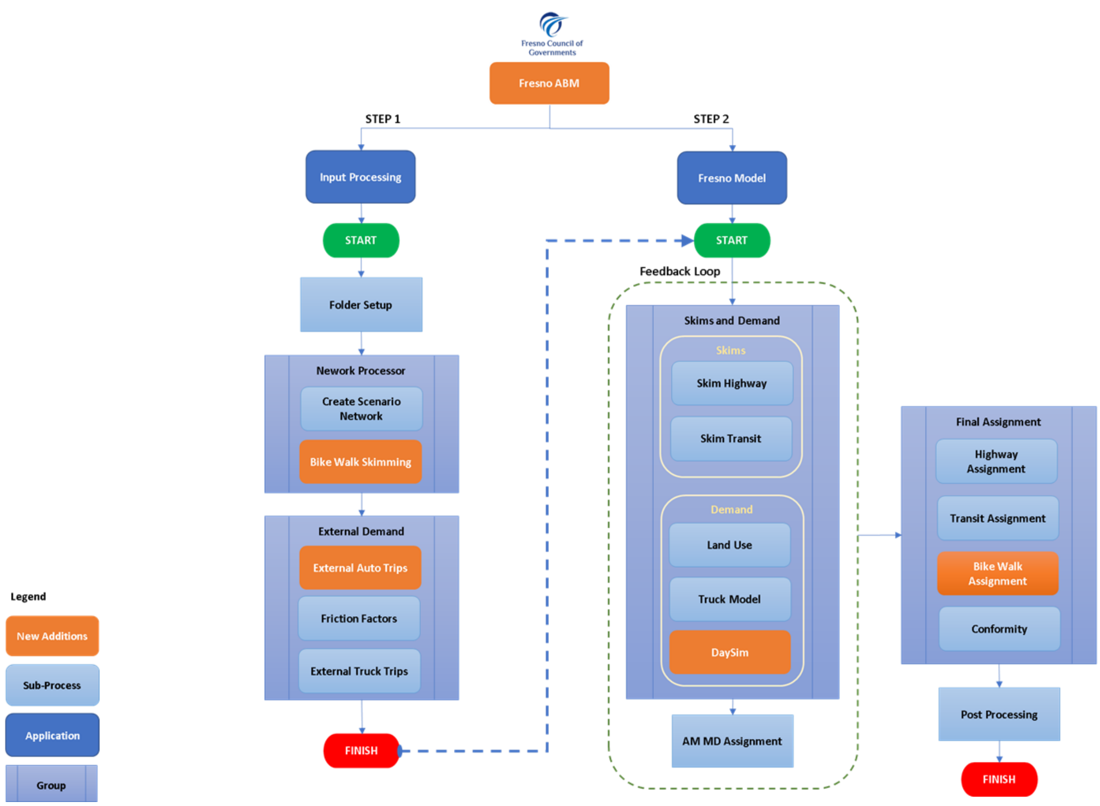


FIGURE 1: MODEL PROCESS FLOW

[2] The repository is private and requires permissions to access. Presently, RSG maintains the repository.

The ABM is run in two sequential steps: input processing and Fresno model. The first step, input processing, process the model network and creates inputs required by the second step, Fresno model, which performs network skimming, demand generation, and assignment in three feedback loops.

The input processing step first converts the input network in a geodatabase to CUBE network formats (highway, transit, and non-motorized) for network skimming and assignment procedures. As non-motorized assignments are not capacity constrained, re-skimming of bike and walk are not required in every feedback loop. Therefore, an initial set of bike and walk skims are generated to use in every feedback loop. Similarly, this step generates external demand matrices (auto and truck) which are kept static across all feedback loops.

After the networks are prepared in required format, the second step, Fresno Model, runs three feedback loops. Each feedback loop starts with generating highway and transit skims to use in demand models. A set of demand models produce demand for two travel markets in the Fresno County region: resident travel and truck travel. The resident travel demand is generated by DaySim, whereas the truck demand is produced by the truck model, which is retained from the VMIP2 model. The two sets of travel demand, along with the external demand (auto and truck) generated in the first step, input processing, are combined into four time-period specific highway demand matrices with each having OD demand by multiple highway classes based on vehicle type and value of time. Similarly, the transit demand from DaySim are also prepared into four time period specific transit demand matrices with each having multiple transit assignment classes based on transit sub-mode and access mode. The non-motorized (bike and walk) demand too come from DaySim model and the corresponding matrices are for daily demand. After the demand matrices are constructed, two highway assignments are run: peak (AM) and off-peak (MD). Note that transit and non-motorized assignments are run only once, in the final assignment step after the feedback loops. The link flows from the two highway assignments then inform new link travel times for the next feedback loop.

After the model system runs three feedback loops of skimming, demand, and assignment, a final assignment step runs assignment for all network systems (highway, transit, and non-motorized). The final highway assignment is run for four time periods (AM, MD, PM, and EV) and the transit assignment (also skimming) is performed for peak and off-peak periods. The non-motorized assignments are run for daily demand. After the network assignments, the last step of the model system summarizes the assignment results and produces files/reports to use for analysis.

Note that before running DaySim, a utility called “Buffer Tool” generates a buffered micro-zone file (land-use and accessibility measure for micro-zones) to use in DaySim. Also, a python tool called “population sampler” updates the input synthetic population to oversample households by a factor of 3.0 to reduce the Monte Carlo variance for an individual DaySim model run. The DaySim model is run four times in each feedback loop; the first three iterations of DaySim produce stable shadow prices that are used as an input for the last DaySim run.

The Fresno ABM system process is comprised of the following key elements: input preparation, network skimming, demand generation, assignment preparation, assignment, feedback loop, and reporting. The following sections discuss these processes in more detail.

2.2 INPUT PREPARATION

The first step of the model system, Input Processing, performs all necessary calculations and conversions to prepare networks and other inputs for processes in the second step, Fresno ABM, of the model system. This input preparation process undertakes the following four primary tasks that are completed sequentially:

- Setup scenario directory
- Create scenario networks
- Generate non-motorized skims
- Generate external demand

Setup Scenario Directory

When the input processing step is started, it first creates a scenario directory with all necessary sub-folders required at any step of the model system.

Create Scenario Networks

After setting up a scenario directory, the process reads the input all-street highway network and transit network from a geodatabase and processes them to convert into Cube networks. In addition to generating necessary fields, this step filters out minor streets to create a coarser auto network for use in skimming and assignment.

Note that the process of creating the auto network does not check for connectivity. The accuracy of the output network is entirely dependent on how well the attributes that are used in the street filtering process are coded in the all-street network. Additionally, a coarser network may miss some streets with substantial traffic, thus causing, on some links, higher model flows than observed traffic counts. Substantial effort was made to track down and resolve such network disconnectivity and missing link issues during the model validation phase of this project.

Generate Non-motorized Skims

New non-motorized (walk and bike) skimming and assignment procedures were implemented in the model system as part of this project. Since non-motorized network models are not constrained by capacity, non-motorized networks are skimmed only once, during network preparation, and non-motorized assignment is run once during the final assignment step after all feedback loops have completed.

The non-motorized skimming procedures use the all-street network and generate distance skims for MAZ OD pairs[3]. The input preparation step creates a separate non-motorized network from the input all-street network and generates one distance skim for each of the two non-motorized modes. The distance skims are generated using a Cube-based shortest path method. The walk distances are based on a distance-based shortest path, whereas the bike distances are based on a shortest path minimizing a generalized cost function that is calculated as equivalent minutes weighted by bike facility class. As shown in Table 1, eight bike facility classes are available in the all-street network. The weights[4] in bike generalized cost are applied by more aggregated bike facility classes as presented in Table 2.

TABLE 1: BIKE FACILITY TYPE IN ALL STREET NETWORK

BIKE_FACTYPE	DESCRIPTION
0	Shared Roadway (No Bikeway Designation).
1	Class I Bikeway (Bike Path). Provides a completely separated right of way for the exclusive use of bicycles and pedestrians with crossflow by motorists minimized.
2	Class II Bikeway (Bike Lane). Provides a striped lane for one-way bike travel on a street or highway.
3	Class III Bikeway (Bike Route). Provides for shared use with pedestrian or motor vehicle traffic.
4	Class IV Separated Bikeway (cycle tracks). On-street bicycle facilities that include a vertical physical barrier between the bikeway and moving traffic.
5	Separate highway overcrossings
6	Unpaved Multipurpose Trails
9	Freeways and Ramps (bicycling not permitted)

[3] Auto trips are significantly longer and calculating impedances at a finer detail (MAZ) than TAZ would not be worth adding substantial run time to the model system.
[4] The parameters are borrowed from SANDAG route choice model, converted from utiles per mile to minutes per mile.

TABLE 2: PARAMETERS BY BIKE FACILITY TYPE

BIKE FACILITY	WEIGHT (MINS PER MILE) *
No bike facility	6.0
Bike trail	2.0
Bike lane	3.1
Bike route	4.9

*Note: parameters are borrowed from SANDAG bike route choice model

Note that implementing non-motorized skimming and assignment procedures required several changes to the original input all-street network (from MIP2). The changes include:

1. Coding bike facility type: a field "BIKE_FACTYPE" is added to the network. A description of the values assigned for the field is provided in Table 1.
2. Adding MAZ centroid and centroid connectors: as skims are maz-to-maz, maz information was required in the network.
    a. First, maz centroids are added to the node file by numbering the new nodes continuously and starting from 100,000. To identify these centroids, set NODETYPE = "MAZ" for these nodes.
    b. Next, in the node file, populate a new field MAZ that would have maz id of the microzone that have the node as its nearest node. So first find nearest nodes to each maz centroid and then join those nodes with the node file and populate MAZ field as the mazid.
    c. Lastly, for each maz, two maz centroid connectors were added to the link file. The corresponding A and B node fields were populated appropriately.

The skimming process first strips out all freeway facilities from the all-street network and then calculates a bike generalized cost (BIKE_MINMILE) for each links in the network. As Cube expects skim zones (nodes) to be numbered from 1, the node ids in the resulting non-motorized network are updated to bring maz centroids in the beginning and the rest of the nodes to start after them. This was done by subtracting 100,000 if node id is greater than 100,000 (maz centroids) and adding 100,000 otherwise (street nodes). Both node and link files were updated to reflect the new node id numbering scheme. Now, this updated network is used to generate bike and walk skims using a Cube utility (BUILDPATH) that, for every OD pair, summarizes an attribute on a shortest path based on a user provided cost function. In this implementation, bike distance is calculated on a shortest path based on generalized cost (BIKE_MINMILE) and the walk distance is determined on distance-based shortest path.

Generate External Demand

The ABM system contains two sets of external demand: auto demand and truck demand. The generation of external truck demand is the same as the MIP2 model, where external-external truck demand is generated in three truck classes (light, medium, and heavy) by interpolating demand from two input demand matrices for year 2007 and 2040. Note that the interpolation is embedded in the model system thus, a future year model run would have updated external truck demand matrices.

The generation of external auto demand is a new procedure developed during this project. The model system takes external demand matrices by time periods as inputs. The demand matrices are created beforehand from the California statewide model run. Figure 2 describes the process of creating these external auto trip table for the Fresno model region. Note that a future year run would need to follow this process and create these demand matrices before running the AB model.

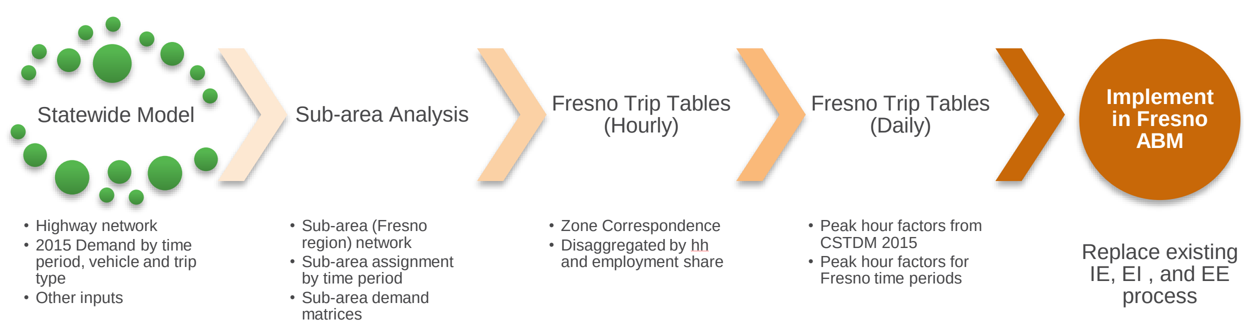


FIGURE 2: EXTERNAL AUTO TRIP TABLE GENERATION PROCESS

The California Statewide Model system includes a highway network, 2015 peak-hour demand by time-period, vehicle and trip type, and other inputs necessary for running highway assignment. In order to generate demand specific to the Fresno region, RSG created a sub-area analysis with the statewide model using a sub-area network developed for the Fresno model region. Each assignment produced a demand matrix for the Fresno region (sub-area) containing demand by the statewide model zones that were identified in the Fresno sub-area network. The matrices contained all four types of demand for the Fresno region: internal-internal, internal-external, external-internal, and external-external. However, the matrices were by statewide model zones and not by the Fresno model zones as required in the Fresno ABM system. The transformation of the demand into Fresno model zones involved disaggregation (or aggregation in some cases) based on a correspondence between the Fresno ABM and the statewide model zones. The process factored[5] the demand based on employment share or population share or combination of both in the Fresno model zone compared to the Statewide model zone. The converted trip tables represent peak-hour demand for the specific time-period, therefore, the trips tables were applied with peak hour factors to scale them to represent demand for the entire duration of the Fresno model time-periods. The peak hour factors were determined by making appropriate adjustments to the original statewide model factors to compensate for differences in time-period durations in the two model systems. The output trip tables, with no internal-internal travel, for the four model time periods are input to the Fresno ABM. As the model system, in addition to the four time periods, runs highway network models for two peak hours (AM and PM), AM and PM peak hour trip tables are also provided as input to the model system. Each input demand matrix contains internal-external, external-internal, and external-external trips by three vehicle classes (drive-alone, shared-ride2, and shared-ride3+).

2.3 NETWORK SKIMMING

The model system generates impedances (skims) for three sets of networks: highway, transit, and non-motorized. The highway and transit skims are (re)generated at the beginning of each feedback loop. The non-motorized skims are generated one-time in the Input Processing step, before the feedback loops.

Highway

The highway skimming process generates 6 skim matrices for two time periods (peak and off-peak) and three auto modes (drive-alone, SR2, and SR3). Each highway skim matrix contains the following impedances in three value of time categories:

- Generalized time (mins)
- Time (mins)
- Distance (miles)
- Cost (dollars)

Note that the highway system reads auto operating cost (AOC) from an input CSV file that contains AOC by scenario years. To maintain consistency between demand and supply models, before each run, the DaySim configuration file is updated dynamically to reflect the same auto operating cost.

[5] Factors are used for the zones internal to the Fresno region and applied at both origin and destination ends. At the origin end, household share is used in AM and MD periods and employment share is used in PM and NT periods. At the destination end, households share is used in PM and NT periods and employment share is used in AM and MD periods.

Transit

The transit skimming process generates 8 skim matrices for two time periods (peak and off-peak), two transit sub-modes (bus and rail), and two access modes (walk and drive). Each transit skim matrix contains the following impedances:

- Transit in-vehicle time (mins)
- Walk access time (mins)
- Walk egress time (mins)
- Transfer walk time (mins)
- Initial wait time (mins)
- Transfer wait time (mins)
- Fare (dollars)
- Boardings
- Transit in-vehicle time for local bus (mins)
- Transit in-vehicle time for express bus (mins)
- Transit in-vehicle time for BRT (mins)

Note that this project made two important enhancements to the existing transit system retained from the VMIP2 model: implementing zone-based fare and adding dwell times to transit stops.

The existing VMIP2 model applies a fixed fare to all transit services (FAX, CLOVIS, and FCRTA) in the Fresno region. However, in reality the FCRTA operated bus routes charge stop-based fare to the riders using the service. The Fresno ABM implemented a zone-based fare system to reflect the actual fare structure for the FCRTA bus services. The new transit system uses two sets of fare systems: fixed fare and zone-based fare. The fixed fare[6] ($1.25) is applied to FAX and CLOVIS operated bus routes, whereas the zone-based fare ($0.75 - $6.75) is applied to the Fresno County inter-city buses operated by FCRTA. The following changes were made to implement the new zone-based fare system in the ABM:

- Added a new field FAREZONE to highway node file and assigned a value from 2 to 31 for stop nodes that are on FCRTA routes and a value of 1 to stop nodes that are not on FCRTA routes
- Assigned the field MODE to 2 for FCRTA routes in the transit line network
- Created a new FARESYSTEM for FCRTA routes
- Created a fare matrix to provide fares by fare zone pairs

[6] The existing MIP2 model has $1.00 as the fixed fare. This value was updated to $1.25.

- Updated Cube scripts for transit skimming and assignment processes

The transit validation exercise discovered that in the VMIP2 transit system, transit travel time on a link was set to the respective auto travel time. This was resulting in unrealistically faster transit times because a transit service generally runs slower than auto traffic due to its nature of making frequent stops for boarding/alighting riders. This project updated calculations of the transit travel times to also include a dwell time at every transit stop[7].

Non-motorized

The new non-motorized skimming process generates two Cube matrices for bike and walk distances by MAZ OD pairs. The walk distances are produced in a text format as well for use in the DaySim buffer tool[8].

2.4 TRAVEL DEMAND

The model system encompasses three sets of travel demand: resident travel, truck travel, and auxiliary travel. DaySim simulates the internal travel for the Fresno residents. Truck travel within the region is produced by the truck model. Auxiliary demand such as external travel (auto and truck) are provided as inputs in the form of trip tables to the model system.

Resident Model (DaySim)

DaySim is an activity-based (AB) travel demand model. DaySim simulates 24-hour itineraries for individuals with parcel-level spatial resolution and minute-level temporal resolution. DaySim can generate outputs at the level of resolution required as input to dynamic traffic simulation. DaySim's forecasts in all dimensions (activity and travel generation, tours and trip-chaining, destinations, modes, and timing) are sensitive to travel times and costs that vary by mode, origin–destination path, and time of day. As a result, DaySim uses the improved network travel costs and times output from a dynamic traffic simulator as inputs. DaySim captures the effects of travel time and cost upon activity and travel choices. These effects are balanced across modes and times of day and are consistent with the econometric theory of nested choice models. DaySim can be used in a distributed manner by running different partitions of the study area population on separate processors and then merging the results.

DaySim comprises several subcomponents and is structured as a series of hierarchical or nested choice models. DaySim's hierarchy places the long-term models at the top of the choice hierarchy and places the short-term models at successively lower levels in

[7] The dwell time values are stored in two new dwell time fields in the transit line file: DWLL_DEFAULT and DWELL_DEFAULT[2] for peak and off-peak periods respectively.
[8] A DaySim input preparation tool that generates a buffered micro-zone file that contains several land-use and accessibility variables.

the hierarchy. The detailed hierarchy and flow through the model is illustrated in Figure 3.

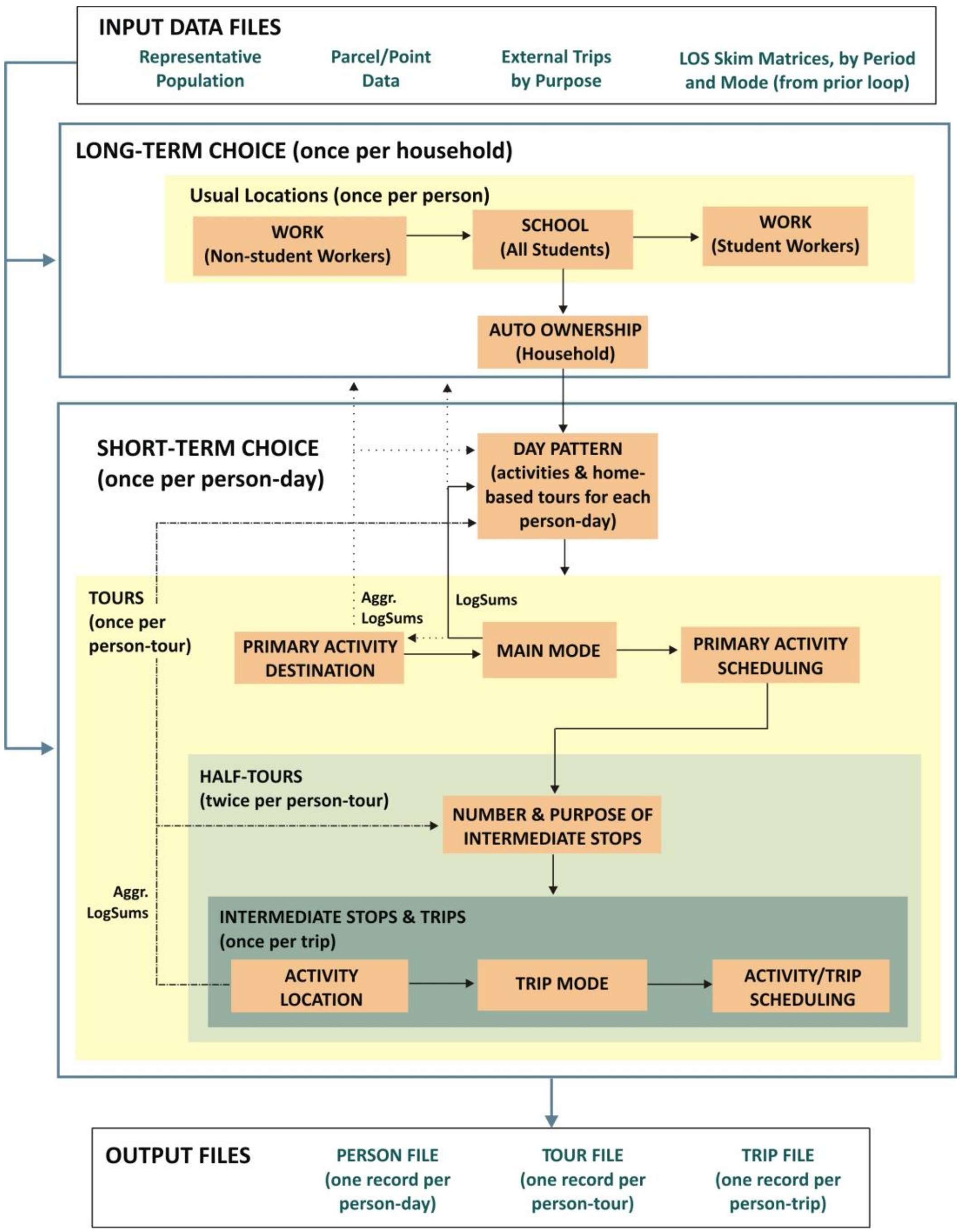


FIGURE 3: DAYSIM SUB-MODELS

The flow generally proceeds in a linear fashion from the long-term models to the short-term models. The choices made in the long-term models influence or constrain the choices made in the lower-level models. For example, household auto ownership affects both day pattern and tour (and trip) mode choice by including auto ownership variables in those component models. In addition to these direct influences, utilities from lower-level models flow upward to higher-level models. "Logsums" (expected utilities) from tour destination and tour mode choice models affect other short-term models and also affect the upper-level, longer-term models. Some of the logsums from lower-level models are aggregated for use in the long-term models in order to reduce the computational time required to use fully detailed disaggregate logsums in a complex nesting structure.

Buffering and Transit Access Preparation

In DaySim, it is important to have measures not only of within a particular micro-zone, but also what lies in the area immediately surrounding each micro-zone. These measures are created by defining a "buffer" area around each micro-zone and counting what lies inside the buffer. These variables can then be used in DaySim, similar to how zonal land use and density variables are used in TAZ-based models, with the advantage that the buffer is defined in exactly the same way for each micro-zone. The buffer variables that DaySim uses include:

- The number of households in the buffer;
- Employment (number of jobs) in the buffer in various employment sectors;
- Enrollment in schools in the buffer, segmented by grade schools and colleges;
- The number of spaces and average price of paid off-street parking in the buffer;
- The number of transit stops within the buffer (segmented by sub-mode, if relevant);
- The number of street intersections in the buffer, segmented by 1-node (dead-end or cul-de-sac), 3-node (T-junction), and 4+node intersections; and

In addition, DaySim also uses the distance from the micro-zone centroid to the nearest transit stop (by transit sub-mode, if relevant), as well as the distance to the nearest open space area while simulating models.

DaySim Buffering Tool

A tool, Buffer Tool, to perform the buffer calculations is implemented in the DaySim component of the model system. This tool calculates all the buffer and transit access variables that DaySim needs, using the following inputs:

- Base micro-zone file
- Street intersections file
- Transit stops file
- Network nodes file (for all streets network based short trip distances)
- Node-to-node shortest path distance file (for all streets network based short trip distances)

Note that it is essential that the buffer measures used in application are consistent with those used for the original model estimation. Thus, when preparing new buffer measures, users should not modify the settings in the buffering tool control file.

Buffer Calculations

As mentioned earlier, buffer variables for a micro-zone are calculated by summing land use variables of all micro-zones within a certain buffer distance of the particular micro-zone. In the past, buffer calculations have used a “flat” buffer, using a certain radius, e.g. ¼ mile, and counting everything within that radius and nothing outside the radius. That approach is simple, but has the disadvantages that (a) it weights all opportunities within the buffer the same, whereas in reality the land use that is very close by will tend to have more influence on behavior than the land use at the edge of the buffer, and (b) there can be “cliff effects” if a large development is located near the edge of the buffer. In the latter case, the measures become sensitive to the somewhat arbitrary specification of the buffer size, and to the rules used to deal with micro-zones that straddle the buffer boundary. This tends to add random “noise” to the buffer measures.

The buffering tool can be set to use flat buffering, or a distance-decay buffer, in which each buffered item is weighted according to the distance from the origin micro-zone centroid. There are two options provided for the weighting function: a logistic function and a negative exponential function. The logistic form is recommended because its shape is more representative of typical behavioral models that use logistic functions. The Fresno ABM uses a distance-decay buffering.

The buffering program simultaneously calculates all the buffer variables for two different buffer sizes. The reason for this is that the DaySim choice models use smaller buffers for some variables (e.g. those that represent attractiveness of walk trips), and larger buffers for some other variables (e.g. those that represent attractiveness of bike trips, or more general neighborhood effects).

For distance decay buffering, the user specifies three parameters for each buffer: (1) the distance parameter, (2) the offset parameter, and (3) the slope parameter (the latter two are used only for logistic buffering). The parameters and equation for the logistic curves used for DaySim model estimation and calibration are listed below. It is necessary that these same parameters be used for model application.

Parameter	**Buffer 1**	**Buffer 2**
Inflection	BDIST1 = 660 (ft)	BDIST2 = 1320 (ft)
Offset	BOFFS1 = 2640 (ft)	BOFFS2 = 2640 (ft)
Slope	DECAY1 = 0.76	DECAY2 = 0.76

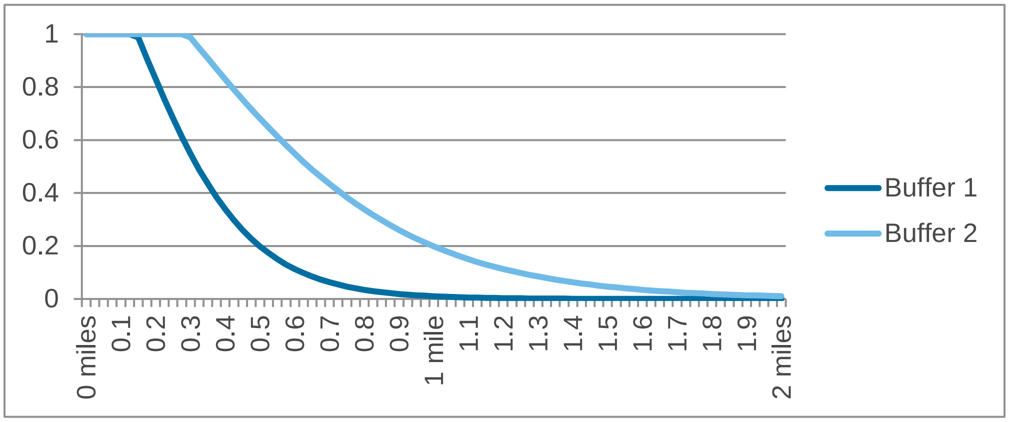
The equation is:

Weight = MIN(1, (1 + Exp(DECAY – 0.5 + BOFFS/5280))

/ (1 + Exp(DECAY * (Distance/BDIST – 0.5 – BOFFS/5280))))

Distance is the distance, in feet, from the origin micro-zone to any other micro-zone whose calculation is explained in the next paragraph..

FIGURE 2.4. BUFFER1 AND BUFFER2 DISTANCE DECAY WEIGHTS



The buffering program also gives the user three options as to how distances are calculated within the buffering program:

1. Use crow-fly distance between the XY coordinates
2. Use interpolation with a "circuity surface" around each micro-zone.
3. Use shortest path distance between the nearest all street network nodes.

Note that in option 1, because the buffer distance is calculated using XY coordinates from centroid to centroid for micro-zones, buffering may not be very accurate for micro-zones that are very large compared to the size of the buffer.

Option 3, used in Fresno ABM, provides most accurate distances that take into account obstacles and directness in the street network and is preferable if the required data exists. The following steps are involved in buffering using distance decay weights and all streets network distances:

1. The buffering tool first associates each micro-zone with the nearest network node and creates a micro-zone-node correspondence.
2. Multiple micro-zones may be associated with a single node and so the base micro-zone file is reduced to node level by aggregating data from all micro-zones that are associated with the same node.
3. Other items such as open spaces/parks and transit stops are also associated with the closest network nodes.
4. At this point, buffering calculations are all done at the node level since node-to-node all street network distance are available. For node pairs that are not within 3 miles of each other, Euclidean distance calculated from XY coordinates is used. Buffer variables for a particular node are calculated by obtaining the

weighted sum of land-use variables of all the nodes with the chosen buffer distance. The calculation of distance weights has been described earlier.

5. Once the buffer calculations at the node level are complete, the buffer variables are transferred to micro-zones using the micro-zone-node correspondence created in step 1.

It should also be noted that in case of option 3, during the buffering process, two binary files that have information about node-to-node network shortest path distances are output so the DaySim can use them for simulation of short trips.

The following steps are involved in buffering using distance decay weights and XY/Euclidean distance:

1. Calculate distance weights using the logistic decay equation described earlier.
2. Calculate buffer variables for each micro-zone by counting land-use attributes of the surrounding micro-zones by getting their centroid distances (Euclidean) from that of the micro-zone under consideration and weighting by the corresponding distance weights.

Population Sampler

A common problem in the application of activity-based models to small urban areas is the effect of Monte Carlo simulation variance on the interpretation of results. The simulation error is inversely proportional to the number of decision-makers or agents in the model. For smaller regions like Fresno COG and for cases where the policy of interest is relatively small scale (e.g., a capacity enhancement on a low-volume facility), the simulation error can exceed the sensitivity of the model, confounding the analysis of benefits. To eliminate this problem, the project applied a tool “Population Sampler” that was originally developed for the Bellevue-Kirkland-Redmond region DaySim model to enable oversampling households by geography. Population Sampler reduces the number of feedback iterations required for full model convergence by replicating households whose results are then averaged prior to assignment. The tool is implemented in a Python script and requires specifying the sample rates by zone along with a distribution of households by income and size to ensure that the sampled households are representative of the actual population. For Fresno ABM, the population sampler oversamples households by a factor of 3.0.

DaySim Inputs

This section provides a brief description of inputs required in DaySim model[9].

[9] DaySim is an open source activity-based model stored on a GitHub repository: https://github.com/RSGInc/DaySim. The repository has a wiki page with extensive documentation describing the model.

Micro-zones

The Fresno activity-based model system uses micro-zones as the fundamental spatial unit for generating travel demand. Use of micro-zones improves the sensitivity of the model system to land use, fine-grained urban form and accessibility attributes. However, use of these detailed measures necessitates the preparation of more detailed and larger model input datasets. The micro-zone data input file contains fields that describe the quantities of households, school enrollment by type and employment by industrial sector within quarter mile and half mile buffers. Note that these buffers are based on “all streets” based network accessibilities and employ decay functions that weight closer opportunities more than distant opportunities. In addition, the micro-zone file contains information about urban form and the transportation system on and close to the micro-zone, such as the number of dead end streets, the proximity of the micro-zone to transit stops, and the price and supply of nearby parking.

FresnoCOG provided a shapefile of micro-zones in the model region with all the necessary attributes. The shapefile is then converted into a format required in DaySim. Appendix C describes steps to do this conversion.

Synthetic Population

The synthetic population is comprised of lists of households and persons that are based on observed or forecasted distributions of socioeconomic attributes and are typically created by sampling detailed Census microdata. These lists function as the basis for all subsequent choice-making simulated in the activity-based model. This project implemented a population synthesizer called “PopSynIII” (originally developed for MTC) to generate synthetic population that are input to the Fresno ABM. Details of the procedures and instructions to run the software for a base year (2014) are provided in Appendix A.

Worker IXXI Fractions

Although the modeling area is defined in such a way as to capture as much “internal” travel by regional residents as possible (that is, travel with both origins and destinations with the modeling area), a certain portion of observed regional travel involves either regional residents travelling to destinations outside the modeling area or people who are not regional residents travelling to destinations within the modeling area. As in a traditional trip-based travel demand model system, these travel markets are typically incorporated into the model through the use of internal-external trip tables, which may be either fixed or dynamic.

A distinguishing feature of the DaySim activity-based model system is that, due to the spatial and behavioral detail embedded in the model, it is sensitive to how this internal-external travel affects the choices made by regional residents. A particular focus of this detail is on ensuring that the right numbers of workers are “out-commuting” to employment locations outside the modeling area, and that the right number of regional jobs are being consumed by non-residents “in-commuting” to the region. At present, this is accomplished by using a file (worker IXXI fractions) that contains TAZ-based shares of workers who are in-commuting and out-commuting, which is provided as an external input to the DaySim model system. The shares either can be held fixed or may be updated by deriving updates shares from the trip-based model outputs.

TAZ Indexes

The TAZ index file enables users to flexibly define non-continuous zones numbering systems, and to identify the availability of external and other zones as destination choices, without impacting DaySim performance.

PNR Nodes

The PNR file provides park and ride locations with corresponding capacity and parking cost. The file is prepared by extracting PNR locations from the highway network (node file). For each location, capacity is set to 100 and cost to 0.

Coefficients

A coefficient file provides a list of variables used in the model and corresponding coefficient values and t-statistics. Each Daysim model component is associated with a coefficient file. For the Fresno ABM model, the model coefficients were borrowed from SACOG model and later calibrated to match observed survey data for the Fresno region.

Roster

A key set of inputs to any travel demand forecasting model system are the files that contain the scenario, mode, user-class, and time period-specific measures of network impedance, often referred to as network “skims.” The roster provides users with the ability to flexibly specify the skims that are associated with the different mode, time period and user classes used in the Fresno activity-based models system, without necessitating changes to the core DaySim model code. For example, a user may want to increase the number of time periods used in the model system to better reflect changes in network impedance by detailed time-of-day. In order to implement such an enhancement, a user would first revise the Cube-based network-processing scripts in order to generate the desired skims and would only need to revise the DaySim impedance roster to make DaySim sensitive to this additional detail.

Roster Combinations

The "Roster Combinations" file gives the possible mode/path type combinations used in DaySim. The file has columns that enumerate the 9 modes used in the current model system (walk, bike, SOV, HOV2, HOV3, transit, park-and-ride, school-bus, other) and 7 rows that enumerate the path types currently used (full-network, no-tolls, local-bus, light-rail, premium-bus, commuter rail, ferry). The path type “ferry” is used for BRT. The cells are TRUE for valid combinations within DaySim and FALSE otherwise.

Configuration

The configuration file is the main control file for DaySim. The configuration file informs DaySim about inputs/outputs and various model settings. These settings include input/output file names, types and locations, sample rates, DaySim pathbuilding weights, and also allow users to specify which DaySim model components should be executed.

The ABM system dynamically generates a configuration file before a DaySim run. The model system uses two sets of DaySim configuration files: one for running only work and school location choice models and the second for running all choice models in DaySim. In each feedback loop, DaySim is run for four iterations. The first three iterations run only work and school location choice models and, in the end, generate stabilized shadow prices. These shadow prices are seed to the fourth DaySim iteration that runs all choice models to produce resident travel in the region.

Shadow Prices

The “shadow_prices.txt” is used to constrain employment and enrollment in a parcel by its actual capacity. Similarly, “park_and_ride_shadow_prices.txt” is used to constrain parking at park and ride locations’ capacity. The shadow prices are intermediate outputs of a DaySim run. They are optional as inputs. However, it is advisable to have starting shadow prices in order to get stable results. Also, having fixed starting shadow prices is helpful in replicating an ABM run with the same inputs

Truck Model

The truck model is retained form the existing VMIP2 model system. The truck model produces a trip table containing truck demand in three truck classes: small (TS), medium (TM), and heavy (TH).

Auxiliary Demand

The auxiliary demand consists of demand external to the model region. Specifically, it consists of external auto demand and external truck demand. The external auto demand for internal-external (IE), external-internal (EI), and external-external (EE) trips by time-period are provided as input trip tables to the model system. The trip tables are created from the California Statewide model as described in Section 2.2 Input Preparation.

The external truck demand is generated in the same way as the VMIP2 model. The truck demand is interpolated for the scenario year using two input demand matrices for year 2007 and year 2040. The interpolated trip table contains external truck demand into four demand classes for distance (short or long) and truck category (medium or heavy): short-medium, short-heavy, long-medium, and long-heavy.

2.5 ASSIGNMENT PREPARATION

Before performing network assignments, the travel demand in the model system are prepared in formats required in the assignments. Specifically, the resident travel from DaySim is converted into vehicle trips and segmented into highway and transit travel in respective assignment classes. Other components of the model demand, internal truck trips and auxiliary demand, are added with the resident travel demand to create time-period specific Cube matrices.

The resident model, DaySim, generates a list of person trips for the residents and outputs them into a trip file in tab separated values format (_trip.tsv). The person trips are segmented into four time periods (see Table 3) using a trip time calculated based on trip's position in the corresponding tour chain. The DaySim trip file contains this information in variable "HALF", which takes value as 1 or 2, indicating if a trip is in first half of the tour or in the second half respectively. If a trip is present in the first half (leg) of the tour (HALF=1), then the trip time is set to trip's arrival time. Otherwise (HALF=2), trip's departure time is considered as the trip time.

TABLE 3: HIGHWAY ASSIGNMENT TIME PERIODS IN THE FRESNO ABM

TIME PERIOD	DURATION
AM	6 am – 9 am
MD	9 am – 4 pm
PM	4 pm – 7 pm
EV	7 pm – 6 am

Highway

Person auto trips generated by the resident model (DaySim) are converted into vehicle trips using vehicle occupancy factors for high occupancy auto modes (SR2 and SR3). Table 4 presents the occupancy factors used to convert DaySim person trips to vehicle trips.

TABLE 4: OCCUPANCY FACTORS TO CONVERT PERSON TRIPS TO VEHICLE TRIPS

MODE	OCCUPANCY
Drive-Alone (DA)	1
Share-Ride 2 (SR2)	2
Shared-Ride 3+ (SR3)	3.5

The DaySim vehicle demand are combined with truck demand and the auxiliary demand to construct four time-of-day demand tables for highway assignment. The demand tables are generated in Cube matrix format and contain demand in 13 assignment classes – 9 auto classes (3 auto classes x 3 value of time), one external class (EE), and three truck classes (small, medium, and heavy). Three auto classes are DA, SR2, and SR3.

Transit

Daysim produces transit trips across all times of day in which transit service is provided. The DaySim trip file (_trip.tsv) reports all transit trips as transit and do not classify them by access mode (walk or drive). However, another DaySim output, tour file (_tour.tsv), details tours made by a person and reports transit tours by access mode as well. DaySim calculates trips from the tours made by a person. Therefore, if the tour corresponding to a trip is reported as drive-to-transit than the trip is a drive-to-transit too. To get the tour information to the trips, tour and trip files are joined by a common identifier (tourid). Presently in the Fresno model, no KNR sub-mode is included in DaySim, thus, all drive-to-transit trips are reported as PNR trips.

The transit trips are aggregated in OD format into two Cube matrices by transit mode (Bus and Rail). Each matrix contains demand in four assignment classes by time-period (peak and off-peak) and access mode (walk and drive). Note that the model base scenario (2014) have only bus mode in the transit system, thus all transit demand is put into the bus mode specific demand matrix. However, the rail mode specific matrix is setup to use as placeholder for future BRT line in the Fresno region.

Non-Motorized

DaySim also produces list of trips made for non-motorized modes. These trips are aggregated by MAZ OD format and generated a cube matrix that contains demand by two non-motorized modes (walk and bike).

2.6 NETWORK ASSIGNMENT

The model system assigns three sets of travel demand to the model network: highway, transit, and non-motorized. Note that highway and transit assignment procedures are retained from the VMIP2 model system. The project implemented the non-motorized components in the ABM system.

Highway

The highway assignment is a multi-user class equilibrium assignment that builds paths on a generalized cost function with the cost function differentiated by user class. The generalized cost is a function of travel time and cost (toll) and is converted to time using value of time.

The assignment employs a user equilibrium method that is an iterative process to achieve a convergent solution where no travelers on the roadway network can improve travel-times by shifting routes. Throughout each of these iterations, Cube computes network-link flows, which incorporate link-capacity restraint effects and flow-dependent travel-times.

Similar to the MIP2 model, the ABM system runs highway assignment for two time periods (AM and MD) in every feedback loop and for four time periods (Table 3) in the final assignment step. Note that the truck trips are assigned as PCEs by applying factors of 1.5 and 2.0 for medium and heavy trucks respectively. As convergence criteria, the assignments use a gap value of 0.00001 and maximum iterations of 50 for the peak and 20 for the off-peak periods.

Each converged assignment run produces a loaded network (*.NET) that contains assigned flows on the links. The four time-period specific loaded networks are then post-processed and combined into one loaded network containing flows for all four time periods. The post process also exports the flow data into a DBF file.

Transit

The model system uses Cube-based Pathfinder assignment to assign transit demand on the model transit network. The transit demand are assigned by time period (peak and off-peak), sub-mode (bus and rail), and access mode (walk and drive). Note that presently the rail mode is used as placeholder for future BRT line.

Each transit assignment run produces a loaded network file (*.NET) and a DBF file that contain on and off at every transit stops in the model system. In all, the transit assignment step generates 8 files for each of the two outputs – two time periods (peak and off-peak), two sub-modes (bus and rail), and two access modes (walk and drive).

Non-Motorized

This project implemented a new assignment procedure to assign non-motorized demand to a network. For both bike and walk, the daily demand is assigned to a non-motorized network created in the Input Processing step. The bike assignment calculates best path using a generalized cost based on bike facility type, whereas the walk assignment is setup to directly use distance as a generalized cost. As the non-motorized assignments are not capacity restraint, only one iteration is run in each of the two assignments.

Each non-motorized assignment run produces a loaded network with flows on the network links. However, as described in Section 2.2, the non-motorized network contains the re-numbered node ids that are required for skimming and assignment by MAZ. Therefore, the loaded networks are processed to set the node ids back to the original all-street network node ids and also the flows in the new loaded networks are exported into DBF files.

2.7 FEEDBACK

Feedback is used in two primary ways in the Fresno activity-based model system: between iterations within the highway assignment process, and between “system” iterations of the model system in which both the DaySim activity-based demand components and the Cube-based network supply components are executed. The feedback provides the next iteration new (congested) travel times calculated based on the previous iteration results.

2.8 REPORTING

This project developed several stand-alone report processes to summarize outputs of DaySim and network assignments. These reports are useful to assess model performance and inform model calibration and validation, as well as other analyses.

DaySim

A DaySim run produces multiple outputs describing daily travel for the residents of the model region. The primary outputs are:

- _household.tsv
- _household_day.tsv
- _person.tsv
- _person_day.tsv
- _tour.tsv
- _trip.tsv

A separate set of R scripts summarize these outputs in various excel spreadsheets. The spreadsheets contain tables and charts summarizing different sub-models in DaySim. The spreadsheets are also utilized during calibration and validation of DaySim. Appendix D describes the process of setting up and running these R scripts.

Highway Assignment

The outputs of the highway assignment are time-period specific flow tables, which are, during a post-process, combined into one flow table. This flow table can be used to report various highway statistics.

This project created a highway validation spreadsheet that takes a combined highway assignment flow table[10] as input and automatically generates various highway statistics and compares them with the same statistics from observed data (traffic counts). The statistics include:

- Scatter plot of estimated flows and observed counts
- Gap by facility type (also by time of day)
- Gap by volume group

[10] A python script generates a combined assignment flow table that is input to the highway validation spreadsheet.

- Root mean squared error
- Gap by screen line

Gap is calculated by taking the difference of estimated traffic volume with the observed count and then dividing it by the observed count.

Appendix E provides step-by-step instructions for updating the highway validation spreadsheet.

Transit Assignment

The transit assignment outputs transit boarding tables by period, transit sub-mode and mode access. A python script processes the assignment outputs and reports a summary of daily boardings by transit route. The summary is in a CSV format and is input to a transit validation spreadsheet that generates the following statistics:

- Scatter plot of estimated boardings and observed boardings
- Boarding rate
- Boardings by route

Appendix E provides step-by-step instructions for updating the transit validation spreadsheet.

Non-motorized Assignments

Similar to the highway assignment, the non-motorized assignment outputs walk and bike flows on network links. Two separate validation spreadsheets summarize and compare estimated non-motorized flows with observed counts. Instead of validating every count location separately, the locations are combined and compared in 13 count groups. The groups are created based on the direction of travel and the neighborhood of the location.

Appendix E provides step-by-step instructions for updating the bike and walk validation spreadsheet.

3.0 CALIBRATION & VALIDATION

A calibration process adjusts the model to ensure that the model generates demand that reasonably follow the behavior depicted in observed data. The demand is defined as frequency of trips by origin and destination pair. The frequency of trips by OD pair can have different segmentation for ex. trip mode, time of day etc. The demand are then loaded on to network (assignment) to determine frequency of trips using each link in the network. For highway, this provides vehicle flows on every link (road) in the highway network and for transit, generates number of people (boardings) using each transit service.

After a model is calibrated to produce demand that reasonably predict observed travel behavior in the region, the model is validated to ensure network-level usage of the demand. The model validation includes, on the highway side, comparing estimated traffic volume from the model with observed traffic counts, and on the transit side, comparing estimated transit boardings from the model with observed transit ridership.

The rest of this chapter presents model calibration and validation in separate sections. For each, first, the observed data are discussed followed by summaries from a final calibration and validated model run. In the end, a summary of the chapter presents key takeaways from the discussions.

3.1 CALIBRATION

In model calibration, alternative-specific constants (ASCs) and other model parameters are iteratively adjusted until the model generates demand that reasonably matches travel patterns in observed data. Typically, models are calibrated according to the following procedure: first, create comparisons between observed data and estimated model results. Next, calculate ASC adjustments by calculating natural log of the ratio between the observed value and the estimated value for each alternative. Then, add the adjustments to the ASCs from the previous iteration. Next, run the model with the updated constants.

Data

Table 5 presents a list of datasets utilized to calibrate the Fresno ABM.

TABLE 5:MODEL CALIBRATION DATASETS

DATASET	YEAR	SOURCE	PURPOSE
California Household Travel Survey (CHTS)	2012		Fresno – Tour Destination, SJV – Other Sub-models
National Household Travel Survey (NHTS)	2009		Fresno - Tour Destination
Transit On-Board Survey	2014	Transit On-Board Survey Program	Transit Tours/Trips
Census Transportation Planning Product (CTPP)	2010	Census	Workers Flow
Longitudinal Employer Household Dynamics (LEHD)	2014		Workers Flow
Streetlight	2014	Streetlight	Internal-External Trips

The present effort used multiple datasets to calibrate the Fresno ABM. The 2012 California household travel survey (CHTS) and the 2009 national household travel survey (NHTS) are the primary datasets used during the calibration. Due to limited sample records within the Fresno region, the calibration also utilized survey records of the entire San Joaquin Valley (SJV). As tour destinations require geocoding of destination locations to Fresno TAZ/MAZ, the calibration of tour destination choice model used survey records that are for the Fresno residents only. Among the two survey datasets, the CHTS is the main model calibration dataset. The NHTS dataset is used as secondary dataset for calibrating tour lengths only. The two datasets (CHTS and NHTS) report different observed tour lengths, therefore, the model calibration tried to match estimated tour lengths to an average value that is somewhere between the values from the two datasets.

Whenever available, additional datasets are utilized to inform more accurate information for particular type of travel. For example, transit survey provides information about transit travel, thus, informed transit travel targets (tours and trips) in mode choice calibration. Also, the 2010 CTPP ACS Journey to Work data and 2014 LODES data provide flow of workers and used to validate estimated work location choice of Fresno residents.

As mentioned before, the external auto travel (IE, EI, and EE) are input trip tables to the model system. The external trip tables are constructed from the California Statewide model demand and are adjusted, if required, based on comparisons with an observed data source, Streetlight, that provided origin-destination GPS-based travel data.

California Household Travel Survey (CHTS)

The 2010-2012 California Household Travel Survey (CHTS) was a collaborative effort lead by Caltrans to gather travel information needed for regional and statewide travel and environmental models using the same instrument and methods across the state. The study was jointly sponsored and funded by Caltrans, the California Strategic Growth Council, the California Energy Commission, and 8 local transportation agencies, including Fresno Council of Governments (Fresno COG).

The CHTS was designed to collect travel information from households in all of California's 58 counties for the year 2012. NuStat's obtained detailed travel behavior information from over 42,500 households, using multiple data collection methods, including computer assisted telephone interviewing (CATI), online, mail surveys, wearable and in-vehicle GPS as well as using on-board diagnostic (OBD) sensors. Out of the completed households, 36,714 were non-GPS households and 5,717 were GPS households.

The collected sampled data was weighted and expanded by first, calculating weights at household and person levels. Then travel of GPS households are used to apply trip correction factors to non-GPS households. The final weights were developed at the county level, but demographic controls were balanced at the statewide level only. Also trip correction factors developed at the statewide level only. During a separate work, Fehr and Peers re-weighted the CHTS data for the Fresno region[11]. This re-weighted data is used in the ABM calibration.

The model calibration further processed the CHTS data to prepare targets for model calibration. First, the data are filtered to keep only weekday travel by removing weekend and holiday travel. Then the travel weights are scaled appropriately to match total travel in the original survey data.

During preparing calibration targets, household and person level summaries use respective weights. Trip level summaries use trip weights that are calculated by multiplying corresponding household weights with trip correction factors. Tour summaries use tour weights determined by calculating weighted average of trip weights within a tour.

The CHTS data are geocoded for origin (home, tour, and trip) and destination locations (work, school, tour, and trip) to assign corresponding MAZ and TAZ in the Fresno model region.

Due to limited sample records within the Fresno region, the model calibration, whenever possible, utilized survey records for the residents of the entire San Joaquin Valley (SJV). As tour destinations require geocoding of destination locations to Fresno TAZ/MAZ,

[11] See Technical Memorandum: Cleaning and Weighting of California Household Travel Survey Data, From Jennifer Ziebarth, Fehr & Peers, To Users of CHTS data, June 23, 2015, and Weighting Report for Fresno County, Jennifer Ziebarth, Fehr and Peers, June 19, 2015

calibration of the tour destination choice model used survey records for the residents of Fresno County only.

National Household Travel Survey (NHTS)

The 2009 National Household Travel Survey collected travel behavior of sampled households in the nation. In a separate effort, the NHTS dataset was filtered to keep travel records only for the residents of the San Joaquin Valley (SJV) and then processed to prepare in formats similar to DaySim outputs.

The model calibration uses the NHTS dataset as a secondary dataset for calibrating the tour destination choice model. The NHTS data in DaySim format are further processed to retain records only for the residents of Fresno County. The travel origin and destination locations within the model region are then geocoded to assign corresponding MAZ and TAZ.

Transit On-board Survey

A transit on-board survey was conducted in 2013 across 16 Fresno and Clovis transit routes, collecting data from approximately 11k transit trips. RSG coded DaySim trip purpose in the data, and inferred DaySim tour purpose from trip purpose using Monte Carlo simulation based on probabilities from other transit on-board surveys in which it was possible to identify both tour purpose and trip purpose (Figure 5). RSG coded mode of access and egress to/from transit, consistent with DaySim, based on reported access egress modes (approximately 82% walk-access, 9% park-and-ride, 9% kiss-and-ride). Unfortunately, number of autos owned by the household was not available in the survey (only auto availability for the trip was asked). The raw data was not expanded to observed boardings, so RSG expanded the data to represent total boardings at a route level based on observed ridership data, and calculated total linked transit trips by dividing the boarding factor by 1+ the reported number of transfers. The expanded data has approximately 26k linked transit trips in Fresno County.

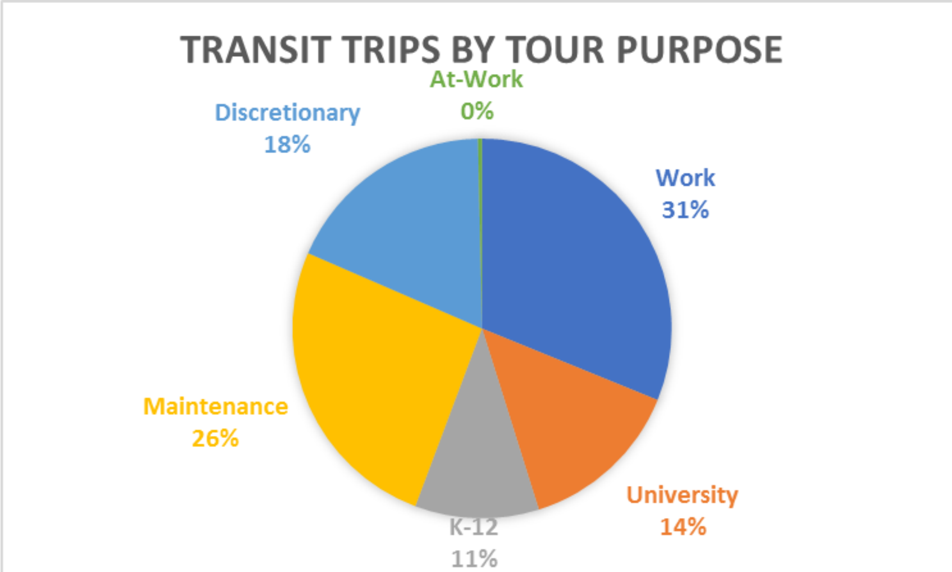


FIGURE 5: TRANSIT TRIPS BY TOUR PURPOSE

Worker Flows

Commute travel is a very important and significant component of any regional travel. Therefore, in addition to verifying number of workers by work location, it is essential to validate workers by their origin (home) and destination (work) locations. This project obtained multiple observed datasets and validated workers flow between home and work districts. The flows are compared at an aggregate level (10 districts) since observed data are likely to show more variance/error at a more detailed geography.

In addition to the CHTS and the NHTS, two other datasets complemented the validation of workers flow: the 2010 CTPP ACS Journey to Work and the 2014 LODES Workers Flow. The datasets are obtained from their respective websites.

Streetlight Data

The project purchased location-based survey (LBS) data from Streetlight for external stations in Fresno County. LBS data is derived from GPS-enabled Smartphone apps that track spatial information (locations and times) when the app is running. Streetlight processes the data to generate trips and makes the data available through a convenient web application. It is not possible to identify the traveler name, exact origin\destination, route trajectory, or other details from the data. To preserve anonymity, the data is only available for user-defined or preset geographic areas and it is aggregated by time-of-day and day-of-week. Streetlight checks user-defined geographies to ensure that they are sufficiently-large to preserve anonymity. The smallest preset geography available is the Census block group. Trips between geographic areas by time of day and day of week are summarized in relative terms (the “streetlight index”).

An origin-destination (OD) analysis extracts trips ("streetlight index") between a set of origins and destinations. To cover all flows for the Fresno region, each set included both internal and external zones. The Streetlight data contract allowed only up to 50 zones for an OD analysis, suggesting a maximum of 25 zones for each of the origin and destination set. To meet this limitation, external and internal zones (census block groups) were aggregated to have total 25 zones for the model region - 18 external stations and 7 internal districts. The internal districts are created by dividing the Fresno County in 7 polygons and the external stations are represented by constructing polygons around streets serving traffic in and out of the region. With this, the OD analysis of the 25 zones provided four set of flows for the Fresno region: internal-internal (II), internal-external (IE), external-internal (EI), and external-external (EE). The analysis data were daily average of Tuesday to Thursday and for spring (March, April) and fall (September and October) in 2017. Note that instead of the base year (2014), the analysis chose the latest year (2017) to have more sample records thus better estimates of daily flows.

The extracted data represents trips in relative terms ("streetlight index") and not actual vehicle trips between an OD pair. The streetlight index is converted into vehicle trips by scaling the flows using actual traffic counts (Caltrans counts and Fresno counts) at external stations. The scaling factored OD flows to match row (origin) and column (destination) totals. Due to the way Streetlight calculates streetlight index, separate scaling is performed for personal and commercial OD flows.

The model results are compared with these scaled vehicle flows from the streetlight data and the internal-external trip tables input to the model system are adjusted accordingly.

DaySim Calibration Summaries

As mentioned earlier, an R-utility summarizes DaySim outputs into statistics that are meaningful and easy to understand. The summaries are prepared by key model components and include work and school location, auto ownership, day pattern, tour/trip destination choice, mode choice, and time of day. These summaries from the final calibrated model are presented below.

Note that the tour/trip destination summaries are compared with the CHTS and NHTS data for the Fresno region only. The other summaries are compared with the CHTS data for the entire San Joaquin Valley (SJV).

Synthetic Population

Table 6 and Table 7 compare synthetic population in the ABM with observed data (CHTS for the SJV and ACS 5-year 2011-2015 for the Fresno County). The population includes both households and Group Quarters. Generally, the synthetic population in the ABM follow the distribution in the Census. Note that compared to the Census data, the synthetic population includes slightly more full-time workers. This was a result of intentional increase in workers population to match total workers in the Census data. Despite the increase, the total workers (full-time and part-time) in the synthetic population (346,258) are still smaller than the Census data (350,551). The CHTS data shows significantly higher share of full-time workers (39%) compared to the Census data (31%) or the ABM (32%).

TABLE 6: POPULATION BY PERSON TYPE

PERSON TYPE	CHTS (SJV)	CENSUS (2011-2015)	ABM
Full Time Worker	1,356,197	291,960	294,596
Part Time Worker	180,629	58,591	51,662
Retired	287,599	88,486	78,389
Non-Worker	646,774	175,781	158,730
University Student	110,156	64,376	46,690
Student Age 16+	141,985	29,117	41,270
Student Age 5-15	565,585	169,007	172,934
Kid under 5	221,045	79,309	81,679
Total	**3,509,970**	**956,627**	**925,951**

TABLE 7: POPULATION BY PERSON TYPE

PERSON TYPE	CHTS (SJV)	CENSUS (2011-2015)	ABM
Full Time Worker	39%	31%	32%
Part Time Worker	5%	6%	6%
Retired	8%	9%	8%
Non-Worker	18%	18%	17%
University Student	3%	7%	5%
Student Age 16+	4%	3%	4%
Student Age 5-15	16%	18%	19%
Kid under 5	6%	8%	9%

Total	**100%**	**100%**	**100%**

Home to Work Distance

As presented in Table 8, the CHTS data indicate an average home to work distance of 9.1 miles regionwide. For the same, the NHTS data suggests a lower value (7.4 miles). The ABM is calibrated to a distance (9.2 miles) close to the higher observed value (CHTS). The distance by worker type are also reasonably calibrated between the two observed datasets. Note that the home to work distances are only for the Fresno County residents that have work location within the Fresno County.

TABLE 8: AVERAGE HOME TO WORK DISTANCE

WORKER TYPE	CHTS (FRESNO)	NHTS (FRESNO)	ABM
Full Time	9.4	8.3	9.4
Part Time	7.8	4.8	8.7
Other	8.2	5.7	6.5
Total	**9.1**	**7.4**	**9.2**

Figure 6 shows a distribution of home to work distances of individuals. The X-axis is distance in miles and the Y-axis is share (%) of the total workers. Due to relatively lower samples, both observed datasets show lumpy distributions. The ABM distribution is smoother and generally follows the two observed distributions from the CHTS and the NHTS.

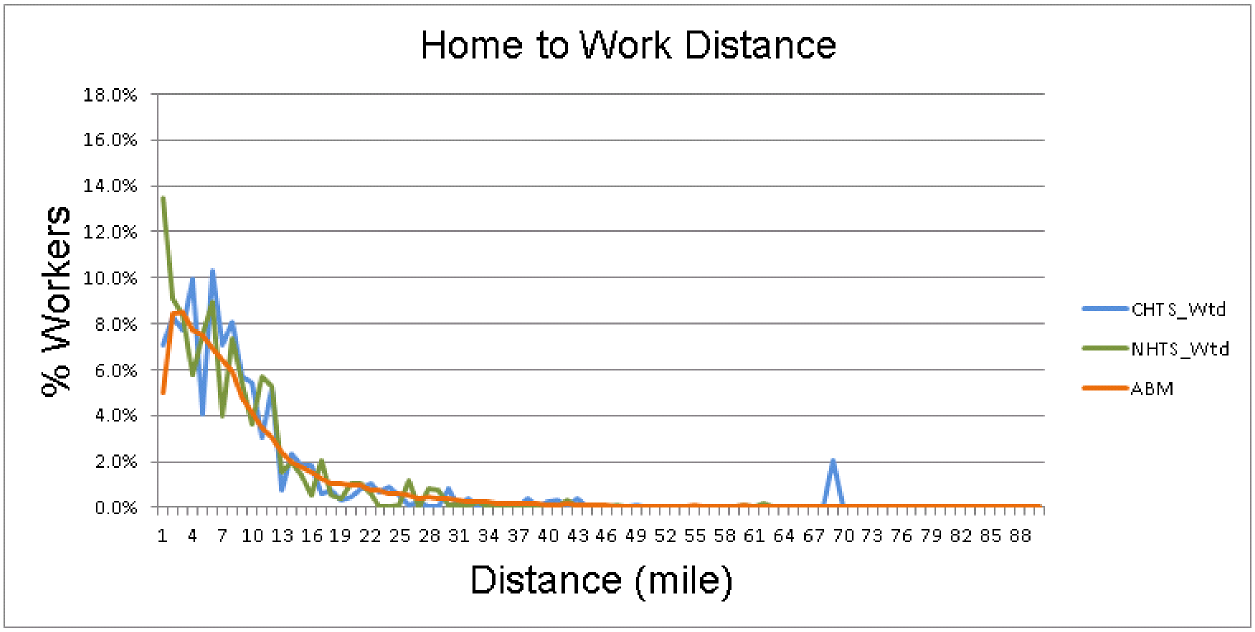


FIGURE 6: DISTRIBUTION OF HOME TO WORK DISTANCES

A comparison of estimated work locations (workers working at a location) and corresponding employment by TAZ, Figure 7, shows a good match with a R-squared value of 0.99. This suggest that the shadow pricing mechanism in DaySim is working well. The shadow pricing iteratively balances number of workers and total employment in a micro-zone.

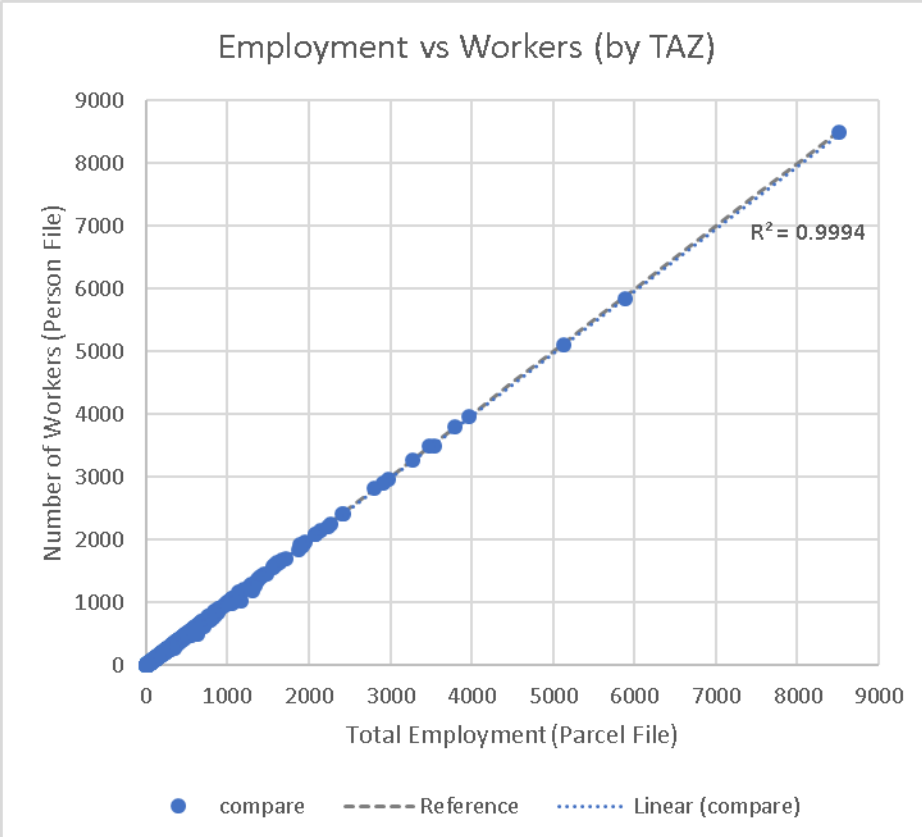


FIGURE 7: EMPLOYMENT VS WORKERS BY TAZ

As shown in Table 9, estimated workers working at a location and the corresponding employment are also compared at an aggregate geography (10 district). A map of the districts is shows in Figure 8. The comparison once again confirmed DaySim's ability to balance workers with the available employment. The comparison is extended to compare estimated workers with observed workers from multiple data sources: LEHD, CTPP, CHTS, and NHTS. Note that the CHTS and the NHTS data are scaled to match estimated workers by home district. Generally, the observed data source among themselves show some significant differences> however, the ABM compares reasonably well with the most datasets.

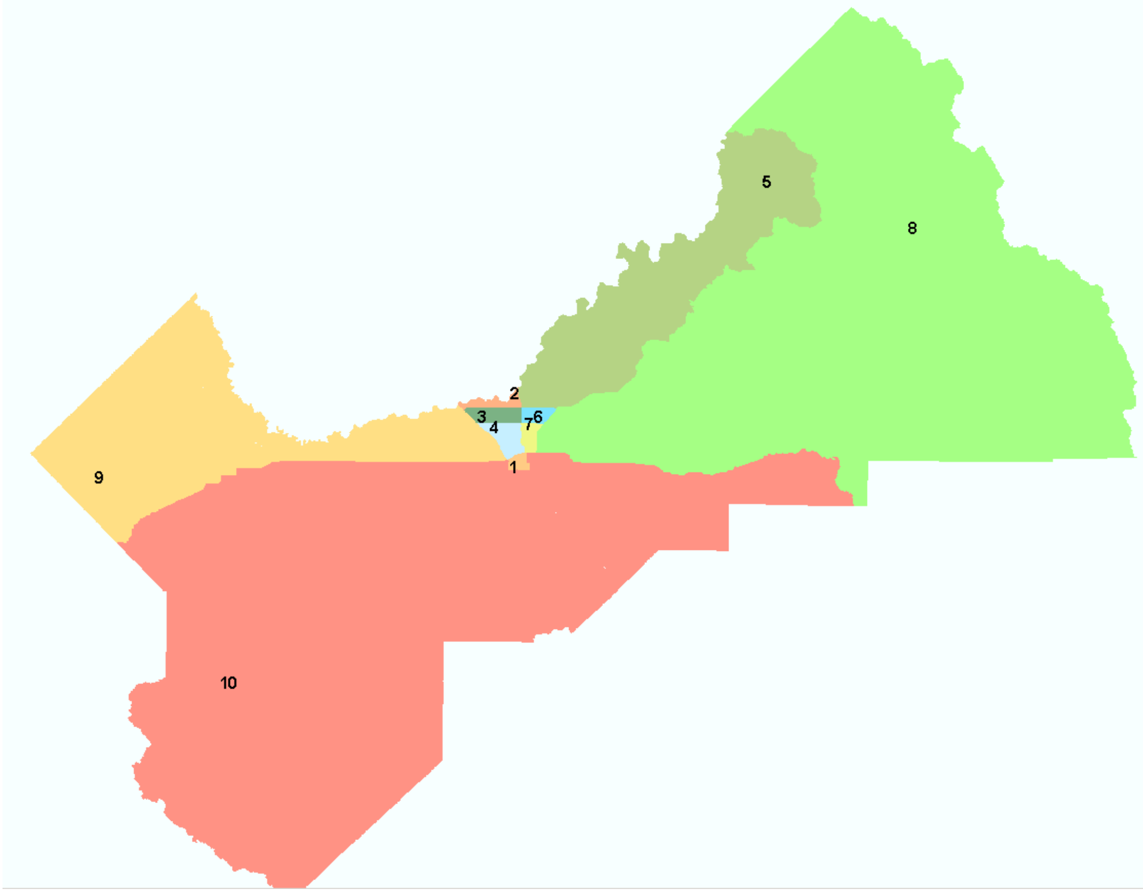


FIGURE 8: MAP OF DISTRICTS

This comparison was proven be very useful during the model validation, where estimated traffic volumes on some freeway corridors showed a systematic pattern compared to observed traffic counts. The observations lead to some district-level demand adjustments in DaySim to generate demand that showed improved comparisons with the observed datasets. The district-level adjustments also improved model validation results to some extent.

TABLE 9: WORKERS BY DISTRICT

WORK DISTRICT	EMPLOYMENT	LEHD	CTPP	CHTS	NHTS	ABM
1	37,710	41,623	30,117	33,572	29,054	37,693
2	17,025	16,553	8,590	9,512	19,926	16,938
3	22,467	23,789	25,873	20,950	24,320	22,421
4	31,619	29,191	42,542	45,088	42,832	31,382
5	26,006	25,982	24,016	31,172	28,409	25,896

6	20,525	19,070	12,946	15,448	11,565	20,510
7	15,179	14,560	10,834	17,736	16,476	15,191
8	43,654	49,212	41,828	52,347	40,645	43,773
9	18,401	14,938	26,478	18,774	19,647	17,551
10	116,515	110,821	105,763	97,836	109,560	110,710
TOTAL	**349,102**	**345,739**	**328,987**	**342,434**	**342,434**	**342,431**

NOTES:
LEHD and CTPP data include workers travelling from out of Fresno
CHTS and NHTS data is only for Fresno workers.
ABM is estimated workers working within the Fresno region only

Home to School Distance

As per CHTS, Table 10, the average distance travelled by students to go from home to school is 3.49 miles. For the same, the NHTS suggests a lower average distance of 2.89 miles. The ABM is adjusted to a value (3.88 mile) close to the CHTS, though slightly higher.

TABLE 10: AVERAGE HOME TO SCHOOL DISTANCE

STUDENT TYPE	CHTS	NHTS	ABM
Kids 5 to 15	2.58	2.42	2.72
Student 16+	3.55	2.05	4.04
University Student	7.98	8.64	8.05
Total	**3.49**	**2.89**	**3.88**

Figure 9 presents a comparison of observed and estimated frequency distribution of trip lengths between home and school. As seen in Table 10, compared to the CHTS, the NHTS data indicates a relatively smaller average distance between home and usual school location. This is also reflected in TLFDs for the NHTS data, more trips are shorter than 3 miles. As the ABM is calibrated to follow the CHTS profile, the ABM TLFD follows the CHTS TLFD closely.

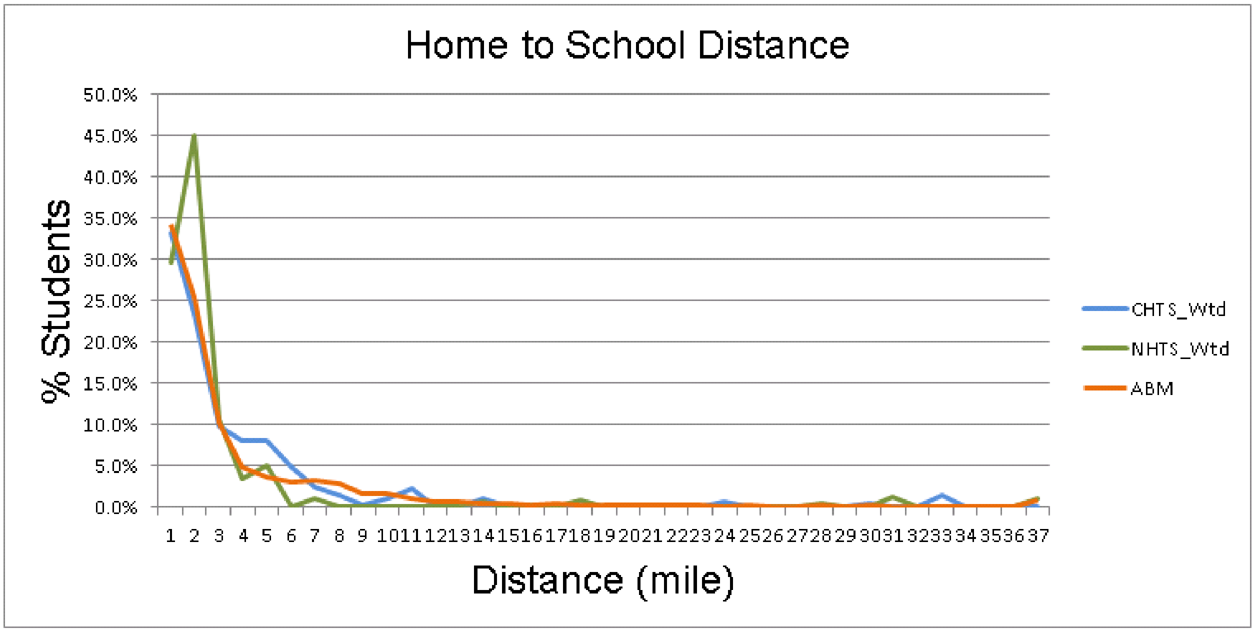


FIGURE 9: DISTRIBUTION OF HOME TO SCHOOL DISTANCE

A comparison of estimated school locations with the corresponding enrollment by TAZ exhibits a good match with a R-squared value of 0.96 and a regressed fitted line following the 45-degree line closely.

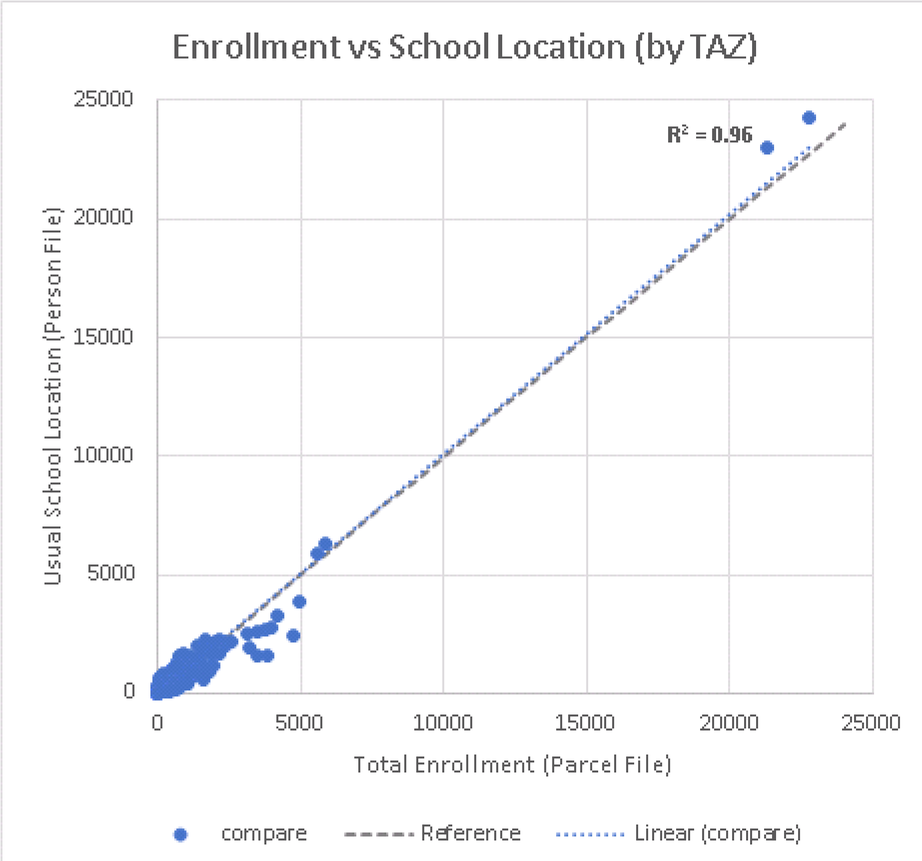


FIGURE 10: ENROLLMENT VS STUDENTS BY TAZ

Auto Ownership

The auto ownership model predicts number of vehicles owned by a household. The auto ownership model is structured as a multinomial logit (MNL) with five available alternatives: 0, 1, 2, 3, and 4+. Key variables are the numbers of working adults, non-working adults, students of driving age, children below driving age and income. Table 11 and Table 12 present share of households by number of vehicles and drivers in the household from the CHTS and the ABM respectively. Difference of household shares between the two datasets are presented in Table 13. The ABM is calibrated to reasonably match the CHTS shares. Note that to calibrate the auto ownership model, the CHTS data for the entire SJV is utilized.

TABLE 11: SHARE OF HOUSEHOLDS BY VEHICLES AND DRIVERS (CHTS-SJV)

	NUMBER OF VEHICLES					
No. of Drivers	0	1	2	3	4+	Total
1	5.4%	16.9%	3.4%	0.5%	0.2%	26.3%

2	2.7%	12.1%	24.2%	5.1%	1.2%	45.2%
3	0.7%	3.1%	5.7%	5.4%	1.4%	16.3%
4+	0.4%	1.4%	3.5%	3.4%	3.4%	12.2%
Total	9.2%	33.5%	36.8%	14.4%	6.2%	100.0%

TABLE 12: SHARE OF HOUSEHOLDS BY VEHICLES AND DRIVERS (ABM)

	NUMBER OF VEHICLES					
No. of Drivers	0	1	2	3	4+	Total
1	5.4%	16.9%	2.9%	0.4%	0.2%	25.7%
2	2.5%	12.7%	22.9%	5.2%	1.2%	44.5%
3	0.6%	3.2%	5.9%	5.1%	1.4%	16.2%
4+	0.4%	1.5%	3.6%	3.5%	4.6%	13.5%
Total	9.0%	34.2%	35.1%	14.3%	7.4%	100.0%

TABLE 13: DIFF IN SHARE OF HOUSEHOLDS BY VEHICLES AND DRIVERS (ABM-CHTS)

	NUMBER OF VEHICLES					
No. of Drivers	0	1	2	3	4+	Total
1	0.1%	0.0%	-0.5%	-0.1%	0.0%	-0.6%
2	-0.2%	0.6%	-1.3%	0.2%	0.0%	-0.7%
3	-0.1%	0.1%	0.1%	-0.3%	0.1%	-0.1%
4+	0.0%	0.0%	0.1%	0.1%	1.2%	1.4%
Total	-0.2%	0.7%	-1.6%	-0.1%	1.2%	0.0%

Day Pattern

Day pattern summaries compare observed and estimated resident travel (tours and trips) by purpose and person type. As the observed data (CHTS) is for the entire SJV region but the ABM predicts travel for the Fresno County only, the comparisons use share of travel.

Table 14 compares tours by tour purpose. As per the CHTS data, of the total tours in in the SJV region, majority are mandatory purpose (work-26% and school-17%) related. About 10% are shopping and 12% are social and recreational. The share of tours in the ABM are calibrated to match closely with the CHTS data.

TABLE 14: TOURS BY PURPOSE

TOUR PURPOSE	CHTS (SJV)	ABM	DIFF (ABM-CHTS)
Work	26%	25%	-0.9%
School	17%	17%	0.4%
Escort	18%	18%	0.2%
Personal Business	11%	11%	0.1%
Shop	10%	10%	0.0%
Meal	3%	3%	0.0%
Social/Recreation	12%	12%	0.1%
Work-based	2%	2%	0.1%
Total	**100%**	**100%**	**0.0%**

A tour rate is calculated as number of tours divided by number of persons. Table 15 compares tour rates by tour purpose. The CHTS indicates on average 1.22 tours per person in the SJV region. The calibrated ABM produces a similar but slightly higher tour rate (1.33) for the Fresno region. The ABM tour rates by purpose match closely with tour rates in the CHTS data.

TABLE 15: TOUR RATE BY PURPOSE

TOUR PURPOSE	CHTS (SJV)	ABM	DIFF (ABM-CHTS)
Work	0.32	0.34	0.02
School	0.20	0.23	0.02
Escort	0.21	0.24	0.02
Personal Business	0.14	0.15	0.01
Shop	0.12	0.13	0.01
Meal	0.04	0.05	0.00
Social/Recreation	0.15	0.17	0.01

Work-based	0.03	0.03	0.00
Total	**1.22**	**1.33**	**0.11**

Table 16 compares observed and estimated tours by person type. Generally, the tours in the ABM match with the CHTS distribution by person type. The differences in tours for some person type are due to differences in population in the two datasets. As seen in Table 7, compared to the CHTS distribution for the SJV region, the ABM is low on number of workers and high on retired person, university students and kids under 5 years of age. This difference in population distribution is reflected in total tours for these person types.

TABLE 16: TOURS BY PERSON TYPE

PERSON TYPE	CHTS (SJV)	ABM	DIFF
Full-Time Worker	41%	34%	-7.9%
Part-Time Worker	6%	6%	0.0%
Retired	5%	7%	1.9%
Non-Worker	17%	17%	0.5%
University Student	3%	6%	2.5%
Student 16+	4%	5%	0.8%
Student 5-15	17%	18%	0.7%
Kid Under 5	6%	7%	1.4%
Total	100%	100%	0%

As presented in Table 17, comparison of tour rate by person type categories also exhibit some differences due to difference in population.

TABLE 17: TOUR RATE BY PERSON TYPE

PERSON TYPE	CHTS (SJV)	ABM	DIFF (ABM-CHTS)
Full-Time Worker	1.31	1.40	0.10
Part-Time Worker	1.45	1.45	0.00
Retired	0.80	1.15	0.35
Non-Worker	1.12	1.36	0.24

University Student	1.18	1.46	0.28
Student 16+	1.20	1.44	0.24
Student 5-15	1.29	1.27	-0.03
Kid Under 5	1.16	1.12	-0.04
Total	**1.22**	**1.33**	**0.11**

The distribution of ABM trips by destination purpose matches well with the CHTS data of the SJV region. The trip shares in the ABM are generally within 1-2% of the CHTS shares.

TABLE 18: TRIPS BY PURPOSE

DESTINATION PURPOSE	CHTS (SJV)	ABM	DIFF (ABM-CHTS)
Work	13%	12%	-1.1%
School	6%	7%	0.2%
Escort	12%	11%	-1.7%
Personal Business	11%	12%	0.5%
Shop	9%	10%	1.1%
Meal	4%	4%	-0.1%
Social/Recreation	8%	9%	0.7%
Home	36%	37%	0.5%
Total	**100%**	**100%**	**0.0%**

As shown in Table 19, according to the CHTS data, a resident of the SJV region makes 3.19 trips in a day on average. The ABM produces slightly higher estimate of the trip rate with 3.55 trips per person in the Fresno region. The estimated trip rates by destination purpose match well with the CHTS trip rates, with slightly higher trip rate for return home purpose.

TABLE 19: TRIP RATE BY PURPOSE

DESTINATION PURPOSE	CHTS (SJV)	ABM	DIFF (ABM-CHTS)
Work	0.41	0.42	0.01

School	0.20	0.23	0.03
Escort	0.39	0.38	-0.02
Personal Business	0.35	0.41	0.06
Shop	0.29	0.36	0.07
Meal	0.14	0.15	0.01
Social/Recreation	0.25	0.30	0.05
Home	1.15	1.30	0.15
Total	**3.19**	**3.55**	**0.36**

As indicated in Table 20, The CHTS data suggests, on average, residents of the SJV region make 2.62 trips on a tour. The ABM produces a similar estimate of 2.67 trips per tour for the residents of the Fresno region. The estimated trips per tour by destination purpose show slight differences but are generally similar to the CHTS data.

TABLE 20: TRIPS PER TOUR BY PURPOSE

DESTINATION PURPOSE	CHTS (SJV)	ABM	DIFF (ABM-CHTS)
Work	1.28	1.23	-0.05
School	1.00	1.03	0.03
Escort	1.84	1.59	-0.25
Personal Business	2.57	2.72	0.14
Shop	2.37	2.70	0.33
Meal	3.37	3.30	-0.07
Social/Recreation	1.66	1.83	0.16
Total	**2.62**	**2.67**	**0.05**

The distribution of model trips by person type categories is similar to the CHTS data for the SJV region, Table 21. As described in comparison of tours by person type, the differences for some person categories (full-time workers, retired, university student, and kids under 5 year) are due to differences in population of those persons in the two datasets.

TABLE 21: TRIPS BY PERSON TYPE

PERSON TYPE	CHTS (SJV)	ABM	DIFF
Full-Time Worker	43%	34%	-9.0%
Part-Time Worker	7%	6%	-0.4%
Retired	5%	8%	2.3%
Non-Worker	17%	18%	0.8%
University Student	3%	5%	2.3%
Student 16+	4%	5%	0.7%
Student 5-15	16%	17%	1.1%
Kid Under 5	6%	8%	2.1%
Total	**100%**	**100%**	**0.0%**

As presented in Table 22, similar to the tour rate by person type (see Table 17), the regional trip rate in the ABM is slightly higher (3.55 trips/person) than the CHTS trip rate (3.19 trips/person) for the SJV region. Comparison by person type categories too show higher trip rates in the ABM.

TABLE 22: TRIP RATE BY PERSON TYPE

PERSON TYPE	CHTS (SJV)	ABM	DIFF (ABM-CHTS)
Full-Time Worker	3.56	3.80	0.25
Part-Time Worker	4.05	3.91	-0.15
Retired	2.05	3.19	1.14
Non-Worker	2.97	3.73	0.75
University Student	2.94	3.65	0.71
Student 16+	3.02	3.62	0.60
Student 5-15	3.08	3.16	0.08
Kid Under 5	2.87	3.13	0.26
Total	**3.19**	**3.55**	**0.36**

Other Tour Destination

A comparison of average tour lengths by purpose between the observed (CHTS and NHTS) and the model data is presented in Table 23. A tour length is calculated as distance between tour origin and primary destination. The comparison includes only non-mandatory tour purposes - mandatory tour purposes (work and school) have already been discussed before (see Table 8 and Table 10). Due to insufficient sample size for each purpose category, shopping and personal business purposes are aggregated into the Maintenance category, and meal and social/recreational purposes are aggregated into the Discretionary category.

For each purpose, the average model tour length is calibrated within a range of the CHTS and the NHTS values. However, the work-based tour purpose exhibits longer tour lengths in the ABM than the observed datasets. Due to smaller sample size in the two observed datasets, the tour length frequency distributions of the work-based tours are very lumpy, Figure 14. This makes it difficult to know the real travel behavior for the work-based tours. As shown in Figure 11, Figure 12, and Figure 13, the observed distributions for other tour purposes are relatively less lumpy due to more samples for these purposes. The ABM distributions are generally smooth and follow distributions from the two observed datasets.

TABLE 23: AVERAGE TOUR LENGTHS FOR OTHER TOUR PURPOSE

TOUR PURPOSE	CHTS (FRESNO)	NHTS (FRESNO)	ABM
Maintenance	2.88	3.38	3.99
Discretionary	4.91	6.40	5.95
Escort	2.88	3.85	3.99
Work-based	3.26	2.88	5.14

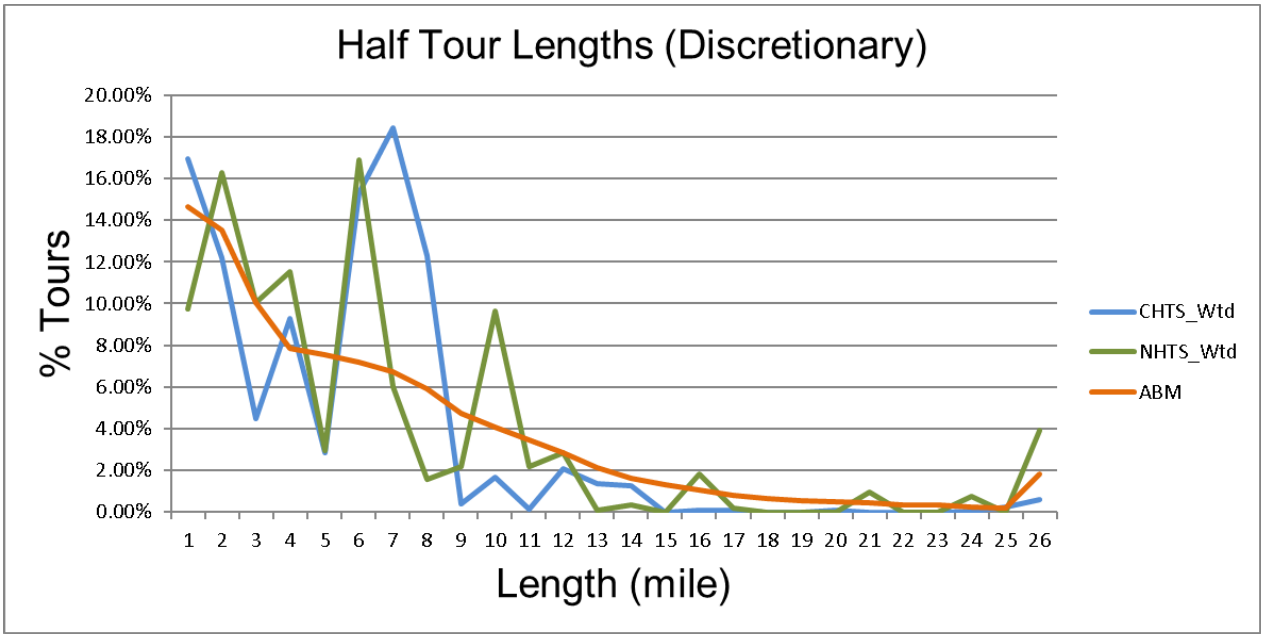


FIGURE 11: TOUR LENGTH DISTRIBUTION FOR DISCRETIONARY TRAVEL

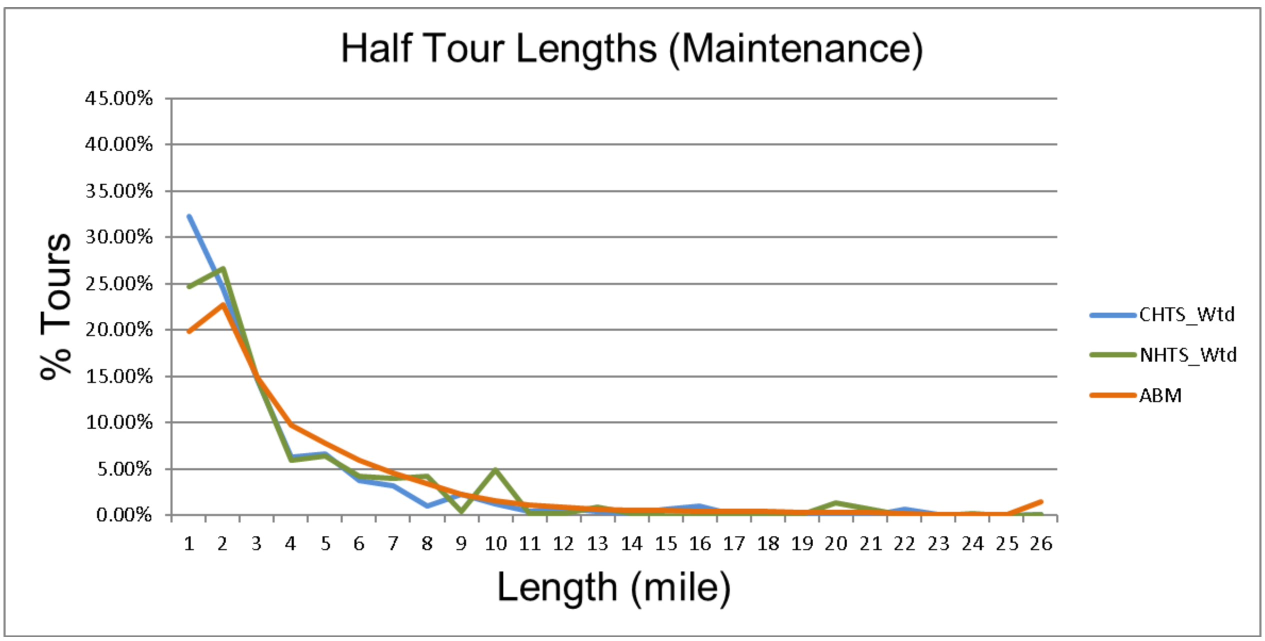


FIGURE 12: TOUR LENGTH DISTRIBUTION FOR MAINTENANCE TRAVEL

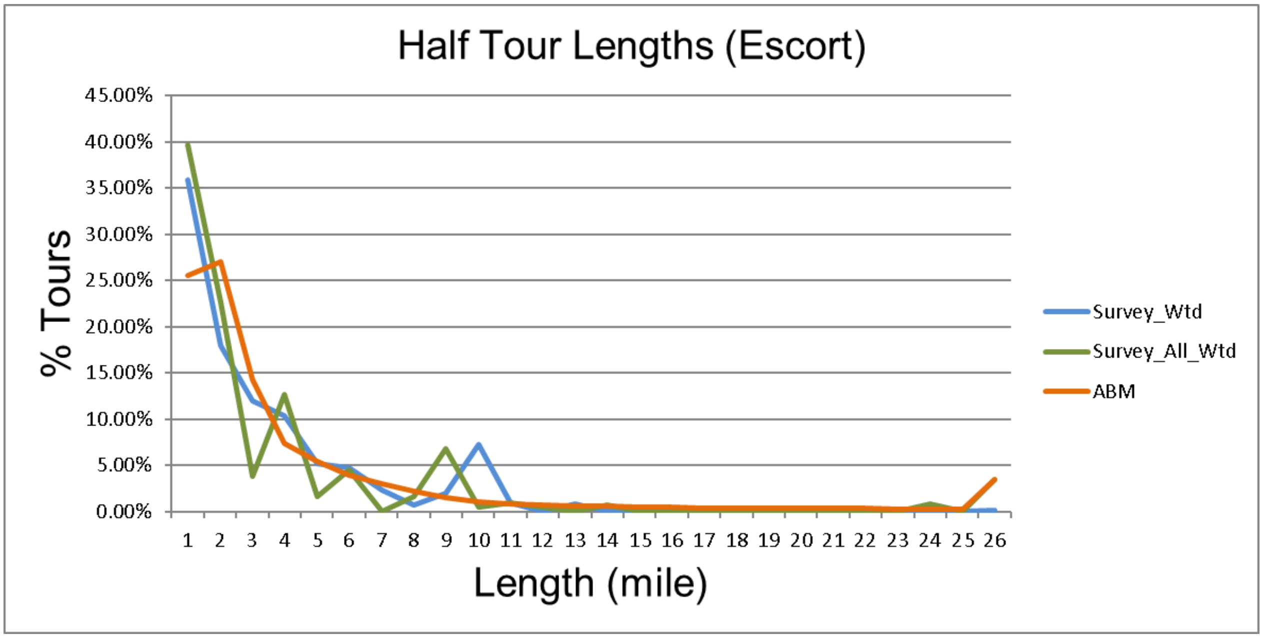


FIGURE 13: TOUR LENGTH FREQUENCY DISTRIBUTION FOR ESCORT TRAVEL

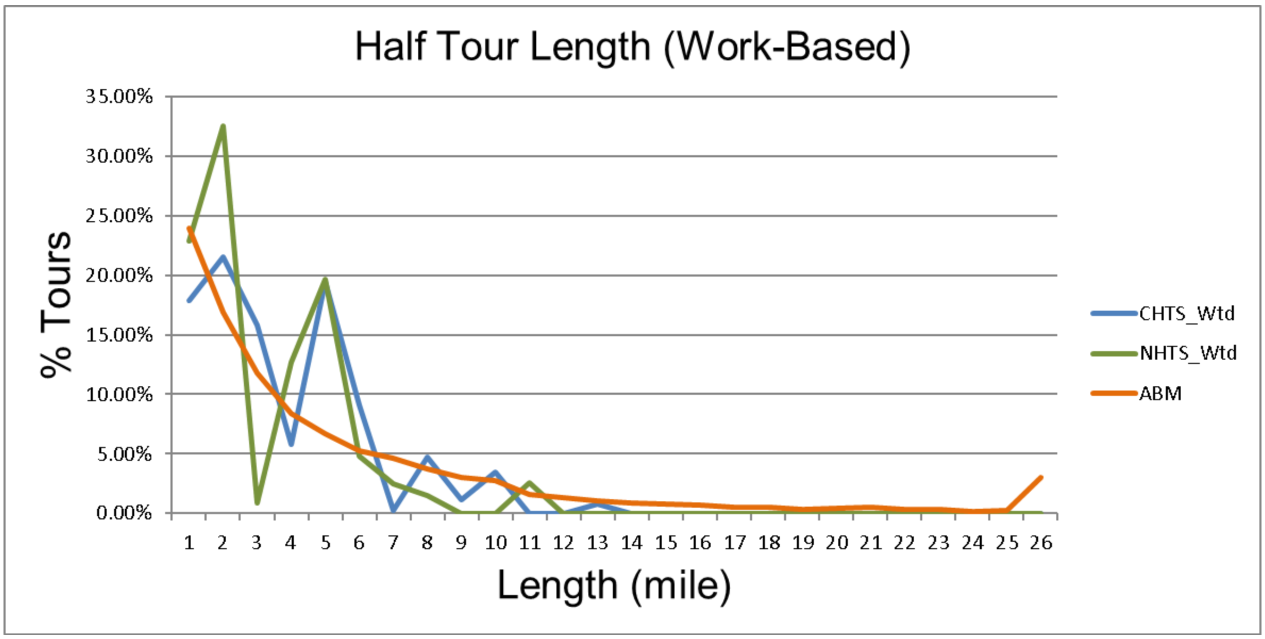


FIGURE 14: TOUR LENGTH FREQUENCY DISTRIBUTION FOR WORK-BASED TRAVEL

Tour Mode Choice

Tour mode is an abstract concept, defined as the main mode of travel used to get from the origin to the primary destination and back. The following 8 tour modes are available in the ABM: drive alone, shared-ride 2, shared-ride 3+, bike, walk, drive-transit, walk-transit, and school bus. The tour mode is coded in the survey based on a set of rules that are dependent on the combination of trip modes used on the tour. The rules can be summarized as follows:

- Any tour with a transit trip is defined as a transit tour
  - Any transit tour with a PNR-transit trip is defined as a PNR-transit tour
  - Any transit tour with neither a PNR-transit trip or a KNR-transit trip is defined as a walk-transit tour
- Any tour with a bicycle trip is defined as a bicycle tour
- Any tour with an auto trip is defined as an auto tour
  - The highest occupancy mode of all auto trips on the tour is used to set the occupancy of the tour
- Remaining tours are walk tours

A similar set of rules is used in tour mode choice to constrain the availability of trip modes based on tour mode. These rules also influence the accessibilities used to choose the locations of intermediate stops on tours; for example, transit and walk accessibilities are used to choose stop locations on transit tours, rather than auto accessibilities.

After scaling the original CHTS targets to accommodate transit targets from the transit on-board survey, the CHTS targets are scaled one more time for tour mode calibration. Generally, a tour mode choice calibration aims to adjust the mode choice model so that the distribution of tours by mode is similar to observed share. Therefore, tour mode choice adjustments are made to alternative-specific constants to match observed mode shares. As transit tour targets are calculated directly from a transit on-board survey, the model needs to be calibrated to the same numbers. However, when calibrated using mode shares, the number of transit tours based on the share of transit mode in the CHTS would result in a different number due to a different value of total tours in the ABM. For example, if a survey says that there are 100 transit tours among 10,000 total tours, then the transit share would be 1%. However, if the model is generating 12,000 total tours then calibrating the model to the survey transit share of 1% will result in 120 transit tours. Since we want to calibrate the model to match the absolute number of transit tours inferred from the on-board survey, we adjust observed tours by mode, keeping the transit tours constant but scaling other modes to match total tours in the model by purpose and auto sufficiency.

Overall, the tour mode shares in the ABM match the CHTS shares reasonably well (Figure 15). The comparison within the tour purpose categories is also similar, Table 24, Table 25, and Table 26. The CHTS observe an overall tour mode share of 32.4% by drive-alone (SOV) and 23.2% and 29.3% by shared-ride 2 and shared-ride 3 respectively. 1.3% of the tours use some form of transit mode with the most (1.1%) use walk to transit. The non-motorized tour modes (walk and bike) make up for 10.7% of the tours in the region.

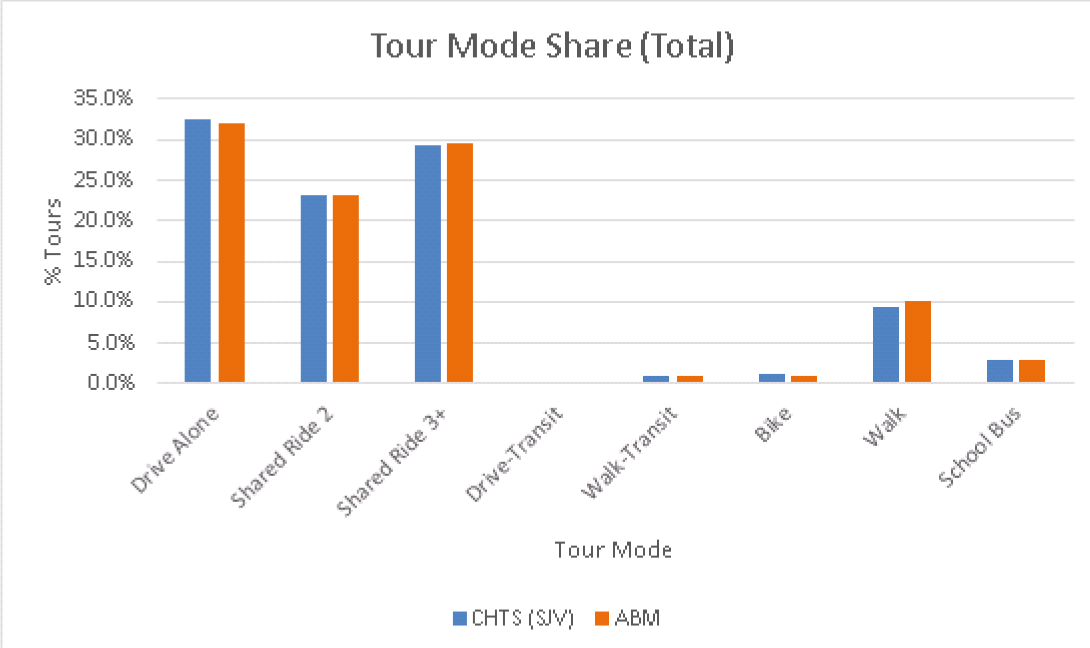


FIGURE 15: TOUR MODE SHARES (TOTAL)

TABLE 24: TOUR MODE SHARES (CHTS-SJV)

MODE	WORK	SCHOOL	ESCORT	OTHER	WORK-BASED	TOTAL
Drive Alone	66%	8%	2%	32%	77%	**32.5%**
SR2	19%	16%	33%	26%	13%	**23.2%**
SR3+	11%	41%	56%	26%	3%	**29.4%**
Drive Transit	0%	0%	0%	0%	0%	**0.2%**
Walk Transit	1%	1%	0%	1%	0%	**0.9%**
Bike	1%	3%	0%	2%	0%	**1.3%**
Walk	2%	14%	9%	13%	7%	**9.5%**
School Bus	0%	17%	0%	0%	0%	**3.0%**
Total	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

TABLE 25: TOUR MODE SHARES (ABM)

MODE	WORK	SCHOOL	ESCORT	OTHER	WORK-BASED	TOTAL
Drive Alone	67%	9%	2%	32%	73%	**32.3%**
SR2	19%	16%	32%	26%	12%	**23.3%**
SR3+	11%	42%	56%	26%	3%	**29.8%**
Drive Transit	0%	0%	0%	0%	0%	**0.1%**
Walk Transit	1%	1%	0%	1%	0%	**1.0%**
Bike	0%	1%	1%	2%	0%	**1.2%**
Walk	2%	13%	9%	13%	12%	**9.3%**
School Bus	0%	17%	0%	0%	0%	**3.0%**
Total	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

TABLE 26: TOUR MODE SHARES (ABM-CHTS)

MODE	WORK	SCHOOL	ESCORT	OTHER	WORK-BASED	TOTAL
Drive Alone	0.4%	0.2%	0.5%	0.2%	-4.2%	**-0.2%**
SR2	0.1%	0.3%	-0.9%	0.3%	-0.7%	**0.1%**
SR3+	0.0%	1.0%	-0.4%	0.4%	0.0%	**0.5%**
Drive Transit	0.1%	-0.3%	0.0%	-0.2%	0.0%	**-0.1%**
Walk Transit	0.1%	0.0%	0.0%	0.1%	0.0%	**0.1%**
Bike	-0.6%	-1.2%	1.4%	0.0%	0.2%	**-0.1%**

Walk	-0.1%	-0.5%	-0.4%	-0.4%	4.8%	**-0.2%**
School Bus	0.0%	0.4%	-0.3%	0.3%	0.0%	**0.0%**
Total	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**

Trip Destination

As presented in Table 27, the CHTS data suggest an average trip length of 4.21 miles regionwide (Fresno). For the same, the calibrated ABM produces only a slightly higher trip length value (4.75 miles). The trip lengths by trip destination purpose match reasonably well too.

TABLE 27: TRIP LENGTHS (MILES) BY DESTINATION PURPOSE

TRIP DESTINATION PURPOSE	CHTS (FRESNO)	ABM
Home	3.97	4.96
Work	7.70	7.61
School	3.24	2.99
Escort	3.47	3.72
Personal Business	3.98	4.25
Shop	2.53	3.46
Meal	4.23	4.61
Social/Recreational	4.18	4.85
Total	**4.21**	**4.75**

Tour Time of Day

Plots of tour arrival and departure times at primary destination are presented in Figure 16 through Figure 23. The ABM distribution generally match well with the CHTS distribution by purpose. The work-based arrival times are more peaked in the survey than in the model, which is common since the model tends to predict smooth distributions.

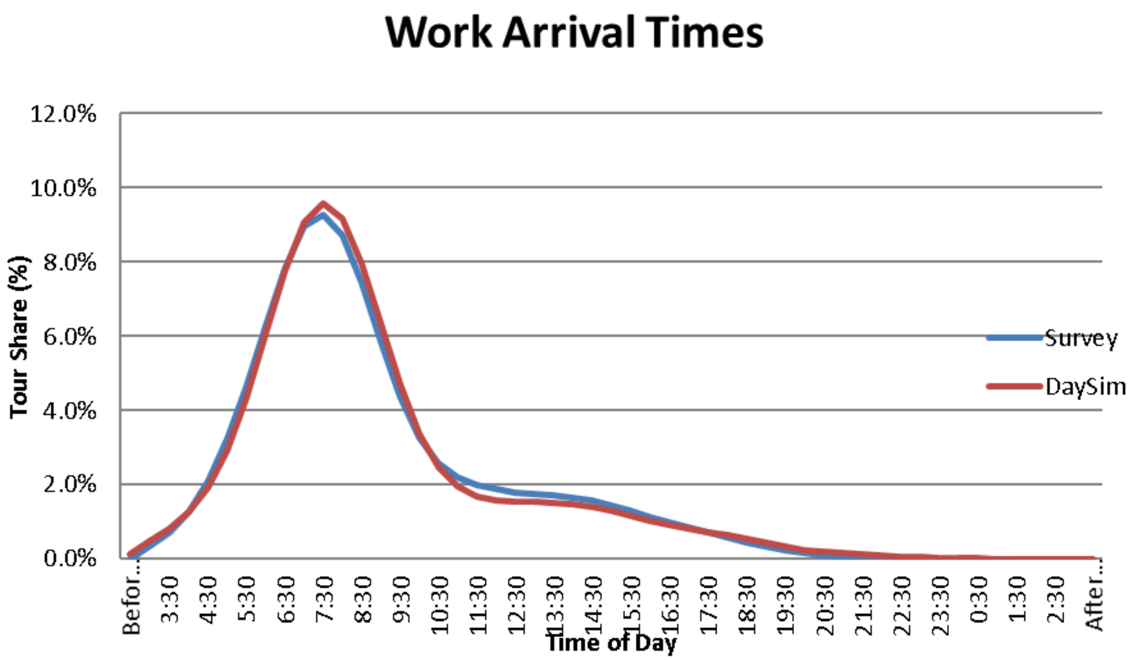


FIGURE 16: TIME OF DAY DISTRIBUTION OF WORK ARRIVAL TIMES

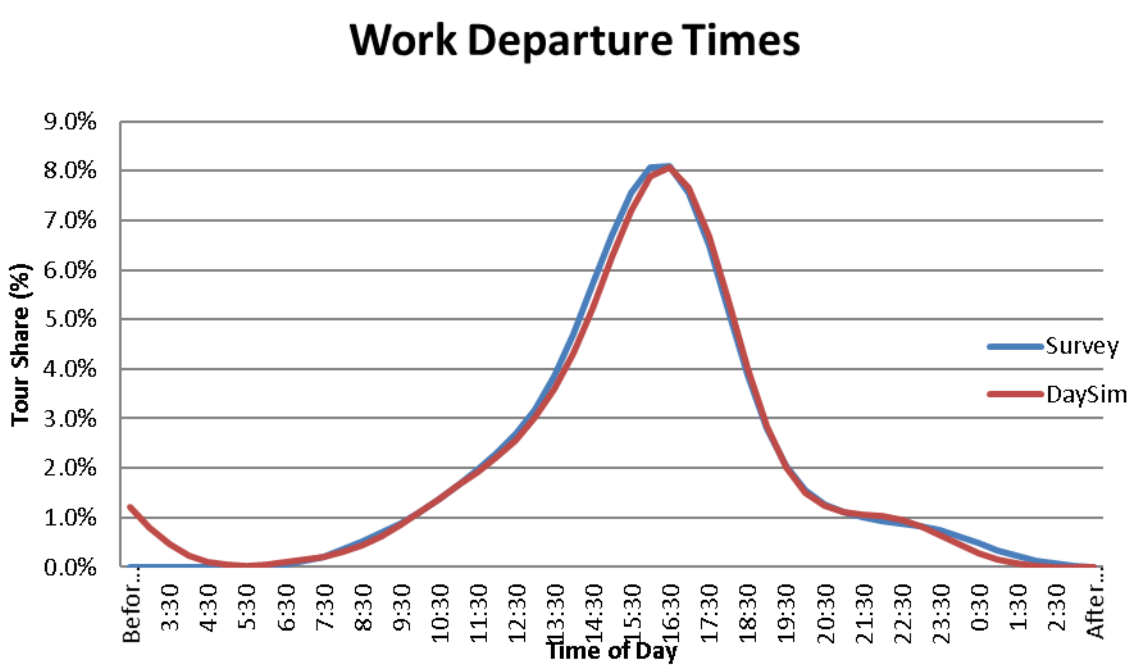


FIGURE 17: TIME OF DAY DISTRIBUTION OF WORK DEPARTURE TIMES

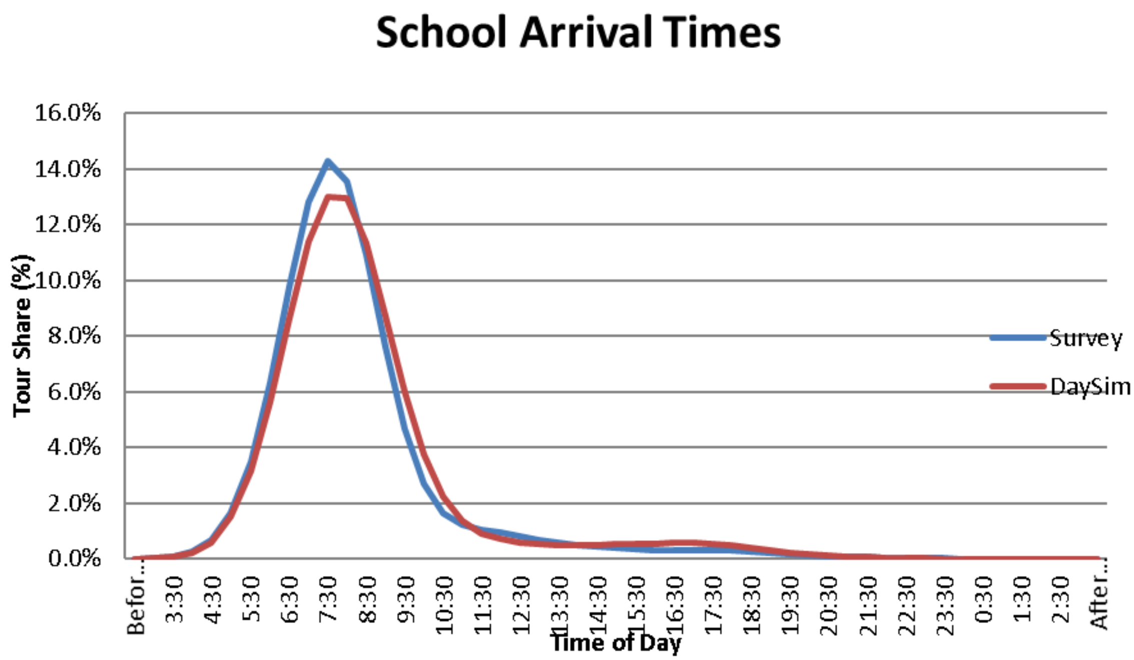


FIGURE 18: TIME OF DAY DISTRIBUTION OF SCHOOL ARRIVAL TIMES

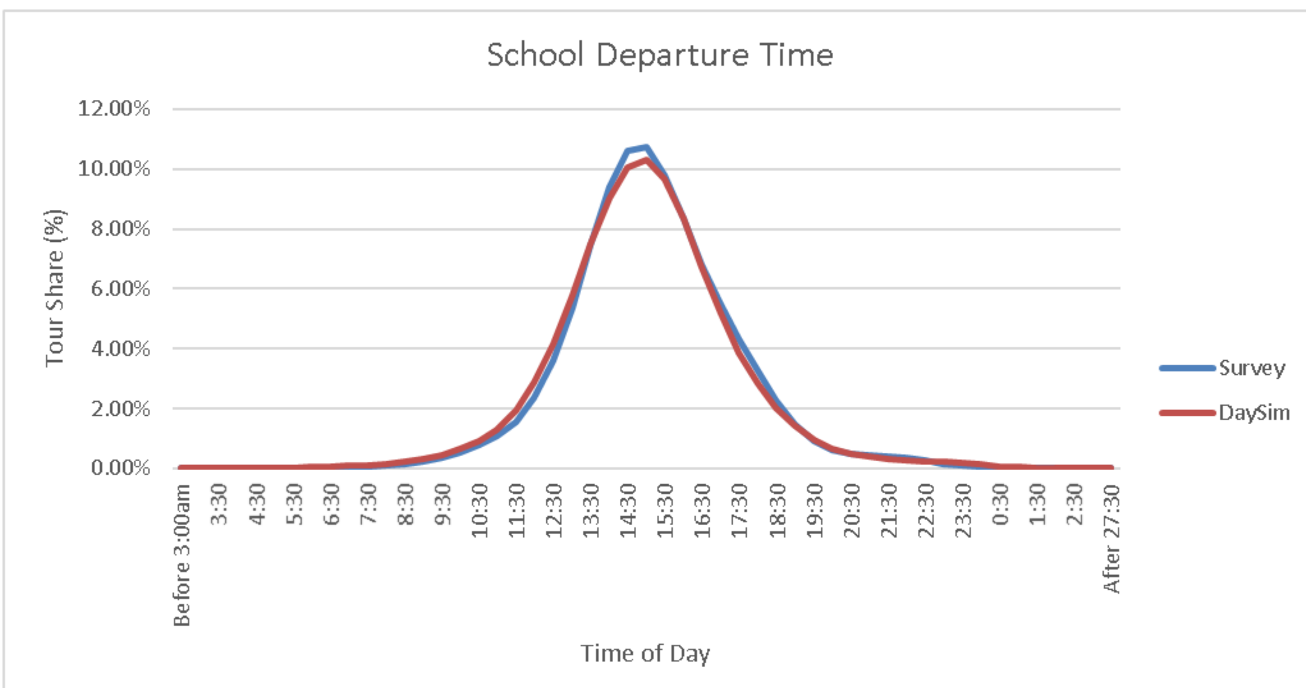


FIGURE 19: TIME OF DAY DISTRIBUTION OF SCHOOL DEPARTURE TIMES

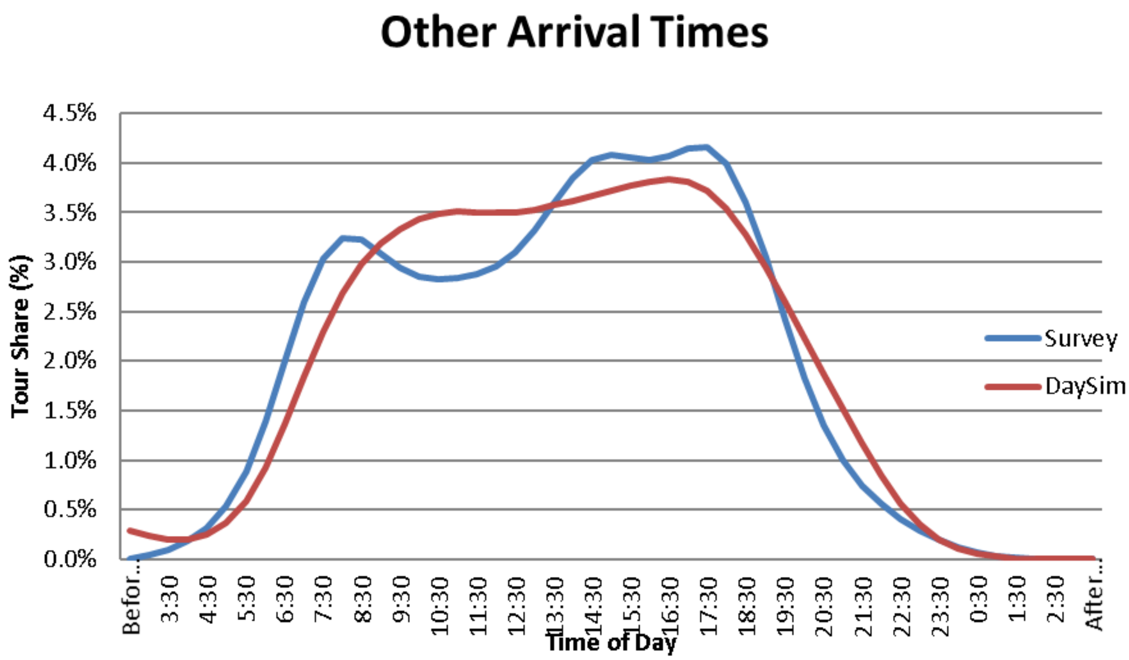


FIGURE 20: TIME OF DAY DISTRIBUTION OF OTHER PURPOSE ARRIVAL TIMES

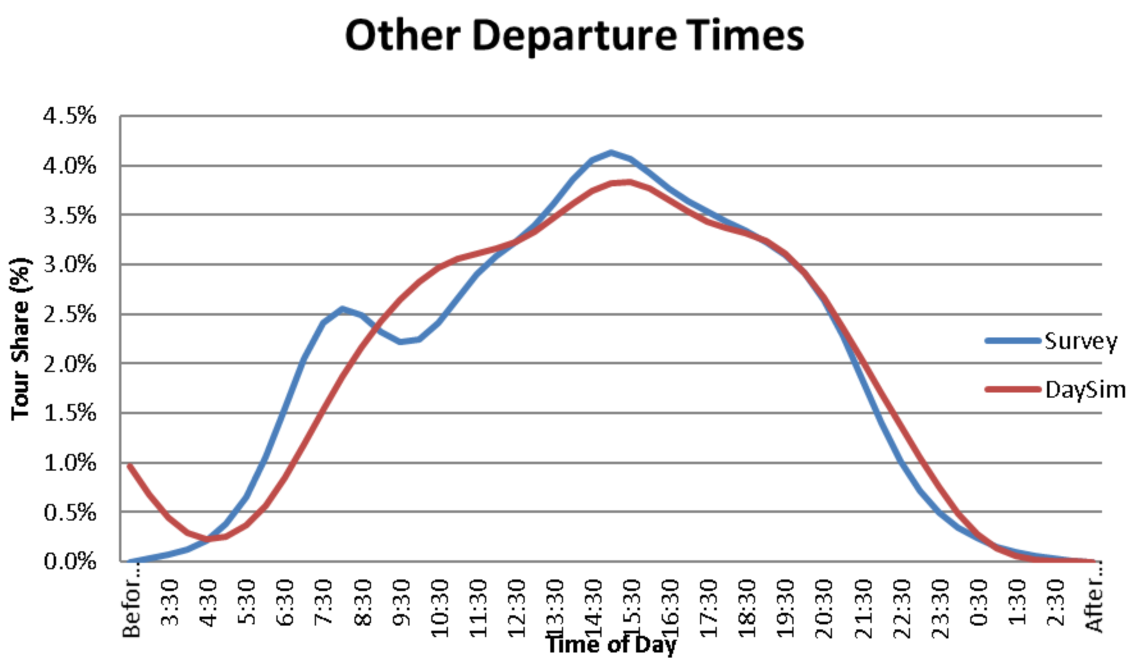


FIGURE 21: TIME OF DAY DISTRIBUTION OF OTHER PURPOSE DEPARTURE TIMES

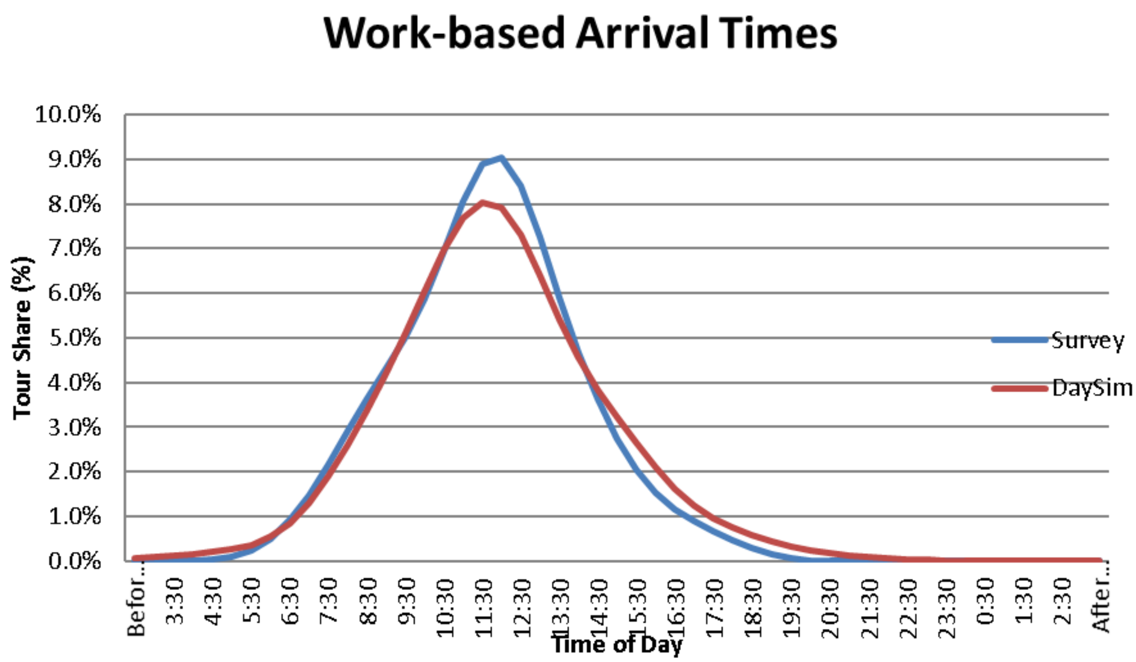


FIGURE 22: TIME OF DAY DISTRIBUTION OF WORK-BASED ARRIVAL TIMES

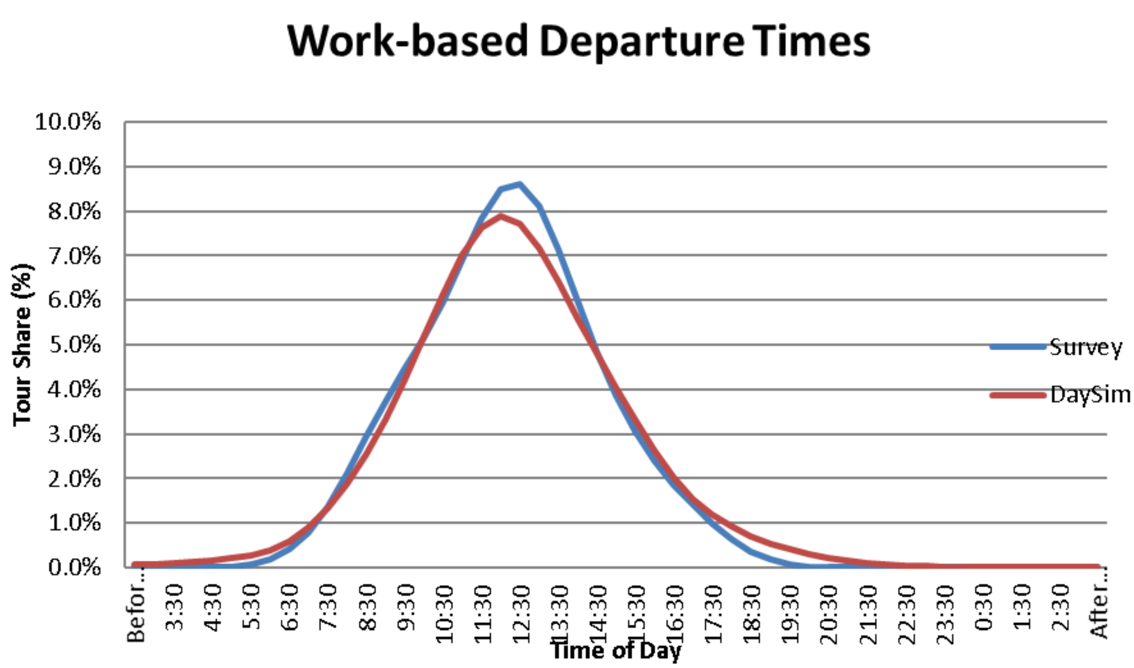


FIGURE 23: TIME OF DAY DISTRIBUTION OF WORK-BASED DEPARTURE TIMES

Trip Mode Choice

Trip mode targets are prepared from the CHTS data for the SJV region and updated with transit trip targets from the transit on-board survey. Other mode targets are appropriately scaled to keep the total trips by purpose the same, similar to the process described above for creation of tour mode choice targets. This ensures that the absolute number of expanded transit trips from the transit onboard survey is matched in calibration.

The calibration process involves adjustment of alternative-specific constants to match observed trips by trip mode and tour mode within each tour purpose. The trip mode choice model can be thought of as a 'mode switching' model, in which the tour mode constrains which modes are available for trips on tours.

Overall, the ABM generates a trip mode distribution which is very similar to observed (Figure 24, Table 28, Table 29, and Table 30). The CHTS data indicate that on an average weekday, 41.5% trips in the SJV region are drive alone and 46.2% are shared-ride (SR2 and SR3), approximately 0.7% of Fresno County resident trips are made by transit, and 10% are made by a non-motorized mode (walk or bike).

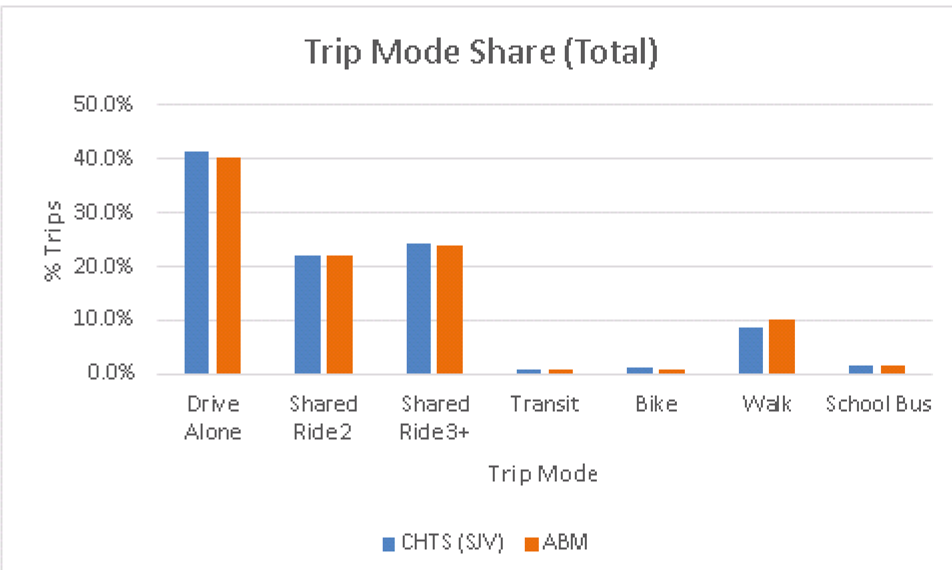


FIGURE 24: TRIP MODE SHARES (TOTAL)

TABLE 28: TRIP MODE SHARES (CHTS-SJV)

MODE	WORK	SCHOOL	ESCORT	OTHER	WORK-BASED	TOTAL
Drive Alone	75%	10%	25%	37%	35%	**41.5%**

SR2	14%	22%	30%	26%	20%	**22.0%**
SR3+	7%	39%	38%	25%	23%	**24.2%**
Transit	1%	1%	0%	1%	0%	**0.7%**
Bike	1%	3%	0%	1%	3%	**1.2%**
Walk	2%	16%	7%	10%	19%	**8.8%**
School Bus	0%	9%	0%	0%	0%	**1.5%**
Total	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

TABLE 29: TRIP MODE SHARES (ABM)

MODE	WORK	SCHOOL	ESCORT	OTHER	WORK-BASED	TOTAL
Drive Alone	76%	11%	24%	34%	75%	**40.7%**
SR2	12%	21%	28%	28%	11%	**22.1%**
SR3+	8%	40%	38%	23%	2%	**24.1%**
Transit	1%	1%	0%	1%	0%	**0.8%**
Bike	0%	1%	1%	1%	0%	**1.1%**
Walk	2%	16%	9%	12%	12%	**9.5%**
School Bus	0%	10%	0%	0%	0%	**1.7%**
Total	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

TABLE 30: TRIP MODE SHARES (ABM-CHTS)

MODE	WORK	SCHOOL	ESCORT	OTHER	WORK-BASED	TOTAL
Drive Alone	1.5%	0.7%	-1.0%	-3.5%	39.7%	**-0.8%**

SR2	-2.6%	-1.0%	-2.5%	2.7%	-8.6%	**0.1%**
SR3+	0.9%	0.9%	-0.1%	-1.8%	-21.0%	**-0.1%**
Transit	0.1%	-0.1%	0.0%	-0.1%	0.0%	**0.1%**
Bike	-0.4%	-1.4%	1.3%	0.3%	-2.4%	**-0.1%**
Walk	0.4%	0.0%	2.3%	2.4%	-7.4%	**0.8%**
School Bus	0.0%	1.0%	0.0%	-0.1%	-0.3%	**0.2%**
Total	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**

3.2 MODEL VALIDATION

A model validation tests the model's predictive capabilities before it is used to produce forecasts. There are two types of model validation; static validation, which compares model outputs against independent data that was not used to build the travel model, and dynamic validation, in which model inputs are systematically varied to assess the reasonableness of model responses. The static validation process compares outputs from model assignment with observed data. Model parameters are adjusted until the model outputs fall within an acceptable range of error.

In the assignment step, model demand (e.g. trips by time period, mode, and vehicle class\valueof-time) are loaded on to network. In highway assignment, the output includes vehicle flows on every link (road) in the highway network and for transit assignment, the output includes the number of boardings on each route. These are compared to observed traffic counts and observed transit ridership respectively. The two observed datasets (traffic counts and transit boardings) used in the present model validation are described in the next section, followed by highway, transit, and non-motorized validation summaries.

Validation Data

Table 31 presents a list of datasets utilized in the validation of the Fresno ABM.

TABLE 31:MODEL VALIDATION DATASETS

DATASET	YEAR	SOURCE	PURPOSE
Traffic Counts	2014	CalTrans and FresnoCOG	Highway Validation
Vehicle Miles Travelled (VMT)	2014	Highway Performance Management System (HPMS)	Highway Validation
Transit Ridership	2014	FAX, Clovis, and FCRTA	Transit Validation
Transit On-Board Survey	2014	Transit On-Board Survey Program	Transit Validation
Bike and Ped Counts	2014	FresnoCOG	Bike and Ped Validation

Highway

Observed traffic counts are used to validate link-level estimated daily traffic flow generated by a model, whereas the observed vehicle miles travelled (VMT) validates the regionwide network-usage as estimated by the model.

Traffic Counts (CalTrans and local)

The observed traffic counts are assembled from two sources: CalTrans and FresnoCOG. The Caltrans traffic census program[12] provides traffic counts on highways (interstates and state routes) in the State of California. These Caltrans traffic counts in year 2014 are downloaded for the Fresno region and in a shapefile format. The count locations (points) in the shapefile are then joined to the model roadway network using a combination of automated and manual review process. The automated process first matched the points in the shapefile to the network links and the manual process reviewed the match and corrected the joins that appeared incorrect. Also, an appropriate count value (Before AADT or After AADT) is assigned to the joined link in the model network.

FresnoCOG provided traffic counts for the facility types (arterial, collector, local etc.) other than highway. The traffic counts already had the corresponding model network link for a count. However, during the model validation, several issues related to suspect wrong link match were discovered therefore some traffic counts were manually reviewed for their correctness of the network link id.

[12] http://www.dot.ca.gov/trafficops/census/

VMT

The 2014 observed VMT is obtained from Caltrans[13]. Caltrans provide daily vehicle miles of travel for COFCG (Council of Fresno County Governments).

Transit

Transit ridership (boardings) by route compare the estimated boardings in the model by transit line. The transit on-board survey provides total number of transit trips and is used to compare transfer rate in the model.

Transit Boardings

The transit boardings are assembled from three sources: FAX, CLOVIS and FCRTA. The three transit agencies provided daily ridership in year 2014 for their transit routes.

Bike and Ped

The bike and ped validation data come from two sets of surveys. Recently, FresnoCOG collected more bike and ped counts in the City. The model validation utilized the combined count database.

Highway Validation

The estimated traffic flows from the model and the observed traffic counts are compared in various dimensions, including:

- Region
- Facility Type
- Volume Group
- Screenline
- Key Corridors
- Time of day

Region

As described before, the observed traffic count database used in this model validation effort encompass 1,433 links on the highway network. The total traffic across these links sum up to 9.92 million vehicles. On the same links, the ABM produce a comparable estimate of traffic volume (10.00 million vehicles) and is only 0.8% higher than the total observed vehicle count. According to the HPMS[14], on an average weekday in year 2014, the roadway travel in the Fresno region resulted in 22.57 million vehicle miles of travel

[13] http://www.dot.ca.gov/hq/tsip/hpms/datalibrary.php
[14] Highway Performance Monitoring System

(VMT). The estimated traffic flows from the ABM produce a daily regionwide VMT value (21.75 million) of within 3.7% of the observed estimate from the HPMS.

Table 32: Highway Validation – Region

MEASURE	OBSERVED	ABM	DIFF	% DIFF
Traffic Volume	9,925,114	10,007,103	81,990	0.8%
VMT*	22,574,620	21,745,004	(829,616)	-3.7%

Note: Observed VMT is from the HPMS estimate of the total VMT in the Fresno region for year 2014

Regionally, the estimated traffic flows are compared with the observed traffic counts by creating a scatter plot, Figure 25. Points in the scatter plot are links where traffic counts are available. A point represents observed traffic count on the X-axis and the corresponding estimated flow on the Y-axis. The scatter plot includes several measures/guidelines assessing accuracy of the model flows with respect to the observed traffic counts.

First, the plot includes a 45-degree line representing a virtual scenario of perfect match between traffic counts and estimated flows. The 45-degree line is useful in quickly identifying overestimation (flow>count) or underestimation (flow<count) of a flow. A highway validation aims to make most points as close to this line as possible. An ideal validation would have all count locations on the 45-degree line. However, perfect match for all count locations is almost impossible to achieve due to various reasons such as error in traffic counts, simulation errors in the model etc. Acknowledging this fact, Caltrans rather provides recommendations on maximum (high and low) deviations of an estimation flow from the corresponding traffic count value. The scatter plot displays these Caltrans high and low deviations as dotted lines above and below the 45-degree line respectively. Lastly, a linear regressed line of all points is also added to the plot. Slope of the regressed line measures regional match between the estimated flows and the traffic counts - a slope of less than 1 means underestimation regionwide and more than 1 indicates overestimation. The plot also displays a R-squared value representing goodness of fit of all data points.

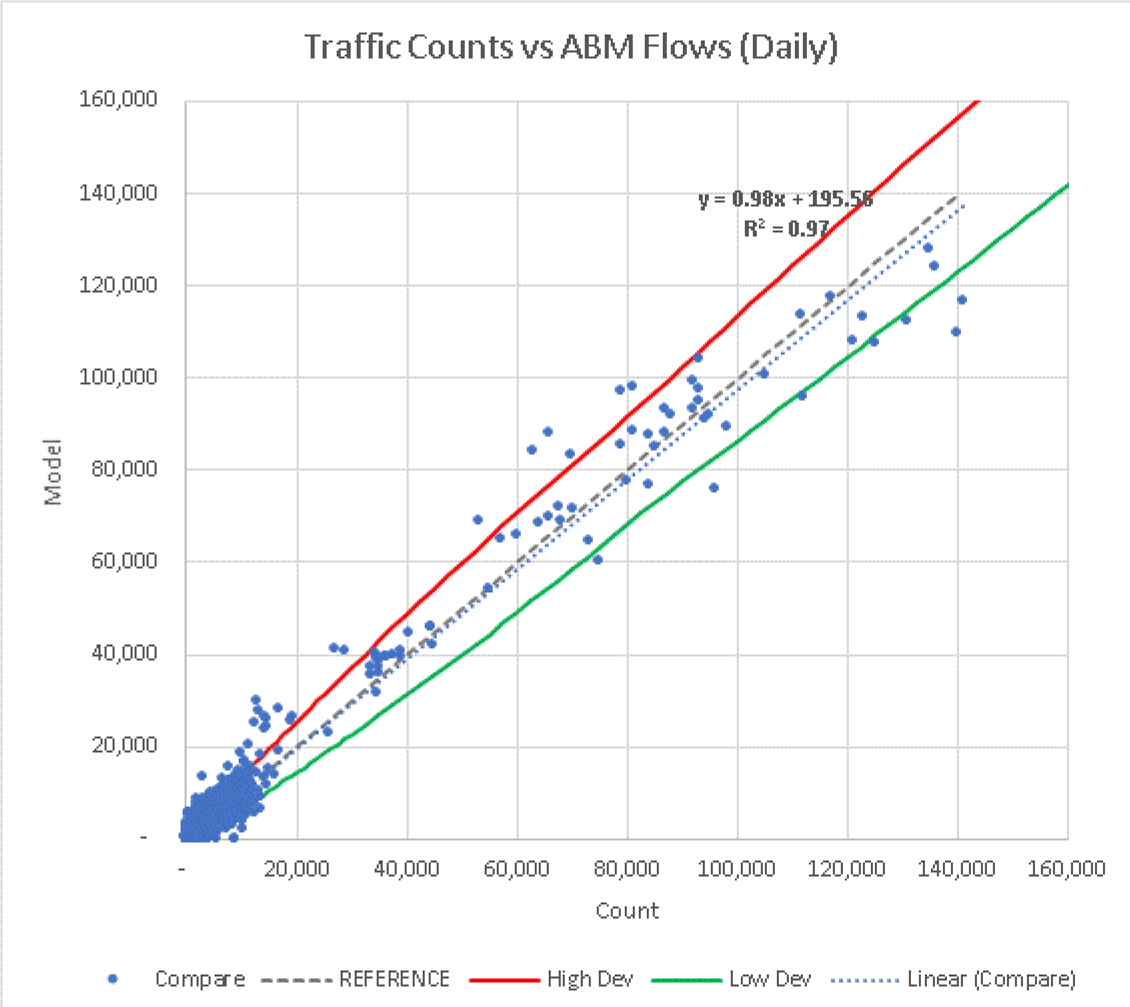


FIGURE 25: DAILY ESTIMATED FLOWS VS OBSERVED TRAFFIC COUNTS

As displayed in the scatter plot, the linear regressed line has a slope of 0.98 and R-squared value of 0.97. The slope indicates a that estimated flows are slightly underestimated compare the traffic counts. The r-squared value close to 1.0 indicates that fitted regression line represents the data well.

Facility Type

Table 33 presents a summary of links by facility type. The facilities in the Fresno region are grouped into three categories: freeway/highway, arterial, and other (collectors, local roads, ramps). The table also contains the FHWA's guidelines of recommended threshold of difference for each facility type. Overall, the estimated traffic volume from the model matches closely (-1%) with the total counts on the compared links. The comparison within the facility type is exhibit good match as well. The other facilities including collector, local roads and ramps, show underestimation. The model's estimates of traffic volume meet the Caltrans' recommended thresholds regionally as well as by facility type.

TABLE 33: HIGHWAY VALIDATION – BY FACILITY TYPE

FACILITY TYPE	COUNT	ABM	DIFF	DIFF (%)	*CALTRANS DIFF
Freeway & Highway	5,854,140	6,106,956	252,816	4%	7%
Arterial	2,242,455	2,264,522	22,067	1%	15%
Other (Collector, local & ramp)	1,828,520	1,635,626	(192,893)	-11%	25%
TOTAL	**9,925,114**	**10,007,103**	**81,990**	**-1%**	**25%**

*Note: the column "Caltrans" is the FHWA's recommended guideline of RMSE for the volume group

Volume Group

The estimated and observed volumes are compared by the level of volume on the links. Table 34, compares the estimated traffic flows and the traffic counts in 7 volume groups that are formed based on the range of the observed traffic counts. Overall, links with lower volumes show bigger differences and RMSE values. This is not surprising given that these links are more likely to be collectors or arterials and as we discussed before, respective traffic counts are less reliable. Links with traffic volume higher than 25,000 outperform the Caltrans guidelines for RMSE. The lower volume (<25,000) links are underperforming with respect to the Caltrans guidelines.

TABLE 34: HIGHWAY VALIDATION – BY VOLUME GROUP

VOLUME GROP		COUNT	ABM	DIFF	DIFF (%)	RMSE	*CALTRANS RMSE
>=0	<1000	131,555	182,306	50,751	39%	150%	60%
>=1000	<2500	668,568	682,886	14,318	2%	75%	47%
>=2500	<5000	1,405,955	1,367,594	(38,361)	-3%	52%	36%
>=5000	<10000	1,938,055	1,889,804	(48,251)	-2%	37%	29%
>=10000	<25000	859,102	931,677	72,576	8%	45%	25%
>=25000	<60000	920,362	1,005,360	84,998	9%	16%	22%
>=60000		4,001,518	3,947,477	(54,041)	-1%	13%	21%
ALL		**9,925,114**	**10,007,103**	**81,990**	**1%**	**43%**	**40%**

*Note: the column "Caltrans" is the FHWA's recommended guideline of RMSE for the volume group

Screenlines

A total of 13 screenlines in the region are compared for validating estimated traffic flows with the observed traffic counts. Figure 26 shows a map of screenline locations in the Fresno region.

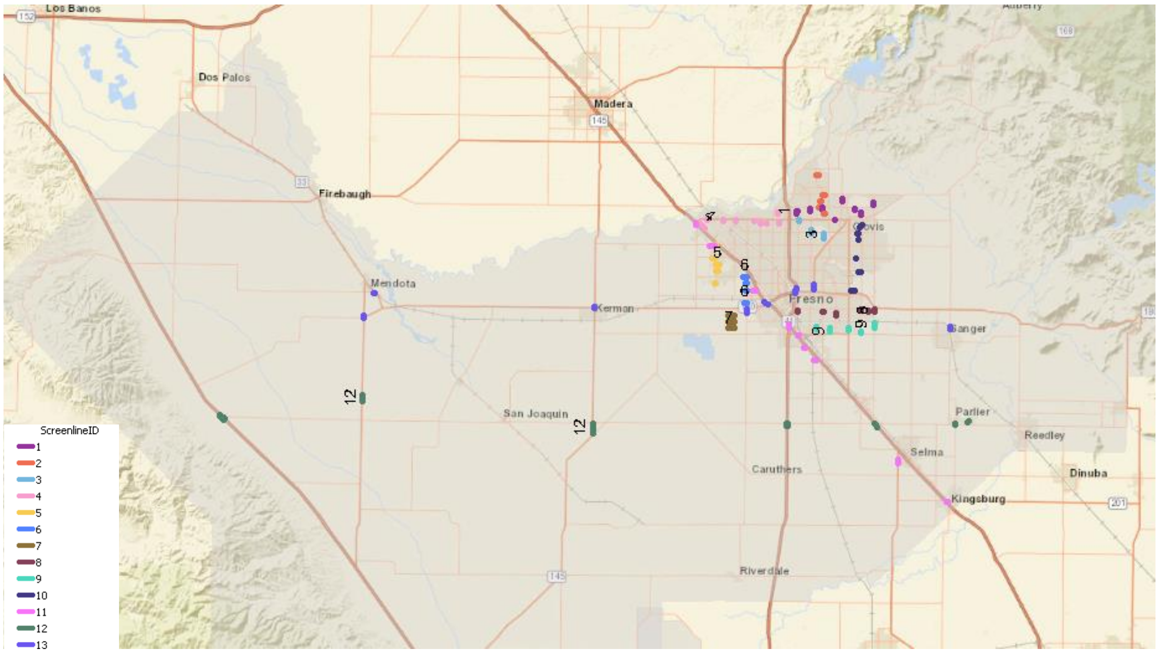


FIGURE 26: SCREENLINE LOCATIONS

Figure 27 presents a bar plot of the comparison by screenline. A tabular summary of the same is provided in Table 35. The figure contains two vertical bars for each screenline representing the observed traffic count and the estimated traffic flow from the ABM. The plot also includes a dotted line showing percentage difference between the two data. The screenlines are placed on the X-axis, traffic volume (in vehicles) is on the primary Y-axis, and percentage difference between the estimated and the observed traffic volume is on the secondary Y-axis.

As shown in Table 35, the 13 screenlines cover a total of 160 links with observed traffic counts. Overall, total traffic volume across all screenlines is slightly underestimated (-4%) in the ABM, which is consistent with the regional highway validation. Generally, the estimated flows on the screenlines compare well with the observed traffic counts. Except three screenlines (2, 5, 7, and 10), estimated flows on all screenlines are within 20% of the observed traffic counts with more than half (6) screenlines being within 10%. Among the three, the screenline 7 includes low volume facilities, therefore a bigger difference is not that alarming. The screenline 2 and 5 are Chestnut and Polk avenue respectively. These screenlines shall be looked at in more detail in future model validation.

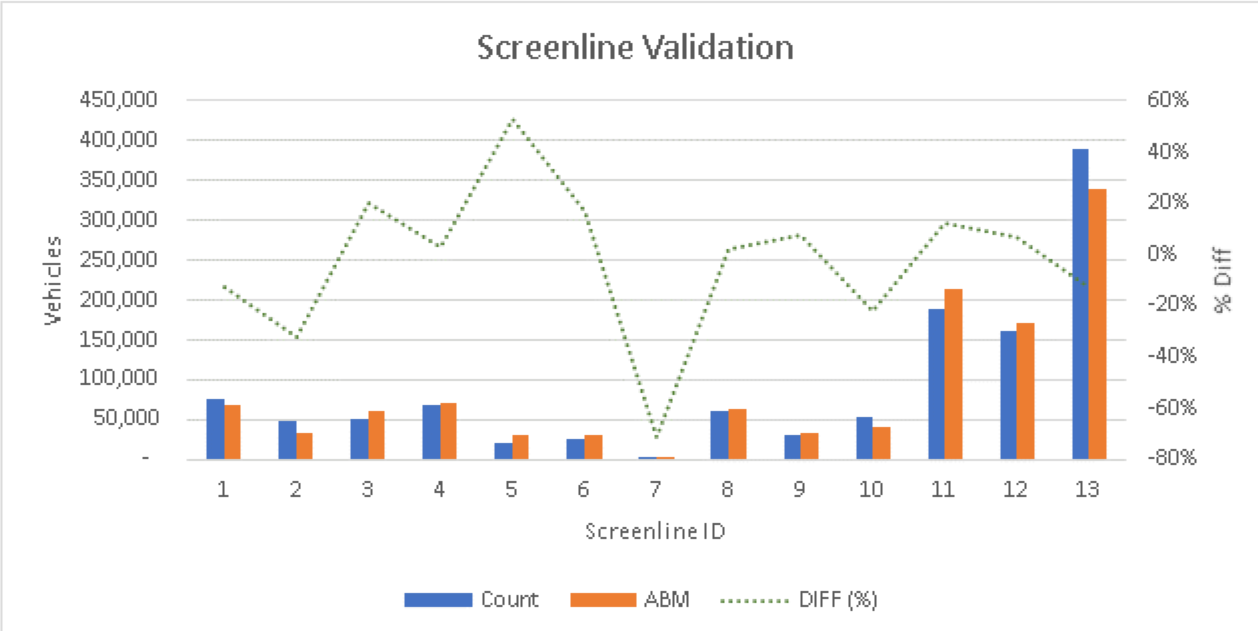


FIGURE 27: SCREENLINE VALIDATION

TABLE 35: SCREENLINE VALIDATION

SCREENLINE ID	LINK COUNT	COUNT	ABM	DIFF	DIFF (%)
1	16	76,870	67,436	(9,434)	-12%
2	10	49,191	33,319	(15,872)	-32%
3	6	50,476	60,752	10,276	20%
4	18	69,623	71,945	2,322	3%
5	10	20,280	30,972	10,692	53%
6	8	26,472	30,891	4,419	17%
7	6	3,182	921	(2,260)	-71%
8	12	61,608	63,036	1,428	2%
9	10	32,081	34,533	2,452	8%
10	12	52,924	41,464	(11,460)	-22%

11	20	190,009	213,099	23,090	12%
12	14	161,039	172,066	11,027	7%
13	18	388,541	338,702	(49,839)	-13%
Total	**160**	**1,182,295**	**1,159,136**	**(23,159)**	**-2%**

Key Highway Corridors

Figure 28 presents a spatial distribution of the count locations on the highways in the Fresno region. The color of a point on the map indicate percent difference between the estimated and observed traffic volume. As shown in the legend, a red means underestimation (<-20%) in the ABM, whereas a green color represents overestimation (>20%). The color becomes darker with increase in overestimation or underestimation. For example, a light green indicates overestimation of 20% to 50% and a dark green indicates overestimation of more than 100%.

Regionally, it is evident that highways outside the urban areas are generally overestimated in the ABM. That is probably due to the coarser representation of the roadway network in those areas and therefore resulting in missing important streets or incorrect location of the centroid connectors for the zones in those areas. Missing important streets in the roadway network would lead the ABM to re-route the traffic to other major streets, therefor assigning them the traffic volume more than actual. Incorrect locations of the centroid connectors would load un-reasonable demand on to the connected facility, thus resulting in higher estimation of traffic volume on the facility. The roadway network shall be reviewed for these issues to improve the ABM's performance in those areas.

The map also reveals two other systematic issues: underestimation on the SR41 north of downtown Fresno and overestimation on the SR145 south of Kerman.

Further investigation into the SR41 corridor revealed that the problematic portion of the SR41 corridor serves several major shopping centers and hospitals in the region. Primarily, a regional shopping center, River Park Shopping Center, is located next to the SR41 near the north end of the county boundary. According to Fresno COG staff, the regional shopping center serves a big portion of the Fresno region and is generally a preferred option for retail by the residents in the county. This suggests that the mall is probably attracting more travel than a regular shopping center and the residents living in far locations too are making significant travel to the mall, thus resulting in longer travel lengths for the location. The model travel for the regional shopping center is calibrated in similar to the regionwide shopping travel behavior. However, as the above discussion established, the travel lengths for the River Parks shopping center are more likely to be longer than a usual shopping travel. This under-estimation of the shopping travel distance could be causing underestimation of estimated traffic volume on the SR 41 corridor north of Fresno downtown. This problem can be attempted in two ways: first, in the ABM, consider the shopping center as a special generator and apply separate parameters for its travel. The parameters can be calibrated using an observed data collected specifically for the travel to/from the shopping center. The San Diego Association of Governments (SANDAG) already use this approach in their regional travel model (SANDAG ABM) to represent travel for several beaches, shopping centers, hospitals, and parks in the San Diego county region. An alternative approach is to apply district level constants to adjust travel between the shopping center and the rest of the region. This approach is already being used in the Fresno ABM, however, in absence of good observed data for travel to shopping centers, it is difficult to correctly represent the travel behavior in the model. Future model improvements should include collecting observed data for this travel and either adjust existing district level constants using the collected observed data or update the ABM to represent the River Parks shopping center as a special generator and calibrate its travel to the observed data.

The overestimation on SR145 south of Kerman is likely due to coarser representation of the roadway network outside urban cities of the Fresno County. As discussed before, missing important streets and/or incorrect centroid connector locations could be causing this overestimation. A closer look into the roadway network in those areas is necessary to improve this part of the SR145.

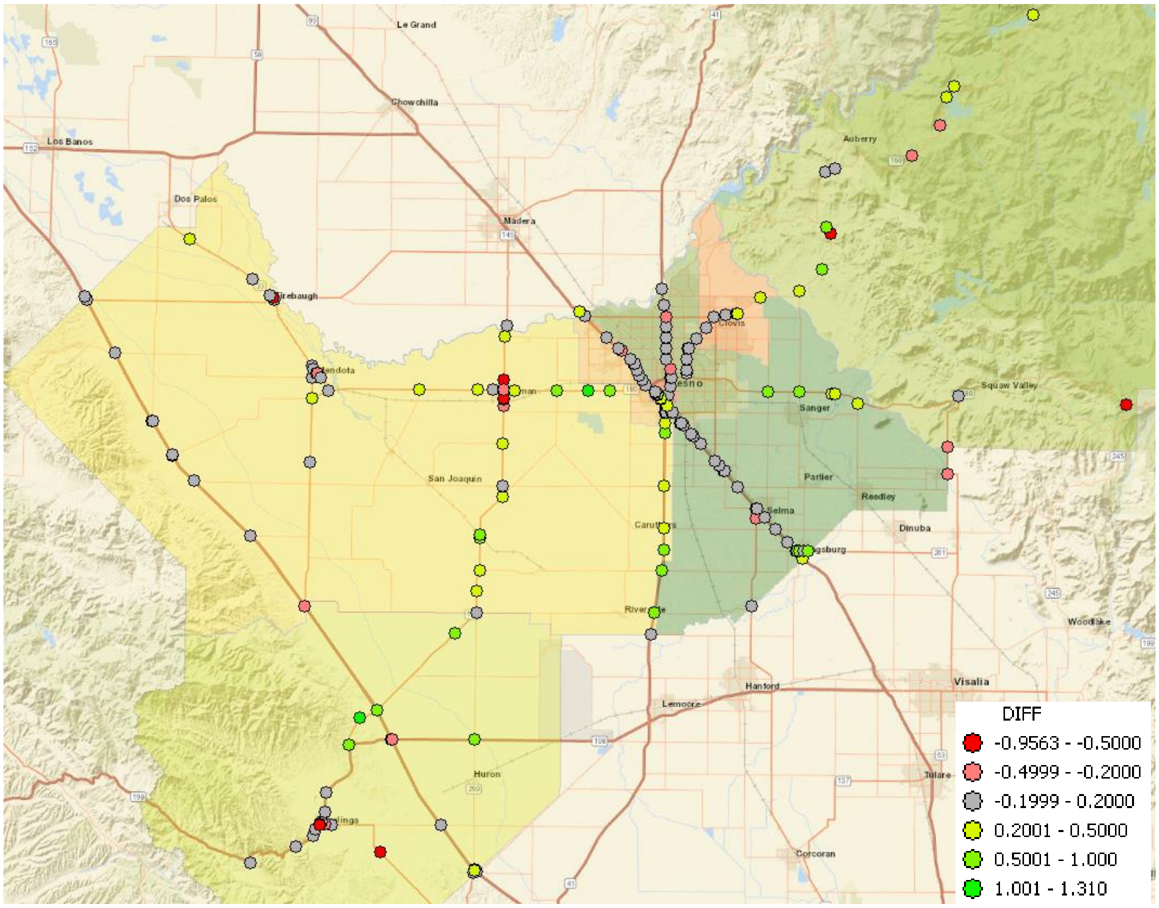


FIGURE 28: HIGHWAY VALIDATION – SPATIAL PERFORMANCE OF THE FRESNO HIGHWAY

Table 36 compares seven key high-volume highway corridors in the Fresno region: I-5, SR33, SR41, SR99, SR145, SR168, and SR-180. As shown, most of the highway corridors (I-5, SR33, SR41, SR99, SR145, and SR168) are doing well with respective percentage difference within 10% of the observed traffic count. Only one highway corridor, SR180, exhibit under performance with overestimation of 29%.

TABLE 36: HIGHWAY VALIDATION – KEY CORRIDORS

ROUTE	CALTRANS_AADT	ABM_VOL	DIFF	DIFF (%)
5	464,362	502,351	37,989	8%
33	108,061	101,587	(6,474)	-6%
41	1,428,200	1,416,350	(14,750)	-1%
99	2,653,518	2,718,865	65,347	2%

145	98,250	99,307	1,057	1%
168	543,390	537,639	(5,751)	-1%
180	193,595	250,018	56,423	29%

To examine the model's performance along a corridor, separate validation plots are prepared for the highway corridors (see Figure 29, Figure 30, Figure 31, Figure 32, Figure 33, Figure 34, and Figure 35). The plots validate estimated model flows at each count location on the corridor. The count locations are arranged sequentially either from south to north or West to East depending on the corridor's travel direction. Note that the direction of the corridor represents only the order of count locations and not the direction of traffic flow. The traffic flows for both estimated and observed data are aggregate of the two flow directions (A to B and B to A). Each figure contains two solid lines representing the estimated traffic flows and the observed counts for count locations on the corridor. It also includes a dotted line showing percentage difference between the two data. The count locations are placed on the X-axis, whereas traffic volume (in vehicles) is on the primary Y-axis and percentage difference between the estimated and the observed traffic volume is on the secondary Y-axis.

The corridor level plots indicate the same two issues as discovered in the spatial map of highway count locations: underestimation on the SR41 north of downtown Fresno and overestimation on the SR145 south of Kerman. In addition, they shed more light into the overestimation of SR180 (can also be seen in the spatial map, Figure 28). When the corridor plot of the SR180 is analyzed in conjunction with the spatial map, it is quite easy to see that the overestimation starts after the city of Mendota (and just before Kerman) and continues on the rest of the corridor towards the east. A possible reason for this overestimation could be seasonal travel in the corridor and thus, raise two questions: if employment in the MAZs served by the state route is represented correctly in the model and also if the observed traffic counts are adjusted for seasonal travel. These two questions need to be looked at in more detail in order to resolve the overestimation on SR180.

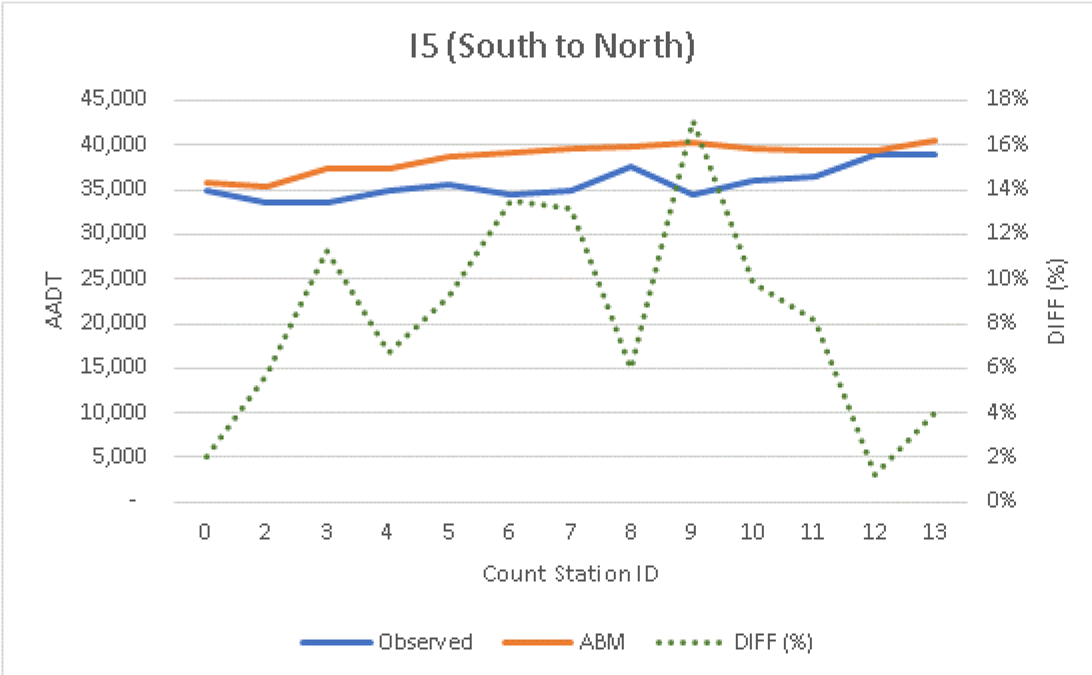


FIGURE 29: HIGHWAY VALIDATION – I5

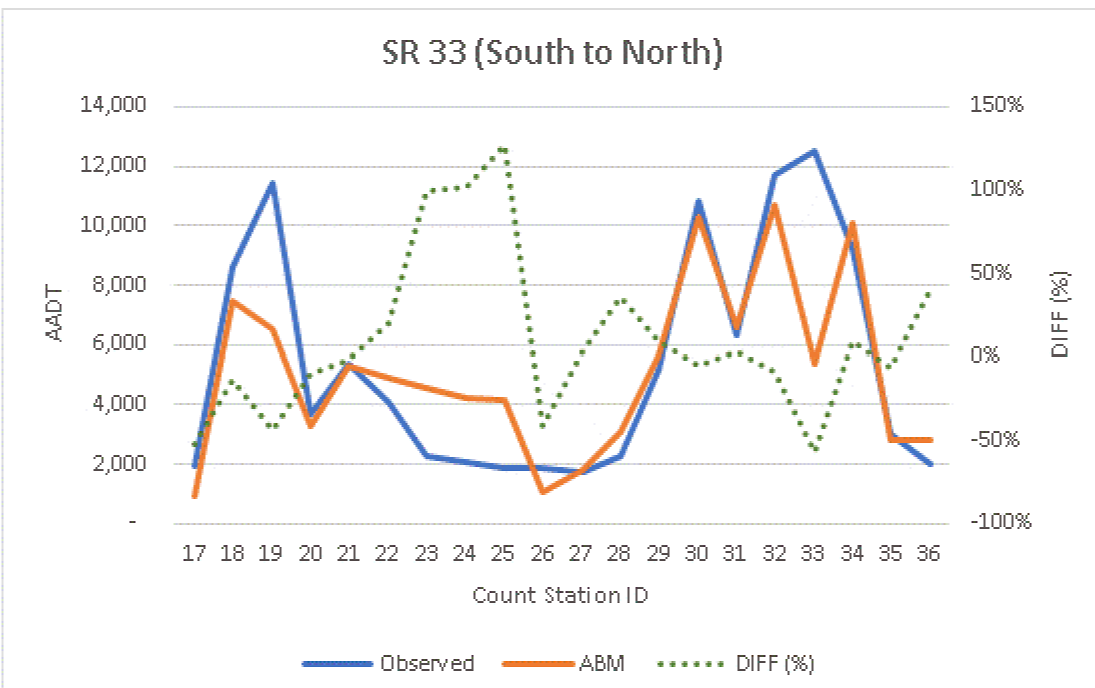


FIGURE 30: HIGHWAY VALIDATION – SR33

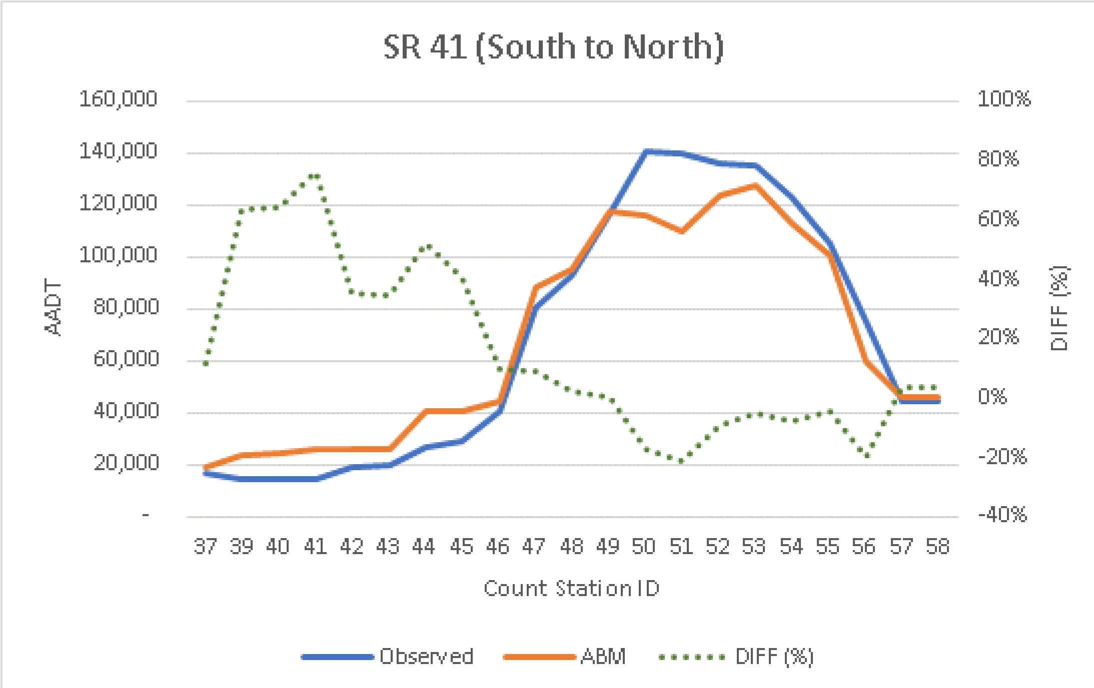


FIGURE 31: HIGHWAY VALIDATION – SR41

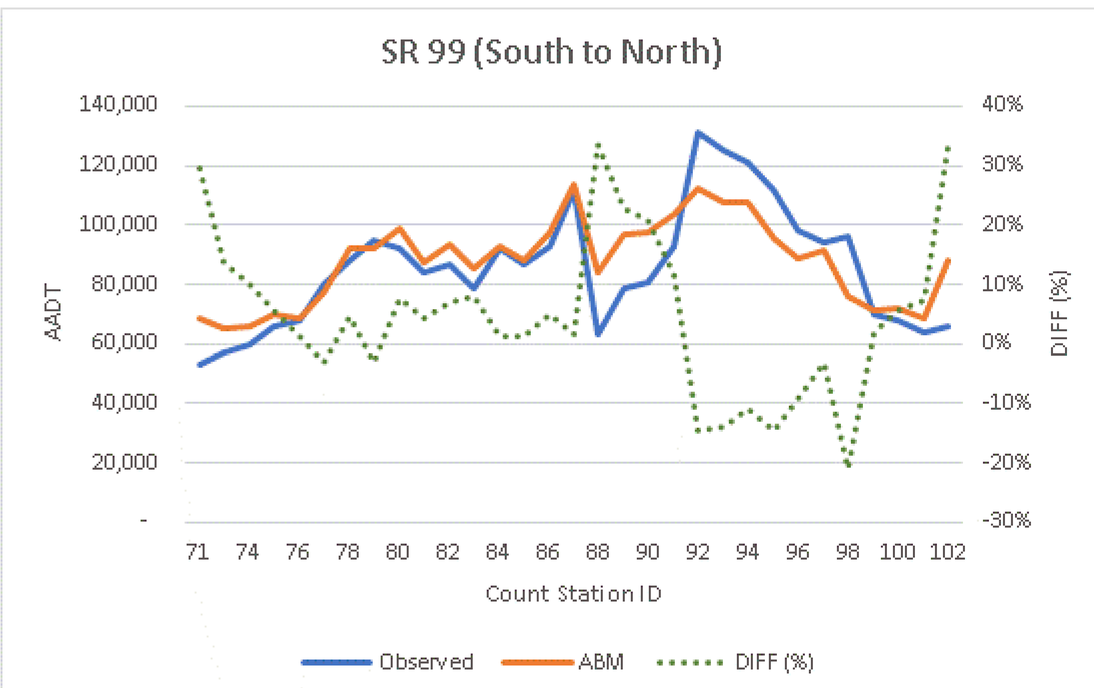


FIGURE 32: HIGHWAY VALIDATION – SR99

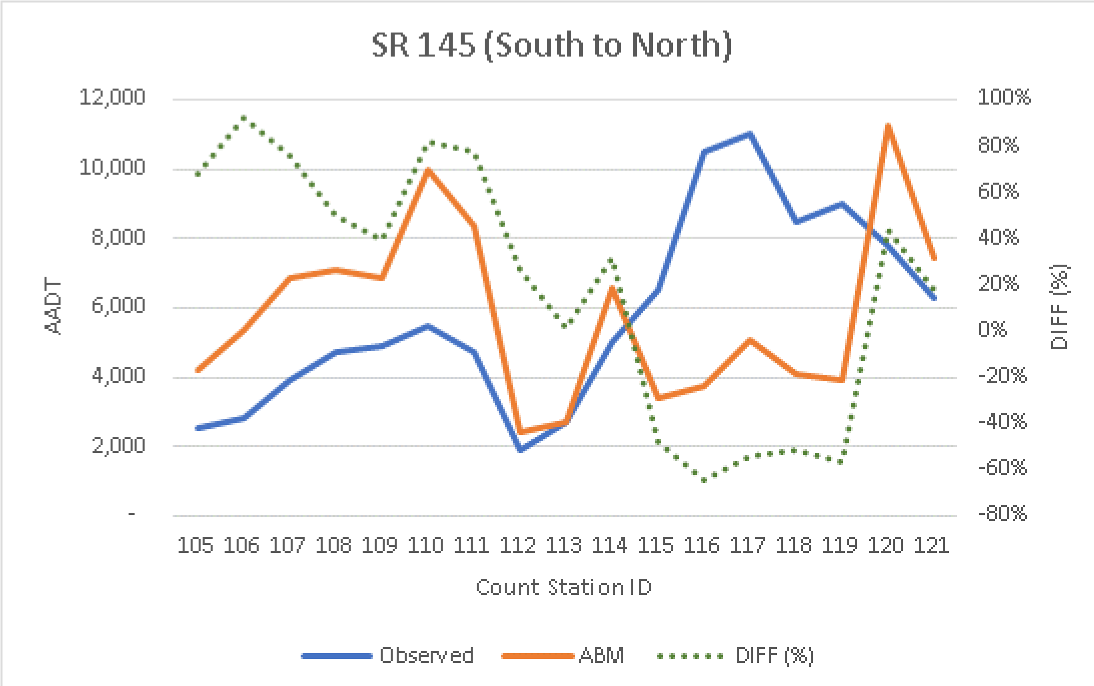


FIGURE 33: HIGHWAY VALIDATION – SR145

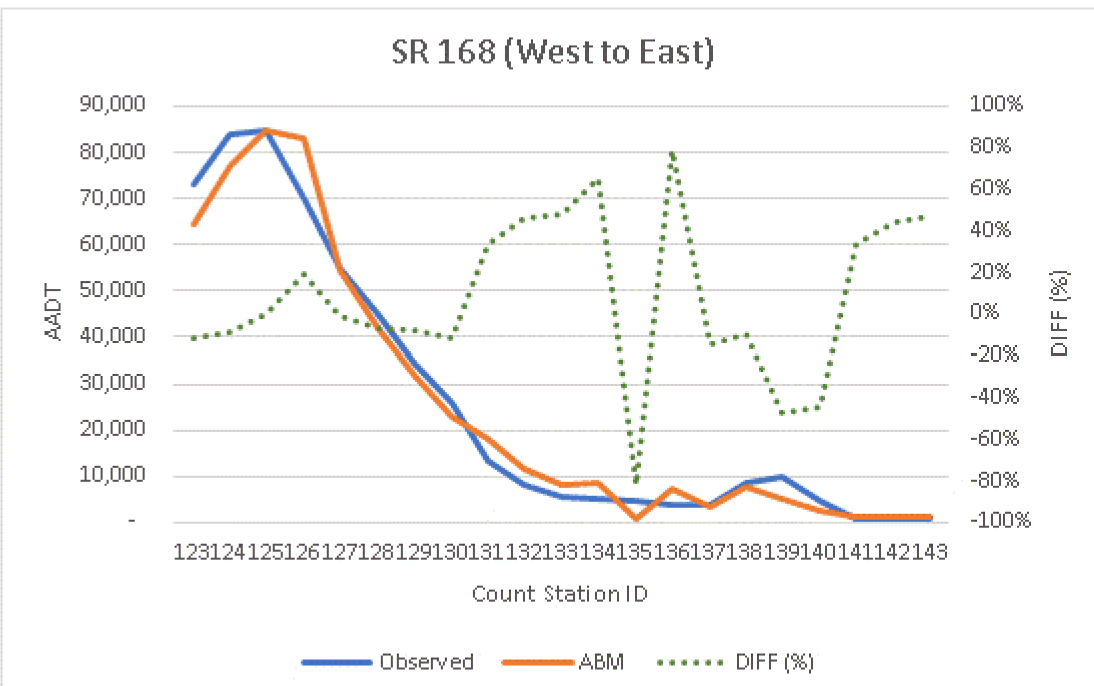


FIGURE 34: HIGHWAY VALIDATION – SR168

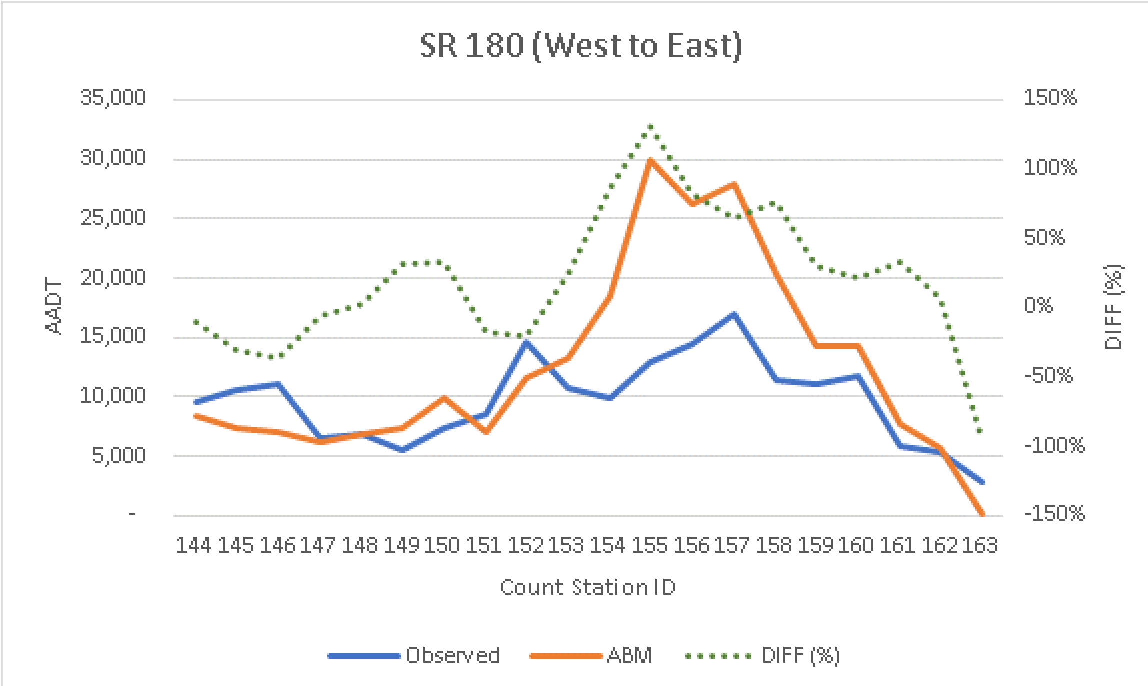


FIGURE 35: HIGHWAY VALIDATION – SR180

Time of Day

Diurnal performance of the model is compared in the four model time periods: AM (6am-9am), MD (9am-4pm), PM (4pm-7pm), and EV (7pm-6am). The comparison of the model's estimates and the observed traffic counts by four model time periods are presented in Table 38, Table 39, Table 40, and Table 41.

The ABM overestimates traffic volumes in the peak periods (AM and PM) and underestimates off-peak periods (MD and EV). Note that the diurnal count data base is primary comprised of counts from FresnoCOG, which appear to be more erroneous. There are only a few freeway counts by time-of-day. Therefore, the time of day comparisons could be biased and may not depicting the true picture of the distribution.

TABLE 37: HIGHWAY VALIDATION – AM AND PM PEAK PERIODS

PEAK PERIOD	COUNT	ABM	DIFF	DIFF (%)
AM (7am-8am)	413,753	433,600	19,847	5%
PM (6pm-7pm)	412,969	477,568	64,599	16%

TABLE 38: HIGHWAY VALIDATION - AM

FACILITY TYPE	COUNT	ABM	DIFF	DIFF (%)
Freeway & Highway	123,640	170,719	47,079	38%
Arterial	363,186	444,669	81,483	22%
Other (Collector, local & ramp)	306,604	327,766	21,163	7%
TOTAL	**793,430**	**943,155**	**149,725**	**19%**

TABLE 39: HIGHWAY VALIDATION - MD

FACILITY TYPE	COUNT	ABM	DIFF	DIFF (%)
Freeway & Highway	293,387	312,729	19,341	7%
Arterial	972,261	879,944	(92,316)	-9%
Other (Collector, local & ramp)	798,200	626,204	(171,995)	-22%
TOTAL	**2,063,847**	**1,818,877**	**(244,970)**	**-12%**

TABLE 40: HIGHWAY VALIDATION - PM

FACILITY TYPE	COUNT	ABM	DIFF	DIFF (%)
Freeway & Highway	146,878	200,665	53,787	37%
Arterial	512,102	553,089	40,987	8%
Other (Collector, local & ramp)	410,648	400,309	(10,339)	-3%
TOTAL	**1,069,629**	**1,154,063**	**84,435**	**8%**

TABLE 41: HIGHWAY VALIDATION - EV

FACILITY TYPE	COUNT	ABM	DIFF	DIFF (%)
Freeway & Highway	150,494	159,963	9,468	6%
Arterial	394,906	386,819	(8,087)	-2%

Other (Collector, local & ramp)	302,989	262,640	(40,348)	-13%
TOTAL	**848,389**	**809,422**	**(38,966)**	**-5%**

Transit Validation

Transit ridership produced by the model is compared against the observed ridership. The ridership (boarding) is compared regionally as well by transit line.

The FHWA provides guidelines to check reasonableness of the transit assignment results from a model. The relevant recommended guidelines are presented in Table 42.

TABLE 42: THE FHWA'S TRANSIT VALIDATION GUIDELINES

METRIC	THRESHOLD
Difference between actual counts and model results for a given year by Transit Mode (e.g. light rail, bus, etc.)	+/- 10%

*Source: The Travel Model Validation and Reasonableness Checking Manual, II Second Edition, September 2010.

Region

Regionally, Table 43, the ABM generates only 4% more transit boardings then the observed data and the corresponding transit trips in the ABM are also well matched with the observed transit trips (5%).The model indicates a boarding rate of 1.59 which is close to the value (1.60) calculated from the observed data.

TABLE 43: TRANSIT SUMMARIES - REGIONAL

MEASURE	OBSERVED	ABM	DIFF	% DIFF
boardings	41,143	42,968	1,826	4%
trips	25,707	26,958	1,251	5%
boarding rate	1.60	1.59	(0.01)	0%

Transit Line

A comparison of ridership by transit line examines the model's ability of producing transit ridership by transit line. A scatter plot in Figure 37 shows the relationship between the transit boardings from the ABM and the observed boarding by transit line. The X-axis in the plot represent the observed boardings and the estimated boardings from the model are presented on the Y-axis.

The regression line fitting all data points shows a R-squared value of 0.76 indicating that it is a reasonable fit. This suggests that the model is predicting the transit behavior reasonably well.

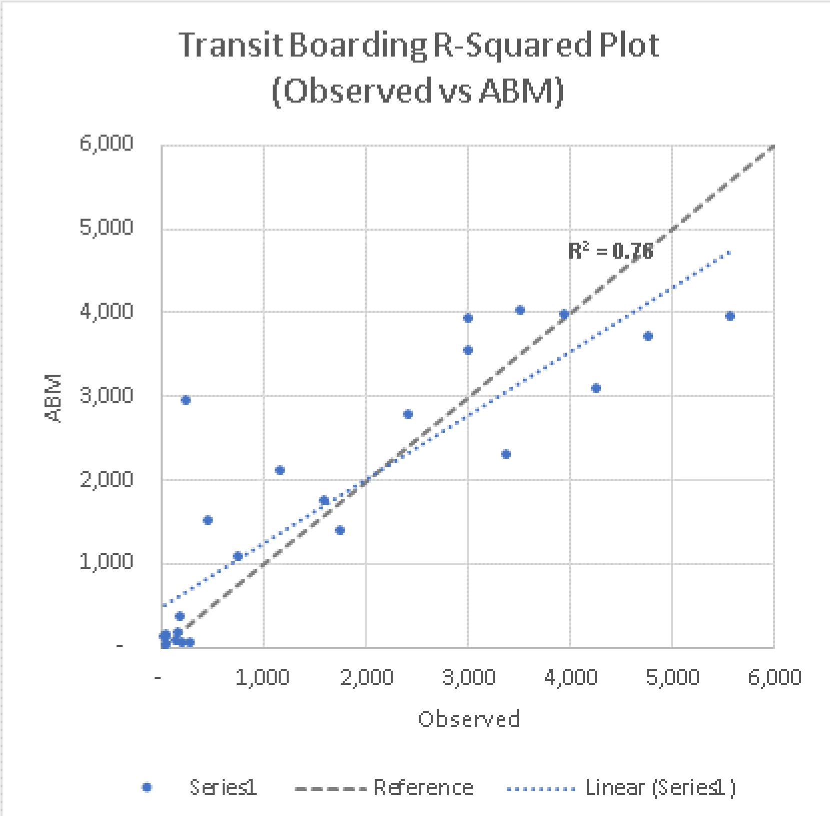


FIGURE 36: OBSERVED AND ESTIMATED TRANSIT BOARDINGS

A comparison of number of boardings by individual transit lines is presented in Figure 37. The X-axis is transit line id and the Y-axis is number of boardings. The transit lines are sorted from high observed boarding to low observed boarding. In general, the plot shows a reasonable match across all transit lines. There are three noticeable differences: FAX 45, Clovis 10, and Clovis 50. These three transit lines are overestimated and would need to be looked into in future model validation.

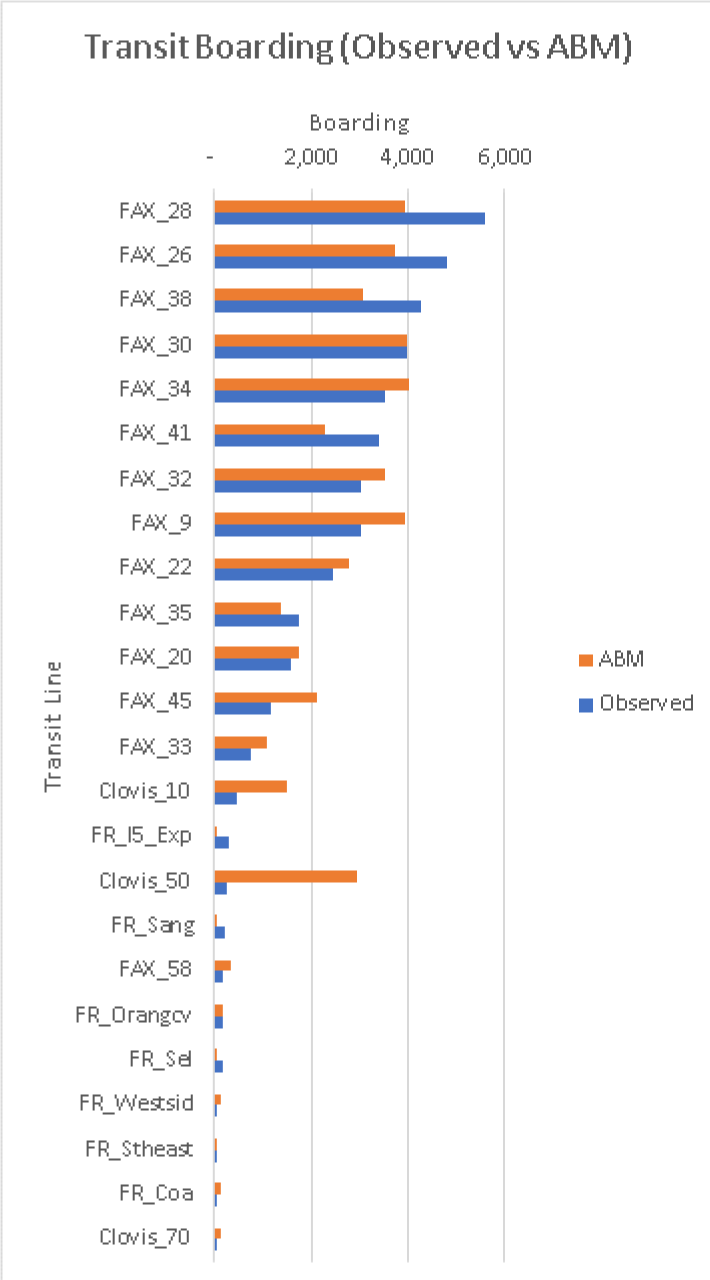


FIGURE 37: ESTIMATED AND OBSERVED BOARDINGS BY TRANSIT LINE

Bike and Ped Validation

Instead of validating every count location separately, the comparisons combine them into 13 groups. The groups are created based on the direction of travel or the neighborhood of the locations. Figure 38 displays the spatial distribution of the count locations and how they are grouped.

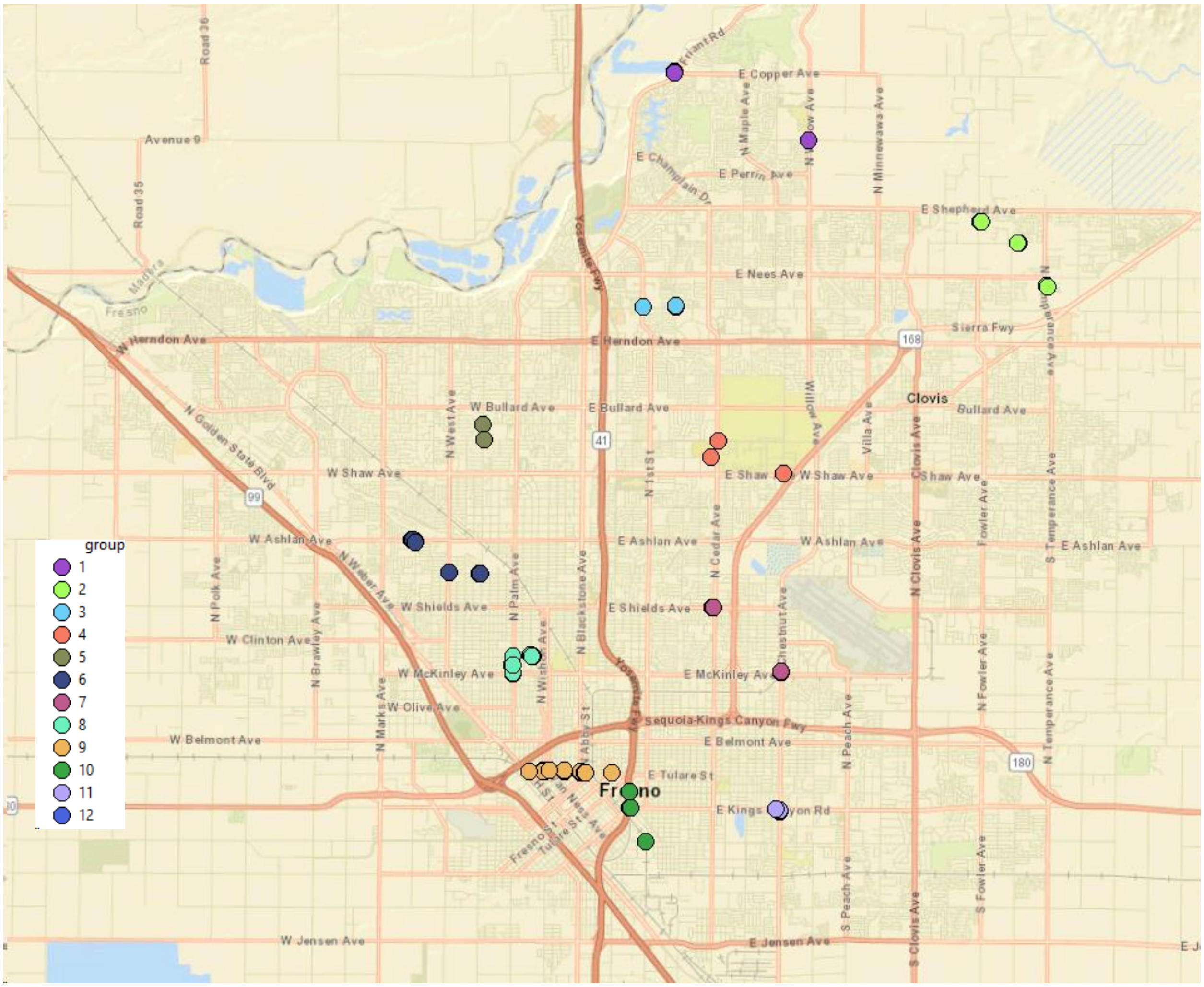


FIGURE 38: BIKE AND PED COUNT LOCATIONS AND COUNT GROUPS

Figure 39 presents a comparison of the estimated bike volume and the observed bike counts by the 13 count groups. A tabular summary of the same is presented in Table 44. Overall, bike volumes are overestimated by over 100% with many of the count groups are overestimated as well with about half of the count groups are overestimated by 500% or more. One group, Group 9 (downtown Fresno representing bike travel in north-south direction), underestimates the bike volumes in the ABM. Given that other count groups show overestimation, the underestimation for Group 9 points to a potential issue with the distribution of bike travel in the region.

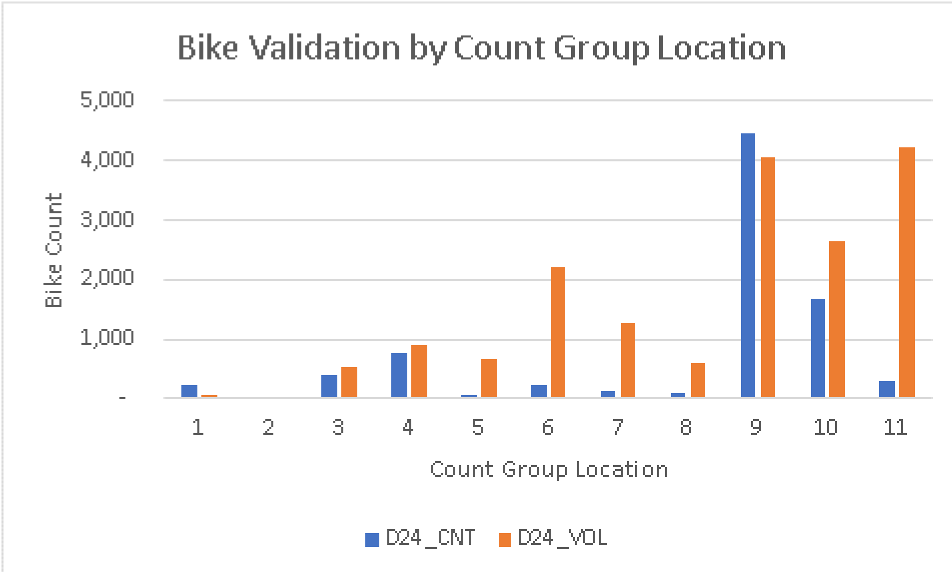


FIGURE 39: BIKE VALIDATION BY COUNT GROUP

TABLE 44: BIKE VALIDATION BY COUNT GROUP

GROUP	OBSERVED	ABM	DIFF	DIFF (%)
1	224	60	(164)	-73%
2	40	21	(19)	-47%
3	401	517	116	29%
4	780	913	133	17%
5	58	667	609	1049%
6	215	2,217	2,002	931%
7	113	1,286	1,173	1038%
8	90	610	520	578%
9	4,446	4,059	(387)	-9%
10	1,683	2,657	974	58%
11	290	4,226	3,936	1357%
Total	**8,340**	**17,233**	**8,893**	**107%**

Figure 40 and Table 44 presents a comparison of the estimated pedestrian volume and the observed pedestrian counts by the 13 count groups. A tabular summary of the same is presented in Table 45. Overall, the comparison of the pedestrian volumes appears alright with a slight underestimation (-1%). Even comparison by group look alright. The count group 4, near California State University, Fresno, significantly underestimated. However, this is not very surprising given that the model represents the university enrollment only in MAZ, therefore generating walk travel to from that MAZ only. However, the real behavior is very different. The university campus comprises of several buildings spread across the campus. In addition to walk trips to or from the campus, a significant number of walk trips are made within the campus to travel between the buildings. These internal walk trips are not represented in the model and therefore, the estimated walk travel is likely to be lower than the observed data.

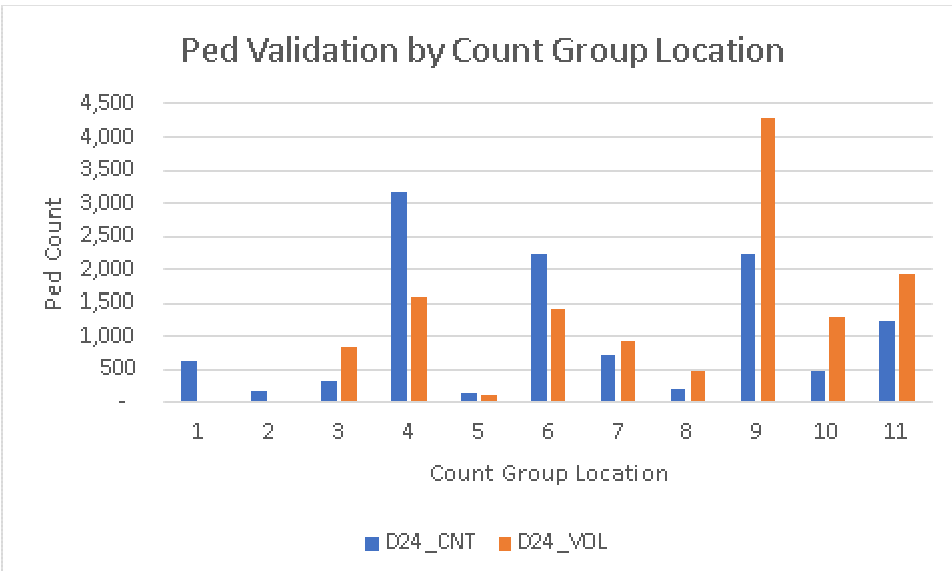


FIGURE 40: WALK VALIDATION BY COUNT GROUP

TABLE 45: WALK VALIDATION BY COUNT GROUP

GROUP	OBSERVED	ABM	DIFF	DIFF (%)
1	639	22	(617)	-97%
2	161	27	(134)	-83%
3	341	834	493	145%
4	3,155	1,603	(1,552)	-49%

5	150	121	(29)	-19%
6	2,226	1,408	(818)	-37%
7	716	927	211	29%
8	203	480	277	136%
9	2,234	4,278	2,044	92%
10	491	1,293	802	163%
11	1,243	1,926	683	55%
Total	**11,559**	**12,919**	**1,360**	**12%**

3.3 SUMMARY

The Fresno ABM is reasonably calibrated to match observed travel behavior reported in CHTS and NHTS datasets. The model validation also showed reasonable match of estimated volumes with observed traffic counts. The model meets most of the recommendations provided by FHWA for model reasonable checks. However, there are a few areas that still needs improvement and shall be addressed in future model development tasks:

- Highway validation showed overall root mean square error (RMSE) of 43% which is slightly higher than the recommended value of 40% or lower. Comparison by key highway corridors showed that the locations around Fresno compare well with observed counts, however, outside areas are not performing that well. That is probably due to the coarser representation of the roadway network in those areas and therefore resulting in missing important streets or incorrect location of the centroid connectors for the zones in those areas. A careful review of the roadway network may help further improve the ABM's performance.
- DaySim is calibrated to represent the observed travel behavior from two surveys: CHTS and NHTS. However, both surveys are older (2012 CHTS and 2009 NHTS) than the base year (2014) and therefore, possibly represent outdated travel patterns. This is evident from highway validation summaries as well. Even though the ABM reasonably predicts travel patterns shown in the calibration datasets, highway validation shows an underestimation of traffic flows north of Fresno downtown. Further, the calibration datasets were collected for the entire California state (CHTS) and the nation (NHTS), so only small number of Fresno residents were included in the surveys. The issue of small sample records was more evident during calibration by different market segments (ex. purpose,

mode). To overcome this issue, the current model calibration effort utilized survey records from the Fresno region for tour/trip destination and the entire San Joaquin Valley for the other model components. A more recent survey with sufficient number of respondents in the Fresno region would contribute greatly in better representing resident travel in the ABM.

- Transit calibration and validation also showed a need for a better transit on-board survey covering all transit services in the Fresno model region.
- Non-motorized flows were validated using count database constructed from multiple sources including a recent data collection effort and the bike count viewer. However, upon review of bike counts, the counts from the viewer were found suspect. A reliable and recent count data is needed to examine and improve, if needed, the model's ability to generate reasonable bike and ped flows.

4.0 SENSITIVITY TESTS

Sensitivity testing assesses the model's sensitivity to changing inputs like fuel prices, transit fares, new land uses, or new infrastructure. During this project, four sets of tests are designed to examine the Fresno ABM's sensitivity and include:

1. Auto operating Cost
2. Transit Fare
3. New Transit Service
4. New Employment Center

The first two set of tests (AOC and transit fare) are undertaken by RSG. The remaining two are conducted by the FresnoCOG staff with guidance from RSG.

4.1 CHANGE IN AUTO OPERATING COST (AOC)

Two model runs, Table 46, examine model's response to change in auto operating cost (AOC): double AOC and half AOC.

TABLE 46: SENSTIVITY TESTS – AUTO OPERATING COST

SENSITIVITY TEST	DESCRIPTION
Double AOC	AOC is 200% of the original AOC
Half AOC	AOC is 50% of the original AOC

The next sections start with describing process of setting up the ABM for these tests and then discuss the results.

Model Setup

The ABM use a consistent auto operating cost (AOC) value across the demand (DaySim) and the supply (network assignment) side. Both model systems read an input CSV file[15] that contains auto operating cost (cents per mile) by year.

The AOC values for the 2014 base year is 22.36 cents per mile. For each test, the input CSV file is updated with a new AOC value. For the double AOC test, the AOC value is increased to 44.71 cents per mile and for the half AOC test, the AOC value is decreased to 11.18 cents per mile.

[15] The input CSV file is here: 1_Inputs\6_Static\AutoOperatingCost.csv

Results

The outputs from the two model runs are compared with the outputs from the base year model run. The sensitivity of the model is measured by change in: tour lengths by purpose, tour/trip mode shares, and regional VMT.

Tour Lengths

As expected, Table 47, increase in auto-operating cost by 100% results in shorter tour lengths. That means residents of the Fresno County are now finding activity locations nearby due to higher cost of travelling by car. Discretionary activities (meal and social/recreation) are affected the most (-12%) and mandatory activities (work and school) the least (-5% and -2% respectively). This is a reasonable response because discretionary activities are generally for leisure purpose and are more likely to be reduced or performed nearby to limit the cost of travel. Whereas, going to work and school support the household and are therefore necessary travel with relatively lower flexibility to alter.

Decreasing the auto-operating cost increased tour lengths. The relative response by tour purpose is similar to the results of the double AOC test; mandatory (work and school) purpose see the least increase (3% and 2% respectively) and discretionary (8%) and escort (10%) purpose see the most increase. The results also suggest that people are likely to make more travel to drop-off/pickup (escort) their children if cost of travel is lower.

TABLE 47: AOC SENSITIVITY - TOUR LENGTHS (MILES) BY PURPOSE

TOUR PURPOSE	DOUBLE AOC	HALF AOC
Work	-5%	3%
School	-2%	2%
Maintenance	-9%	6%
Discretionary	-12%	8%
Escort	-9%	10%
Work-based	-8%	5%

Mode Share

As presented in Table 48, the drive alone tour mode sees the most decrease (-4.9%) due to a higher auto operating cost (double AOC). The higher occupancy auto modes show relatively lower sensitivity with shared-ride (SR) 2 losing only 1.8% of the original

SR2 tours. The shared-ride 3+ mode even received more tours (0.4%). Mostly travelers switch to non-auto modes to minimize increase cost of travel. All non-auto modes including transit (walk or drive), bike, and walk see more than 10% increase in their tours. The bike mode sees the most relative increase in tours (23.4%). The results suggest that higher cost of travel pushes travelers to either choose a higher occupancy auto mode or travel by non-auto modes (transit, bike and walk).

Lowering the operating cost (half AOC) results in more tours by auto mode and fewer by non-auto modes. The travelers are making more drive-alone tours (2.8%) because of its lower cost. Even higher occupancy (SR3+) travelers are choosing for a lower occupancy travel (drive-alone or SR2). The non-auto modes generally respond similarly with bike mode observing the most relative decrease in tours.

TABLE 48: AOC SENSITIVITY - TOUR MODE SHARE

TOUR MODE	DOUBLE AOC	HALF AOC
Drive Alone	-4.9%	2.8%
Shared Ride 2	-1.8%	0.8%
Shared Ride 3+	0.4%	-0.6%
Drive-Transit	11.2%	-6.1%
Walk-Transit	10.4%	-5.7%
Bike	23.4%	-8.2%
Walk	10.2%	-5.0%
School Bus	-1.2%	0.6%
Total	**-0.5%**	**0.3%**

The change in trip mode shares, Table 49, also exhibit the same behavior as the tour mode shares.

TABLE 49: AOC SENSITIVITY - TRIP MODE SHARE

TRIP MODE	DOUBLE AOC	HALF AOC
Drive Alone	-3.2%	1.3%
Shared Ride 2	0.1%	-0.5%
Shared Ride 3+	2.2%	-1.8%

Transit	12.4%	-7.3%
Bike	24.9%	-8.4%
Walk	11.4%	-5.8%
School Bus	-1.0%	0.2%
Total	0.7%	-0.7%

Regional VMT

The regionwide total vehicle miles travelled (VMT), Table 50, responds expectedly to the change in auto operating cost. Doubling the auto operating cost reduce the VMT by 6.28%, whereas halving the cost increase the VMT by 4.10%.

To examine magnitude of the VMT response, a mid-link elasticity value in response to the AOC change are calculated and compared with the elasticities reported in the 2013 ARB report[16],. The ARB analyzed several studies determining the impact of gap prices on VMT. Across the examined studies, the report observed short-run elasticities in the range of -0.03 to -0.10. The ABM is producing elasticities of -0.10 and -0.06 for double AOC and half AOC. These elasticities are within the range observed in the ARB report, suggesting that the ABM is appropriately sensitive to the auto operating cost.

TABLE 50: AOC SENSITIVITY – REGIONAL VMT

MEASURE	DOUBLE AOC	HALF AOC
VMT	-6.28%	4.10%
Elasticity	-0.10	-0.06
ARB Elasticity Range	-0.03 to -0.10	

4.2 CHANGE IN TRANSIT FARE

Two model runs, Table 51, examine model's response to change in transit fare: double transit fare and half transit fare.

TABLE 51: SENSITIVITY TESTS – TRANSIT FARE

SENSITIVITY TEST	DESCRIPTION
Double Transit Fare	Transit fare is 200% of the original fare

[16] See the following document: https://www.arb.ca.gov/cc/sb375/policies/gasprice/gasprice_brief.pdf

Half Transit Fare	Transit fare is 50% of the original fare

The next sections start with describing the process of setting up the ABM for these tests and then discuss the results.

Model Setup

The 2014 base year ABM apply two sets of fare for the transit services in the Fresno County. The first set[17] is a flat fare of $1.25 for the local buses operated by FAX and Clovis. The other set[18] is a distance-based fare for the inter-city buses by FCRTA. Both sets of fares are updated in the transit fare sensitivity tests.

Results

The outputs from the two model runs are compared with the outputs from the base year model run. The sensitivity of the model is measured by change in: tour/trip mode shares, transit trips by income groups and regional transit ridership (boardings).

Mode Shares

As presented in Table 52, the tours made using transit see a significant decrease due to increase in transit fare. Further, the transit tours that are driving to transit (-8.6%) are affected less than the transit tours walking to transit (-14.2%). This is expected as drive to transit tours are more likely to be mandatory travel and made by travelers that are from relatively higher income households, therefore, such travelers are less sensitivity to change in transit cost.

The tours leaving transit are now choosing other modes as the non-transit modes see increase in their respective tours, though the increase is not big due to very low share of transit tours overall (1.2%) – see 3.1 Calibration.

TABLE 52: TRANSIT SENSITIVITY - TOUR MODE SHARE

TOUR MODE	DOUBLE FARE	HALF FARE
Drive Alone	0.2%	-0.1%
Shared Ride 2	0.2%	-0.1%
Shared Ride 3+	0.1%	-0.1%
Drive-Transit	-8.6%	3.0%
Walk-Transit	-14.2%	9.4%

[17] This is provided in “1_inputs\4_Transit\FC_BASE_TRAN_FAR.FAR”
[18] This is provided in 1_Inputs\4_Transit\fareMatrix.txt

Bike	0.5%	-0.5%
Walk	0.2%	-0.1%
School Bus	0.1%	-0.1%
Total	0.0%	0.0%

The change in trip mode shares, Table 53, also exhibit the same behavior as the tour mode shares.

TABLE 53: TRANSIT FARE SENSITIVITY - TRIP MODE SHARE

TRIP MODE	DOUBLE FARE	HALF FARE
Drive Alone	0.1%	-0.1%
Shared Ride 2	0.1%	-0.1%
Shared Ride 3+	0.1%	-0.1%
Transit	-14.9%	9.9%
Bike	0.2%	-0.1%
Walk	0.5%	-0.3%
School Bus	0.1%	-0.1%
Total	0.0%	0.0%

Transit Trips by Household Income

Table 54 examines traveler's sensitivity to change in transit fare by household income. The table looks at transit travelers within 6 household income categories and calculates percentage change in their transit trips due to change in transit fare. For both increase and decrease in transit fare, travelers in lower income categories are the most sensitive to the price change. The sensitivity to the transit fare change decreases with increase in average household income. The results make sense as higher income people have high value of time and are less sensitive to price. The lower income people have low value of time and are more sensitive to price. Due to change in price, such travelers are probably switching between transit and non-motorized modes (bike and walk) to make their travel.

TABLE 54: TRANSIT FARE SENSITIVITY – TRANSIT TRIPS BY HOUSEHOLD INCOME

HOSEHOLD INCOME	DOUBLE FARE	HALF FARE

$0-$15,000	-18%	12%
$15,000-$35,000	-17%	11%
$35,000-$50,000	-15%	8%
$50,000-$75,000	-13%	9%
$75,000-$100,000	-9%	8%
$100,000+	-8%	5%
Total	**-15%**	**10%**

Transit Boardings

The regionwide transit ridership, Table 55, responds appropriately to the change in transit fare; doubling the transit fare reduce the ridership by 14.51% and halving the transit fare increase the ridership by 9.36%.

To examine magnitude of the ridership response, a mid-link elasticity value in response to the AOC change are calculated and compared with the elasticities reported in the 2013 ARB report[19]. The ARB analyzed several studies determining the impact of change (decrease) in transit fare on transit ridership. Across the examined studies, the report observed short-run elasticities close to 0.40. The ABM is producing elasticities of -0.23 and -0.13 for double transit fare and half transit fare. These elasticities are lower than the elasticities observed in the ARB report.

TABLE 55: TRANSIT FARE SENSITIVITY – TRANSIT BOARDINGS

MEASURE	DOUBLE FARE	HALF FARE
Boardings	-14.51%	9.36%
Elasticity	-0.23	0.13
ARB Elasticity Range	-0.40	

4.3 NEW TRANSIT SERVICE

To evaluate the model's response to a new transit service, a new Bus Rapid Transit (BRT) line is added to provide service to the Blackstone and Kings Canyon corridors in

[19] See the following document: https://www.arb.ca.gov/cc/sb375/policies/gasprice/gasprice_brief.pdf

Fresno for the base year 2014. The BRT line is coded with peak period headways of 5, 10, and 15 to test model sensitivity. Figure 41 shows the route of the BRT line.

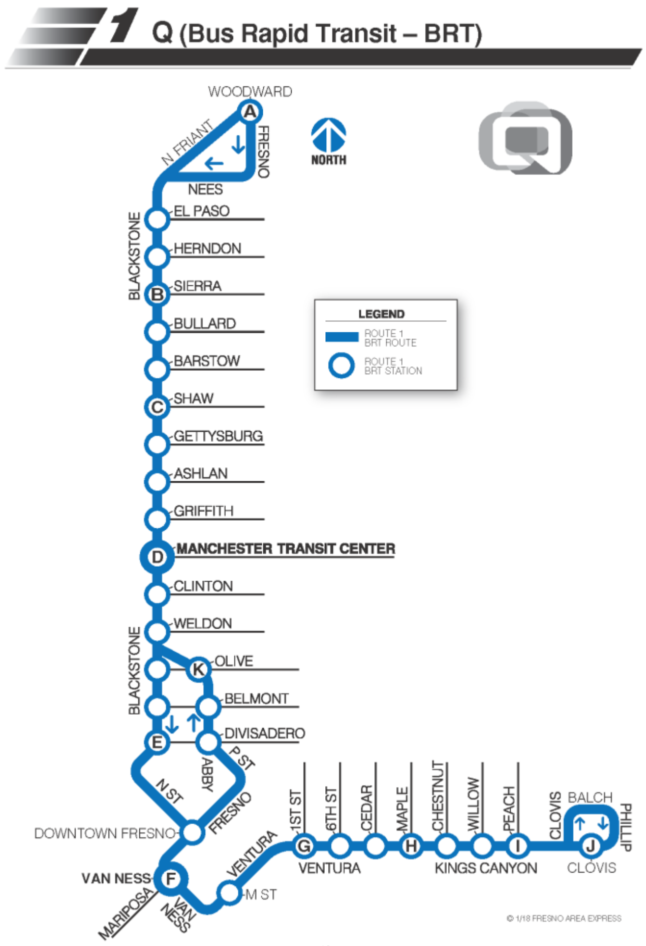


FIGURE 41: SENSITIVITY TEST - NEW BRT SERVICE

Model setup

The ABM already has a BRT mode as a placeholder mode for future scenarios. The new BRT line was simply coded to the transit network[20] to make it effective in the ABM.

Results

Addition of the BRT line into the 2014 base year transit system results in only slight increase in transit ridership regionally. However, the new BRT line itself received reasonable daily boardings – 2,516 boardings with 10mins headway. The low increase in systemwide ridership is reasonable given that there is competition among routes in the

[20] The transit network is here: 1_Inputs\3_highway\FresnoNetwork_Deliverable3.gdb\PTNetwork_14

BRT corridor. Therefore, the boardings on BRT are mostly on the expense of the competing routes and are not the result of new riders.

Model runs with different BRT headways further showed model's sensitivity to transit service frequency, Table 56, decrease in frequency (5mins) resulted in more boardings and increase (15mins) in fewer boardings.

TABLE 56: SENSITIVITY TEST – NEW BRT SERVICE BOARDINGS

BRT HEADWAY	TRANSIT BOARDING
5 mins	3,517
10 mins	2,516
15 mins	1,828

4.4 NEW EMPLOYMENT CENTER

Sensitivity to change in employment is examined by adding a new employment center in the industrial park south of downtown Fresno. The employment center is assigned with 2,708 additional employment, concentrated on three MAZs, Table 57, which covers an area defined by North Avenue to the north, East Avenue to the west, Central Avenue to the south, and State Route 99 and Cedar Avenue to the east (see Figure 42).

TABLE 57: SENSITIVITY TEST – NEW EMPLOYMENT CENTER

MAZ_ID	EMP_NEW	EMP_BASE	EMP_CENTER
12090015001053	1,146	4	1,142
12090015001054	230	0	230
12100015001052	1,337	1	1,336

Two more model runs further tested model's sensitivity with respect to magnitude of employment by decreasing (50%) and increasing (150%) employment.

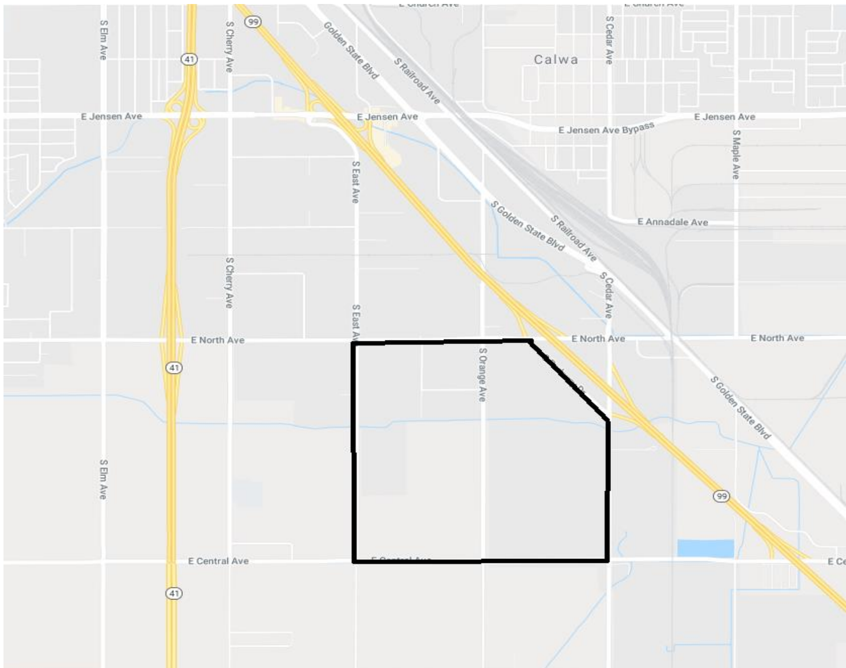


FIGURE 42: SENSITIVITY TEST – NEW EMPLOYMENT CENTER

Model setup

Adding the employment center required only updating the DaySim micro-zone file to include additional employment for the center. As this test made edits to three MAZs, the micro-zone file was updated manually to represent total employment, as well as by 8 DaySim employment categories. The instructions on how to create a new DaySim micro-zone file are provided in Appendix C.

Note that the buffering tool needs to be re-run as well to create a new buffered micro-zone file. However, as the tool is embedded in the model system, it is automatically re-run before a new DaySim run.

Results

The new employment center results in additional 7,000 daily trips regionwide. Change in employment amount (50% and 150%) produced intuitive results for peak hour and daily trips, Table 58.

The spatial distribution of the extra travel, Figure 43, shows a reasonable pattern on the network. Due to its proximity to the new employment center, most of the travel is using SR-99 to access the new center.

TABLE 58: SENSITIVITY TEST – NEW EMPLOYMENT CENTER TRIPS

EMPLOYMENT LEVEL	AM PEAK HOUR	PM PEAK HOUR	DAILY

50%	361	354	3,918
100%	654	624	7,058
150%	950	888	10,034

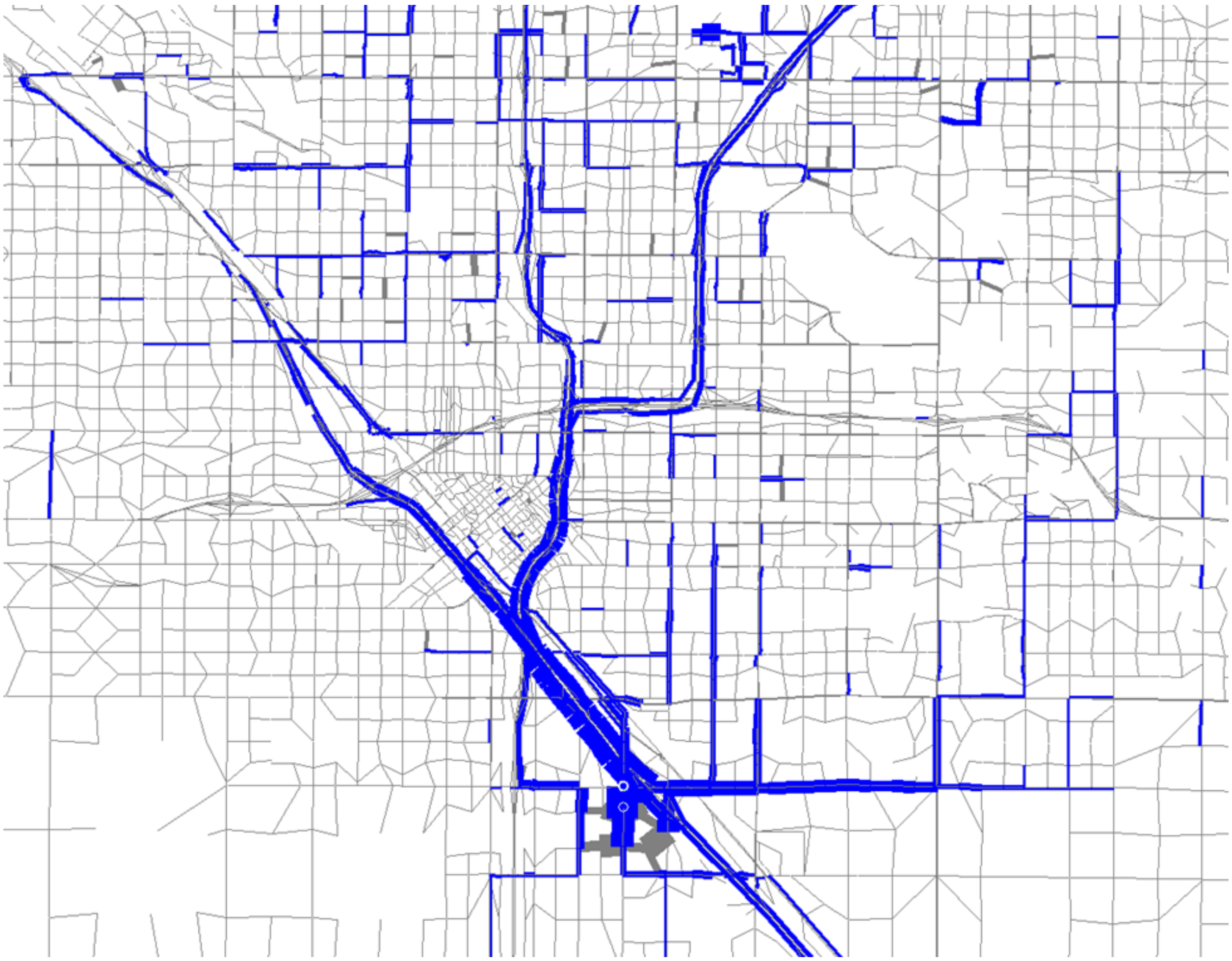


FIGURE 43: SENSITIVITY TEST - CHANGE IN DAILY TRAFFIC VOLUME

4.5 SUMMARY

Four sets of tests are conducted to examine the Fresno ABM's sensitivity: auto operating cost, transit fare, new transit service, and new employment center. The results of change in these inputs were compared with base scenario and appropriate measures were calculated to analyze change in results. Whenever, possible the change in results (elasticity) were compared with ARB guidelines to assess reasonableness of the change. In general, the model responded appropriately to change in various input changes. Model elasticities were also within ARB guidelines except for change in transit fare where the model showed lower sensitivities compared to the ARB guidelines.

5.0 FRESNO ABM USER'S GUIDE

This chapter describes model setup and steps to run the Fresno ABM for base year 2014. It also describes structure and content of the model directory.

5.1 SOFTWARE REQUIREMENTS

The instructions are for a machine with Windows operating systems (Windows 7). To setup PopSynIII, following software are required:

Cube Voyager

Cube version 6.4.2 should be installed on the machine.

Python - Anaconda

Anaconda is a python data science platform. It contains a lot of useful python libraries. We are using a python script in our CUBE work flow that would require some python libraries. Instead of installing libraries separately it is better to install an Anaconda package as that would contain all the required libraries.

Latest Anaconda can be downloaded from here:

https://www.anaconda.com/download/

Here are instructions on installing Anaconda on windows:

https://docs.anaconda.com/anaconda/install/windows

5.2 SETUP ABM

System Setup

Once all the required software are installed (5.1Software Requirements), user need to copy following files to their respective directories:

OMX DLL

The AB model use an OMX DLL to convert Cube skims to OMX format, which are then read in DaySim. User need to copy the OMX DLL ("OMXLib.dll") to cube voyager installation directory (C:\Program Files\Citilabs\CubeVoyager). The DLL is available under the *App* directory in the model setup.

Python Resource File

The ABM model uses a Python script to run it within Cube environment. To do that Cube creates a python resource file ("PYTHON.Rsc") that primarily contains python program path. User needs to update that path (see Figure 44) and copy the resource file to the following location:

C:\Users\<username>\AppData\Roaming\Citilabs\Cube\User

The resource file is available under the *App* directory in the model setup.

```
* RSC file for User Program PYTHON
*
#PYTHON

&ProgPath=C:\Program Files\Anaconda2\python.exe
&ProgDesc=Python Software for Python Programming Language Programs Inclusion
&ProgType=1
&ProgUI=0
&UsePath=0


 &FILES
1100
 0,"Script File   ","*  ",1,PYT,"      ",1
 1,"Data File 1   ","*  ",1,PYT,"      ",1
 2,"Data File 2   ","*  ",1,PYT,"      ",1
 3,"Data File 3   ","*  ",1,PYT,"      ",1
 4,"Data File 4   ","*  ",1,PYT,"      ",1
 5,"Data File 5   ","*  ",1,PYT,"      ",1
 6,"Data File 6   ","*  ",1,PYT,"      ",1
 7,"Data File 7   ","*  ",1,PYT,"      ",1
 8,"Data File 8   ","*  ",1,PYT,"      ",1
 9,"Data File 9   ","*  ",1,PYT,"      ",1
10,"Data File 10  ","*  ",1,PYT,"      ",1

 &PARAM
 &PARHLP
 &PAREG

 &OPTION

 &OPTHLP

 &COMLIN
Inputs
Outputs

&END
```

FIGURE 44: PYTHON RESOURCE FILE – UPDATE PROGRAM PATH

Model Setup

No additional steps are required to setup the AB model.

5.3 RUN ABM

To run ABM, open the catalogue file (*Fresno_BASE_20170208.cat*) in Cube. A catalogue file is a group of applications (steps) run in a pre-defined order. When opened in Cube, user should see four frames on the left side (see Figure 45): Scenario, Data, App, and Keys.

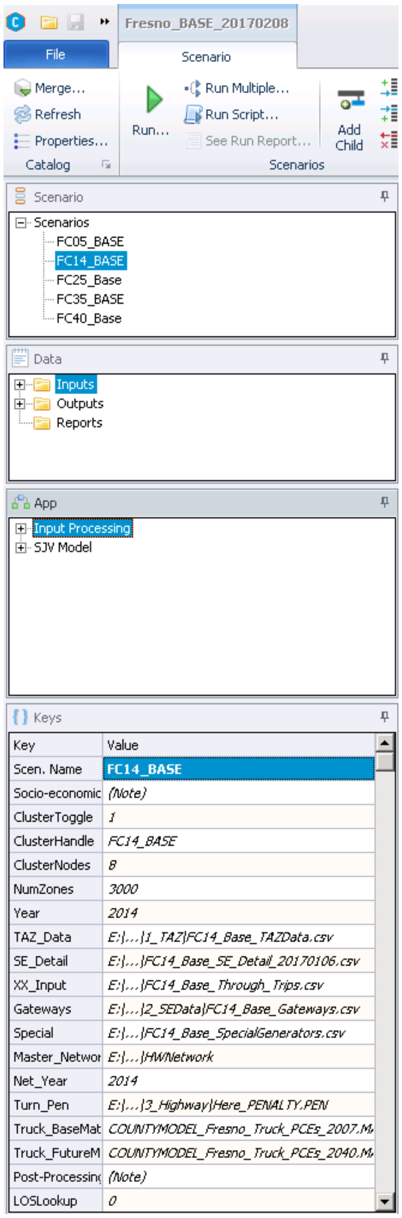


FIGURE 45: MAIN CUBE CATALOG FILE

The *Scenario* frame contains a list of scenarios to pick for the current run. User need to pick a scenario under this frame. For 2014, user would select FC14_BASE. The *Data* frame contains a list of inputs and outputs. User do not need to do anything here. The *Keys* section contains a list of parameters. Again, user do not need to anything here.

The *App* section contains two groups: Input Processing and SJV Model, which are run sequentially and separately.

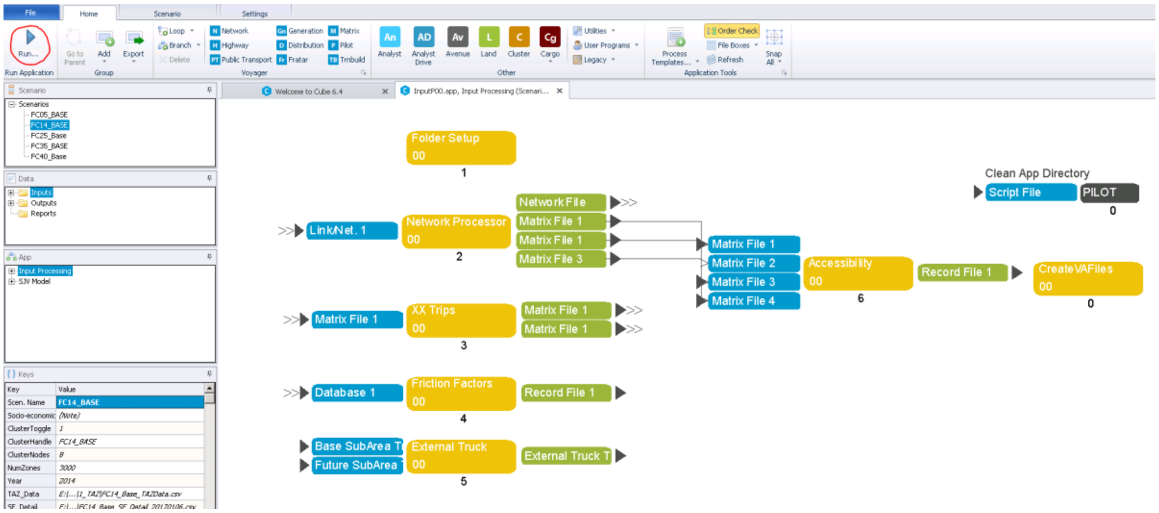
The Fresno AB model, similar to MIP2, is run in two steps:

- First step, input processing, prepares initial inputs for model to use in feedback loops. The processing includes setting up the working directory, creating model networks, initial skimming, external trips, truck model etc. As bike and walk skimming is performed only once, it is also included in the input processing.

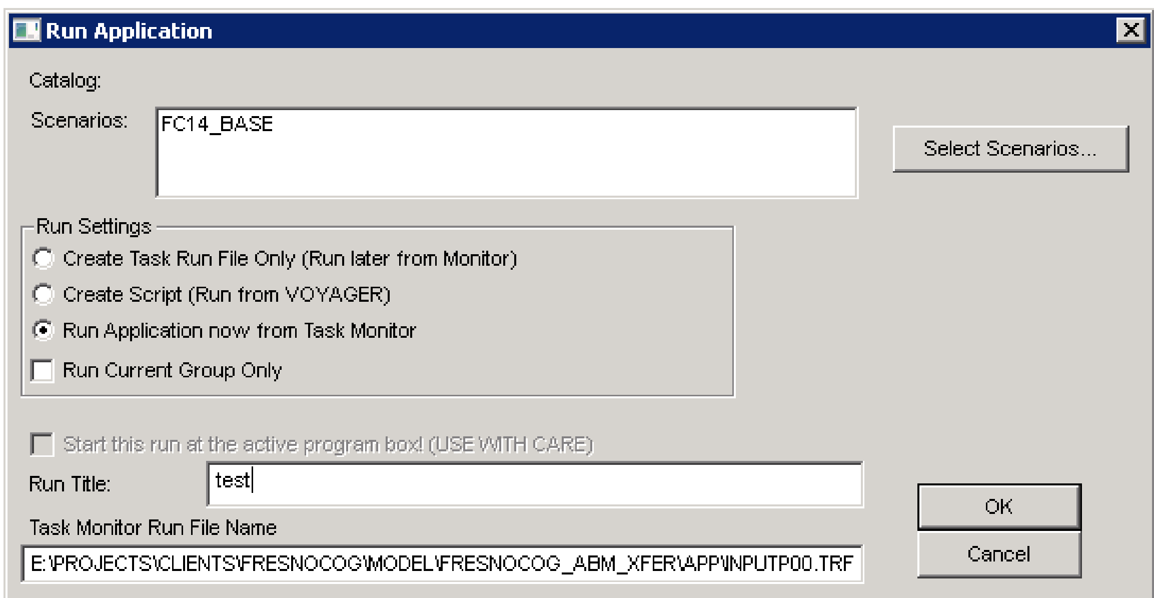
- The second step, SJV Model, includes feedback loops and final assignment. A feedback loop start with generating highway and transit skims. After generating external and truck demand, it runs DaySim. This produces demand matrices by mode and time period. The demand matrices are then assigned on highway and transit networks for AM and MD periods. After three feedback loops, a final assignment of all four periods is performed. Assignment also include bike and walk assignments.

IMPORTANT: Here are sequential steps to run ABM:

1. Open *Fresno_BASE_20170208.cat* in Cube.
2. Select a scenario under *Scenarios* section.
3. First run “Input Processing” by double clicking on *Input Processing* under *App* and click Run on the Home ribbon. (Note: if “Input Processing” is open for the first time in a new setup then Cube will prompt you to update model working directory to the current catalog directory. Accept the change and proceed.)



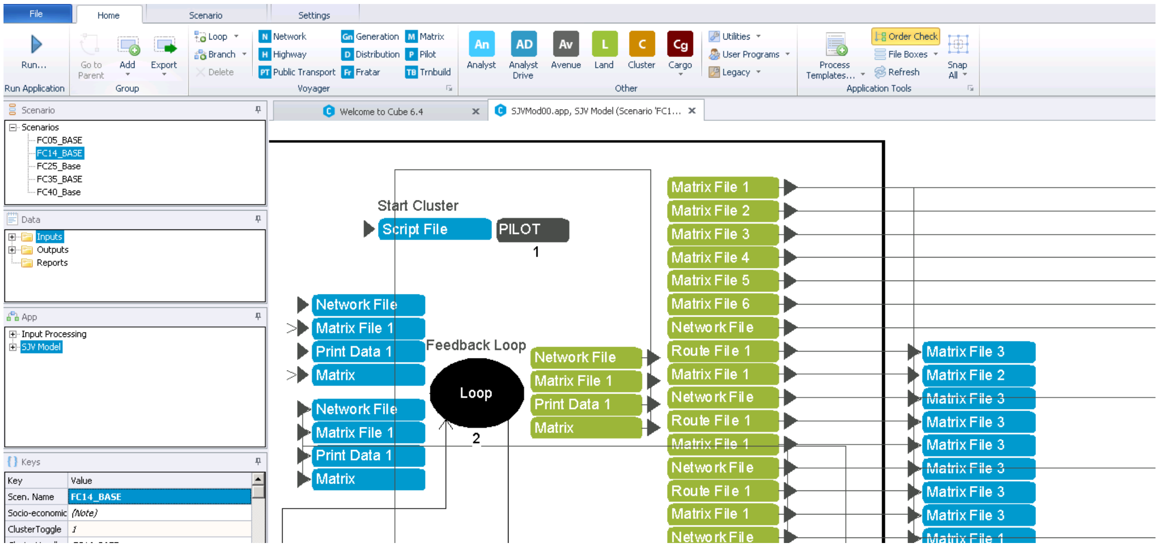
4. This will bring up a Run Application window. Enter “Run Title” of your choice and click OK. This will start Cube processing.



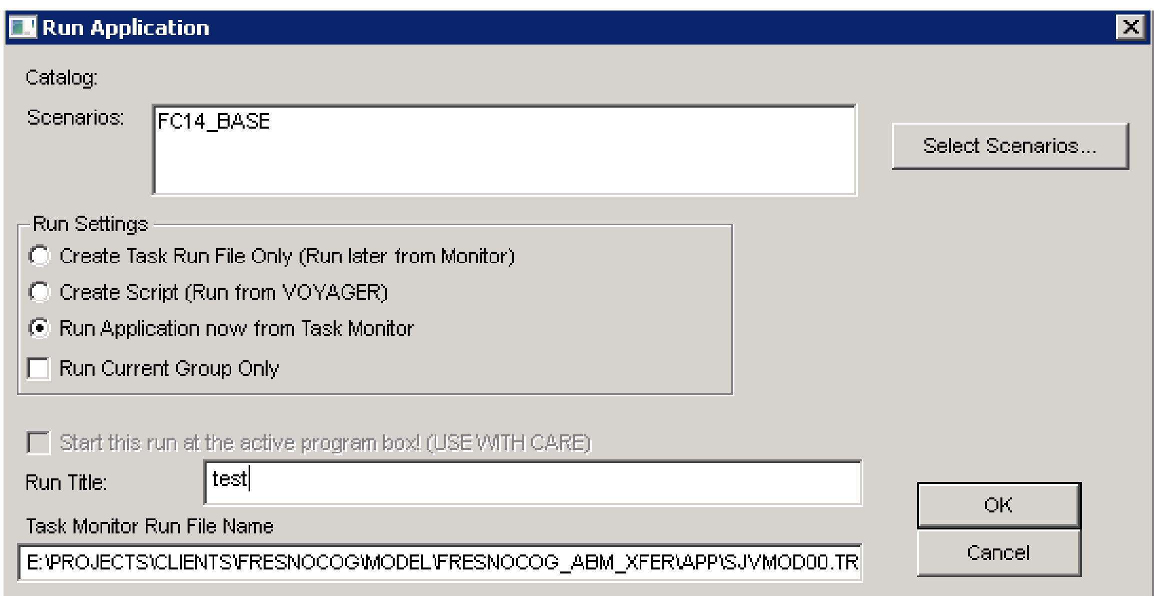
<u>**Note:**</u> after you click OK, Cube may prompt a message saying that errors have been found while checking for files. Ignore the error and choose “Create Job Anyway”.

5. Now run “SJV Model” by double clicking on *SJV Model* under *App* and click Run on the Home ribbon. (Note: if “SJV Model” is open for the first time in a new

setup then Cube will prompt you to update model working directory to the current catalog directory. Accept the change and proceed.)



6. This will bring up a Run Application window. Enter “Run Title” of your choice and click OK. This will start Cube processing.



Note: after you click OK, Cube may prompt a message saying that errors have been found while checking for files. Ignore the error and choose “Create Job Anyway”.

5.4 MASTER DIRECTORY STRUCTURE

Figure 46 presents directory structure for the Fresno Activity-Based Model setup. The structure includes initial setup, as well as directories after a model run is finished.

The primary model file linking the entire ABM system is in the root folder and is called as “Fresno_ABM.cat”. This catalogue file is used to view and run the model using Cube Voyager.

The directory “*1_Inputs”* holds all the input data used by the ABM system. DaySim inputs are also stored in an appropriate sub-directory (*8_DaySim*) within the inputs directory. The *App* directory houses cube scripts and other applications (DaySim executable, parcel buffer, and population sampler). The *GIS* folder contains geodatabase and ArcMap documents.

Once a model run is finished, a new directory *Scenarios* is added to the initial model directory. The output directory (*Scenarios*) contains scenario specific sub-folders (ex. FC14_BASE for year 2014). The scenario specific folders contain outputs organized by model components. In addition, a temp folder contains intermediate files produced during a scenario run.

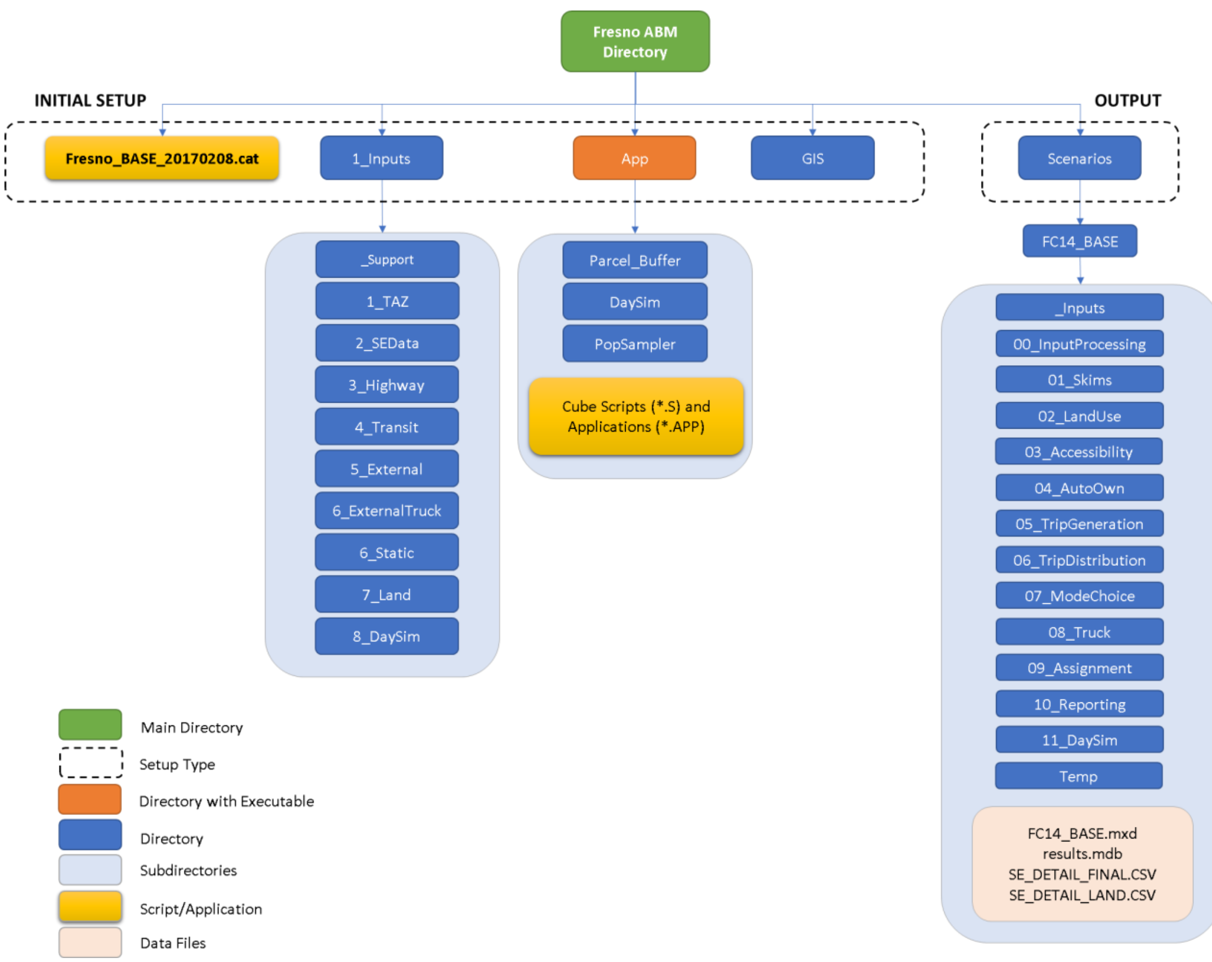


FIGURE 46: FRESNO ABM DIRECTORY STRUCTURE

5.5 INPUTS

This section provides description and structure of input files required in Fresno ABM.

DaySim

This section describes input files required for a DaySim run. The DaySim input files include: synthetic population (household and person files), micro-zone land-use, zone index, worker IXXI fraction, PNR nodes, coefficients, roster, roster combinations, and configuration.

Micro-zone Land-use File

The input micro-zone file provides information on location and land-use for micro-zones in the Fresno region. The file is in comma separated value format and is located at “1_Inputs/8_DaySim/02_Parcel/maz_2014_parks.csv”. This file is created in-house by Fresno COG. Table 59 presents a list of fields available in the input micro-zone file.

Off-street parking location and pricing information is used in the activity-based model system to influence mode and other choices. Note that this parking information is focused on publicly accessibility off-street locations and does not consider private off-street parking locations (such as those available only to workers in an office building), nor does it consider on-street parking location. Future year parking locations and costs can be easily added to the model system by simply updating the input micro-zone file to identify parking capacity and costs for individual micro-zones. Where data is unavailable or unknown, the parking attributes should be set to 0.

TABLE 59: MICRO-ZONE FILE FORMAT

HEADER LABEL	VALID VALUES	DESCRIPTION, COMMENTS
parcelid	1 – 9999999	The parcel ID number. Values must be unique positive integers, in ascending order. (Gaps are allowed, but not efficient for memory.)
xcoord_p	1-999999999	The x coordinate of the parcel centroid, in integer length units (typically SPF).
ycoord_p	1-999999999	The y coordinate of the parcel centroid, in integer length units (typically SPF).
sqft_p	0-999999999	The area of the parcel, in thousands of square length units (typically sqf, does not need to be an integer)
taz_p	1-9999999	The zone that the parcel is in. Must be a valid zone_id in the “zone” file

HEADER LABEL	VALID VALUES	DESCRIPTION, COMMENTS
block_p	0-9999999	This variable is used to store park area in square feet
hh_p	Real >=0	The number of households residing on a parcel.
stugrad_p	Real >=0	The number of grade school (K-8) students enrolled in schools on a parcel
stuhgh_p	Real >=0	The number of high school (9-12) students enrolled in schools on a parcel. If this is not available separately, then set to 0 & put the number of K-12 students in stugrd_p
stuuni_p	Real >=0	The number of university/college students enrolled in schools on a parcel.
empedu_p	Real >=0	The number of educational employees working on a parcel
empfoo_p	Real >=0	The number of food service employees working on a parcel
empgov_p	Real >=0	The number of government employees working on a parcel
empind_p	Real >=0	The number of industrial employees working on a parcel
empmed_p	Real >=0	The number of medical employees working on a parcel
empofc_p	Real >=0	The number of (other) office employees working on a parcel
empret_p	Real >=0	The number of retail employees working on a parcel
empsvc_p	Real >=0	The number of (other) service employees working on a parcel
empoth_p	Real >=0	The number of other sector employees working on a parcel. Typically contains construction, agriculture, mining.
emptot_p	Real >=0	The total number of employees working on a parcel. Should equal the sum of the 9 previous fields.
parkdy_p	Real >=0	The number of paid public off-street parking spaces on a parcel with per day pricing. (May overlap with parkhr_p if have both types of pricing.)
parkhr_p	Real >=0	The number of paid public off-street parking spaces on a parcel with per hour pricing. (May overlap with parkdy_p if have both types of pricing.)
ppricdyp	Real >=0	The average price of public off-street parking spaces on a parcel with per day pricing. (In cents per day)
pprichrp	Real >=0	The average price of public off-street parking spaces on a parcel with per hour pricing. (In cents per hour)

Synthetic Population

DaySim requires household and person files in ASCII delimited format with a header record. The PopSynIII software produces synthetic population in two CSV files: household and person. A python script then converts these files into DaySim format to use in a DaySim run. The Household file is "1_Inputs\8_DaySim\03_Household\Fresno_household.dat" and the person file is "1_Inputs\8_DaySim\03_person\Fresno_person.dat"

Table 60 provides a list of fields available in the household file. Of the variables listed below, only five (hhno, hhsize, hhincome, hhparcel, and hhexpfac) are strictly needed by DaySim as inputs on the raw data file. One (hhvehs) is predicted by DaySim. The rest can be computed based on other data files or aren't currently used in model application. DaySim also adds two variables to the working and output household file ("fraction_with_jobs_outside", which is a property of the residence zone, and "zone_id" which is DaySim's internal zone ID corresponding to "hhtaz").

TABLE 60: HOUSEHOLD FILE FORMAT

HEADER LABEL	VALID VALUES	DESCRIPTION, COMMENTS
hhno	1 - 9999999	The household ID number. Values must be unique positive integers, and should be in ascending order.
hhsize	1 – 99	The number of persons in the household. Must equal the number of person records for the household in the raw person file
hhvehs	0 - 99	The number of autos in the household. (This could be made optional as input, as it is predicted by DaySim.)
hhwkrs	0 – 99	The number of workers in the household. (This could be made optional as input, as it is computed by DaySim from the person records.)
hhftw	0 – 99	The number of HH members with person type=full-time worker. (This could be optional as input, as it is computed by DaySim.)
hhptw	0 – 99	The number of HH members with person type=part-time worker. (This could be optional as input, as it is computed by DaySim.)
hhret	0 – 99	The number of HH members with person type=retired adult. (This could be optional as input, as it is computed by DaySim.)
hhoad	0 – 99	The number of HH members with person type=other non-working adult. (This could be optional as input, as it is computed by DaySim.)

hhuni	0 – 99	The number of HH members with person type=university student. (This could be optional as input, as it is computed by DaySim.)
hhhsc	0 – 99	The number of HH members with person type=grade school student age 16+. (This could be optional as input, as it is computed by DaySim.)
hh515	0 – 99	The number of HH members with person type=child age 5-15. (This could be optional as input, as it is computed by DaySim.)
hhcu5	0 – 99	The number of HH members with person type=child age 0-4. (This could be optional as input, as it is computed by DaySim.)
hhincome	-1 – 9999999	The household annual income, in integer dollars. (A negative value is interpreted as missing data in DaySim estimation mode.)
hhownrent	1 – 9	Household own versus rent status. (This could be optional as input, as it is not currently used in the DaySim model code.)
hhrestype	1 – 9	Household residence building type. (This could be optional as input, as it is not currently used in the DaySim model code.)
hhparcel	1 – 9999999	The ID of the parcel on which the household lives. Must be a parcel ID found in the raw parcel file.
hhhtaz	1 – 9999999	The ID of the zone in which the household lives. (This could be optional as input, as the Parcel file has the parcel-zone correspondence.)
hhexpfac	Real >= 0	The expansion factor for the household – a non-negative real number. (Is typically 1.0 in a synthetic population.)
samptype	0 - 99	The type of sample used. (This could be optional as input, as it is not used in the DaySim model code, but can be useful with survey data in model estimation to identify different sample types.)

Table 61 provides a list of fields available in the person file. Of the variables in the list, only seven (hhno, pno, pptyp, pagey, pgend, pwtyp, pstyp) are needed by DaySim as inputs in the raw data file. Four (pwpcl, pspcl, ptpass and ppaidprk) are predicted by DaySim. The rest are computed based on other data files or aren't currently used in model application and can be coded as -1 by the user. DaySim also adds one variable to the beginning of each record in the output person file ("ID" which is a sequential, unique person ID created by DaySim)

TABLE 61: PERSON FILE FORMAT

HEADER LABEL	VALID VALUES	DESCRIPTION, COMMENTS
hhno	1 - 9999999	The household ID number. Values must be unique positive integers, and should be in ascending order. Must be present in the Household file.
pno	1 – 99	The person sequence number within the household. Values must be unique positive integers within a household, and in ascending order from 1 up to "hhsize" on the Household file.
pptyp	1 - 8	Person type (1= full time worker, 2 =part time worker, 3=non-worker age 65+, 4 = other non-working adult, 5 = university student, 6 = grade school student/child age 16+, 7 = child age 5-15, 8 = child age 0-4. (There could be a switch to make this optional and compute it within DaySim for synthetic populations based on ACS PUMS. For other survey data, the coding and rules may be more variable and better done outside DaySim.)
pagey	0 – 99	Age in years (integer)
pgend	1 – 9	Gender (1=male, 2=female, 9=missing data for estimation)
pwtyp	0 - 2	Worker type (0=non-worker, 1=full time worker, 2=part time worker)
pwpcl	-1 - 9999999	Usual work location parcel ID. -1 for none/missing, otherwise must be a valid parcel ID present in the Parcel file.
pwtaz	-1 - 9999999	Usual work location zone ID. (This could be optional as input, as the Parcel file has the parcel-zone correspondence.)
pwautime	-1 - 9999999	The 1-way peak auto travel time between the residence and usual work parcels (a real number of minutes, -1 if no usual work location. Could be optional as input, used as output for calibration.)
pwaudist	-1 - 9999999	The 1-way peak auto travel distance between the residence and usual work parcels (a real number of miles, -1 if no usual work location. Could be optional as input, used as output for calibration.)
pstyp	0 - 2	Worker type (0=non-student, 1=full time student, 2=part time student if known – part-time distinction not used in DaySim code)
pspcl	-1 - 9999999	Usual school location parcel ID. -1 for none/missing, otherwise must be a valid parcel ID present in the Parcel file.

HEADER LABEL	VALID VALUES	DESCRIPTION, COMMENTS
pstaz	-1 - 9999999	Usual school location zone ID. (This could be optional as input, as the Parcel file has the parcel-zone correspondence.)
psautime	-1 - 9999999	The 1-way peak auto travel time between the residence and usual school parcels (a real number of minutes, -1 if no usual school location. Could be optional as input, used as output for calibration.)
psaudist	-1 - 9999999	The 1-way peak auto travel distance between the residence and usual school parcels (a real number of miles, -1 if no usual school location. Could be optional as input, used as output for calibration.)
puwmode	-1 – 9	The usual mode used to work. (This is optional, as it is a placeholder for possible models that may be added to DaySim in the future.)
puwarrp	-1 – 9	The usual arrival period at work. (This is optional, as it is a placeholder for possible models that may be added to DaySim in the future.)
puwdepp	-1 – 9	The usual departure period from work. (This is optional, as a placeholder for possible models that may be added to DaySim in the future.)
ptpass	0 – 1	Transit pass ownership (0=no, 1=yes. This is predicted by DaySim, so could be an optional input in application mode.)
ppaidprk	0 – 1	Worker has to pay to park at work (0=no, 1=yes. This is predicted by DaySim, so could be an optional input in application mode.)
pdiary	0 – 1	Survey respondent used their diary? (0=no, 1=yes. This is only relevant for survey data in estimation, so could be optional in application mode.)
pproxy	0 – 1	Survey responses by proxy? (1=no, 2=yes, 3=by mail, 9=missing. This is only relevant for survey data in estimation, so could be optional in application mode.)
psexpfac	Real >= 0	The expansion factor for the person – a non-negative real number. (In application mode, this could be optional, since it is set equal to hhexpfac)

Zone Index File

This is often referred to as the "taz index" file. Its main purpose is to indicate to DaySim which zone numbers are valid. If there are gaps in the zone numbering (unused zone numbers), then this file is used to set up a mapping from the nominal zone numbers to an internal zone numbering that is used to compress the amount of memory used to store zone-to-zone skim matrices in memory. The file is ASCII delimited with a header record and located at "1_Inputs\8_DaySim\01_TAZ_Index\Fresno_taz_indexes.dat". Table 62 presents format of the file.

TABLE 62: ZONE INDEX FILE FORMAT

HEADER LABEL	VALID VALUES	DESCRIPTION, COMMENTS
Zone_ID	1 – 9999999	The zone ID number used in the network software that produces skims. Values must be unique positive integers, in ascending order.
Zone_ordinal	1 – 9999999	A zone index number internal to DaySim, which is mapped to Zone_id. Values must be unique positive integers in ascending order. Value will generally begin at 1 with no gaps in numbering, although gaps are allowed.
Dest_eligible	0 or 1	A binary variable indicating whether or not a zone is eligible as a destination in Daysim. Zones that are not eligible as destinations are external zones or special park and ride lot zones.
External	0 – 99	This variable was originally used as a binary variable to indicate external zones, but was not used in the DaySim code, so it is now used to indicate a District mapping of the zones. Including a district mapping is optional – only necessary if ODShadowPricing is used, or other region-specific variables that are District-based
xcoord	1 – 9999999	**OPTIONAL** –only needed if Transit Stop Areas are used for transit skims. The x coordinate of the zone centroid, in integer length units (typically SPF).
ycoord	1 – 9999999	**OPTIONAL** –only needed if Transit Stop Areas are used for transit skims. The x coordinate of the zone centroid, in integer length units (typically SPF).

Workers IXXI Fractions

This is a file that DaySim uses for the work location model, to set the percent of workers living in each zone that work outside the region, and the percent of jobs in each zone that are filled by workers living outside the region. DaySim does not select a usual workplace or simulate internal home-work tours for the I-X fraction of workers, and makes the X-I fraction of jobs unavailable as usual work locations for region residents. Both fractions tend to be larger towards the edges of modeled regions.

The file is ASCII delimited without a header record and located at "1_Inputs\8_DaySim\05_ixxi\Fresno_worker_ixxifractions.dat". Table 63 presents format of the file.

TABLE 63: WORKERS IXXI FRACTIONS FILE FORMAT

HEADER LABEL	VALID VALUES	DESCRIPTION, COMMENTS
Zone_ID	1 - 9999999	The zone ID number. Values must be unique positive integers, in ascending order. There must be the same number of records in the same order as in the raw Zone file.
Worker_IXFrac	0.00 - 1.00	The fraction of workers living in the zone that have a usual work location outside the modeled region (not in a destination-eligible zone)
Jobs_XIFrac	0.00 - 1.00	The fraction of jobs in the zone that are filled by workers living outside of the region (not in a destination-eligible zone)

Park and Ride File

This file is for park and ride lot/path choice in DaySim. It is **optional**, and not needed if either (a) the park and ride mode is not included in the model for the region, or (b) the park and ride skim matrices are prepared in the network software rather than using the path choice in DaySim.

The file is ASCII delimited with a header record and located at "1_Inputs\8_DaySim\10_ParkAndRide \p_r_Nodes_2014.dat". Table 64 presents format of the file.

TABLE 64: PARK AND RIDE FILE FORMAT

HEADER LABEL	VALID VALUES	DESCRIPTION, COMMENTS
Node_ID	1 - 9999999	The park and ride node ID number. Values must be unique positive integers, in ascending order.
Zone_ID	1 - 9999999	The zone that the lot is associated in. Must be a zone ID present in the raw Zone file. May be either an internal (destination-eligible zone) or a special park and ride zone, which allows more accurate zone-to-zone skims for park and ride
xcoord	1 – 9999999	The x coordinate of the lot, in integer length units (typically SPF).

ycoord	1 – 9999999	The y coordinate of the lot, in integer length units (typically SPF).
capacity	0 – 9999999	The number of parking spaces in the lot. A value of 0 makes the lot unavailable as a choice option, but can be useful for including a lot as a placeholder for future/alternative scenarios
cost	0 – 9999999	The daily parking cost for the lot, in hundredths of monetary units (typically this is cents)

Node-to-Node Distance File

The node-to-node distance file is an input to the DaySim buffer tool. The ABM Cube process automatically generates maz-to-maz walk skim in a text format as described in Table 65. Other input, INPUT_NODE.csv, required for the buffer tool is also generated by the Cube process. With these two distance related inputs, the buffer tool generates a binary version of the distance file which is then used by DaySim for short trip distance calculations. In addition, the buffer tool generates a maz to node id correspondence, maz_node_{year}.dat, and an index file, {node_distance_file_index.txt} for DaySim to use node to node distance data.

TABLE 65: NODE TO NODE DISTANCE FILE FORMAT

FIELD	DESCRIPTION
RECORD_ID	Record id number
FROM_NODE_ID	Origin node id
TO_NODE_ID	Destination node id
DISTANCE	Shortest path distance in miles

Intersection Data File (Buffer Tool)

This file is an input to the buffer tool that generates buffered micro-zone file for DaySim. A unique measure of urban form that DaySim incorporates is the number of intersections or nodes of different types around a micro-zone. These intersection types include, dead-ends (1 link), T-intersections (3-links), and tradition intersections (4+ links), and help characterize the pattern of urban development.

The intersection data is in a CSV file and resides at "1_Inputs\8_DaySim\02_Parcel\intersection_2014_nohwys.csv". Table 66 presents contents of the file.

TABLE 66: INTERSECTION DATA FILE FORMAT

FIELD	DESCRIPTION
Id	Intersection ID number

Links	Number of links associated with node
xcoord_p	X coordinate – state plane feet
ycoord_p	Y coordinate – state plane feet

A largely automated process has been developed to calculate these urban form measures based on detailed GIS street centerline files. This is more detailed than the modeled network, which does not include all streets. This GIS process analyses the GIS street centerline file to locate nodes and assigns an intersection type code to them based on the number of links joined to the node.

Appendix C provides details of the procedure to create this input file.

Transit Stops File (Buffer Tool)

This file is an input to the buffer tool that generates buffered micro-zone file for DaySim. In addition to using zone-level information on access times to transit, DaySim also incorporates detailed micro-zone-level information on the distance to transit by transit sub-mode. This file is created from transit network by extracting stops on transit routes.

The file is located at "1_Inputs\8_DaySim\02_Parcel\stops_transit_2014.csv". Table 67 summarizes the contents of this file.

TABLE 67: TRANSIT STOPS FILE FORMAT

FIELD	DESCRIPTION
Id	Transit stop ID number
Mode	Transit sub-mode code 1=local bus 2=express bus 3=commuter rail 4=ferry (BRT) 5=light rail
xcoord_p	X coordinate – state plane feet
ycoord_p	Y coordinate – state plane feet

When developing or updating forecast year or project alternative networks, careful consideration should be given to the location of individual bus stops. In addition to the bus tops located in urban areas of the county, it is also necessary to incorporate bus stop locations for rural transit routes into the model. This fine-grained information is used by DaySim to develop micro-zone-level estimates of access time to transit. Ideally, forecast year transit networks would include a similar level of detail. Forecast year travel model transit network do include information on stop locations as part of the network coding. However, these stop locations are constrained by the coarser travel model roadway networks, and thus may tend to make transit access times appear longer by not including stops that are on major roads included in the roadway network. Model users should ensure that the future year transit stop location file used as input to the micro-zone preparation contains information consistent with expected future year alignments and stop spacing assumptions.

Appendix C provides details of the procedure to create this input file.

Parks/Open Spaces Data Stops File

This file is an input to the buffer tool that generates buffered micro-zone file for DaySim. A unique feature of DaySim is that it incorporates measures of access to publicly accessible open space. Although open space is clearly an attractor of travel for recreational, social and other purposes, typically open space is not included in travel models because the traditional "size" measures used as input to travel models, such as employment and population, are not good indicators of the attractiveness of open space (i.e. a popular park will often have no employment and no population). The open space measures incorporated into DaySim capture the proximity of each micro-zone to the nearest open space, and the amount of open space present in the buffer area around the micro-zone.

The open space park file resides at "1_Inputs\8_DaySim\02_Parcel\openspace_active_fresno.csv". Contents of the file are presented in Table 68.

TABLE 68: OPEN SPACE DATA FILE FORMAT

FIELD	DESCRIPTION
Id	Open space ID number
xcoord_p	X coordinate – state plane feet
ycoord_p	Y coordinate – state plane feet
Sqft	Open space grid cell size in sq ft

The individual records in the open space file are based on converting a shapefile of regional, publicly accessibility open spaces into a smaller set of open space grid cells.

Appendix C provides details of the procedure to create this input file.

Coefficient Files

Coefficient for each model is a separate text file (.F12 format) that can be edited by the user for calibration purpose. There are a total of 23 files corresponding to all the models described in section 5.9. All coefficient files are in the same directory at 1_Inputs\8_DaySim\07_Coefficients. For example, the person day pattern model coefficient is a file named IndividualPersonDayPatternCoefficients_Nash-v1.8.F12. Figure 47 provides an example of a coefficient file.

FIGURE 47: COEFFICIENT FILE EXAMPLE

```
Auto availability auto27.ALO
Created by ALOGIT version 4                              11:16:44 on  3 Apr 12
END
   1 Beta00001  F   -3.74713956271        .485432341916
   2 Beta00002  F   -2.12971158522        .120667788954
   3 Beta00003  F   -3.96633437255        .181623902077
   4 Beta00004  F   -11.21939263133       .347582961233
   5 Beta00005  F   -5.01252239349        .485933923292
   6 Beta00006  F   -1.85637454651        .171204853891
   7 Beta00007  F   -1.39642394328        .108778102305
   8 Beta00008  F   -2.39429368035        .162199862000
   9 Beta00009  F   -4.92018331565        .687312315918
  10 Beta00010  F   -2.29401614977        .261316908165
  11 Beta00011  F   -1.44463523641        .204749856644
  12 Beta00012  F   -0.81535992724        .203518529877
  13 Beta00013  F   -5.32165437066        1.22007612531
  14 Beta00014  F   -1.12612768117        .393371210208
  15 Beta00015  F   -1.63721987659        .302144582330
  16 Beta00016  F   -0.59991414111        .302706275384
  18 Beta00018  F   0.39953921787         .124608088395
  19 Beta00019  F    .699030446312       .402977980122
  20 Beta00020  F    .324854469598       .204037451392
  22 Beta00022  F   -.438561287720       .318213047774
  24 Beta00024  F    .239019734309       .118684885778
  25 Beta00025  F   -.301156533677       .157265779465
  26 Beta00026  F   -.563735159108       .253286330074
  28 Beta00028  F    .319598840569       .228619271418
  29 Beta00029  F    .570686967046       .282854170991
```

Roster File

The Roster file is a very convenient and flexible way to define which skim matrices are used for all modes/path types and level-of-service (LOS) variables, for all times of day. The roster file is here: “1_Inputs\8_DaySim\06_Roster\roster_mz.csv”.

The Roster file must also be in CSV format and must include the columns listed below and shown in the example below.

- **#variable**: A variable label, as referred to in DaySim code
- **mode**: A mode label, as referred to in the DaySim code and present in the RosterCombinations file
- **path-type**: A path type label, as referred to in the DaySim code, and the mode/path type combination must be TRUE in the RosterCombinations file In the example, the “walk”, “bike”, “sov”, “hov2” and “hov3” are input only for the “full-network” path type, while separate “transit” skims are input for five different path types.

- **vot-group**: A value of time class, with boundaries set in the configuration file. Valid names are very-low, low, medium, high, very-high, or all.
- **start-minute** and **end-minute:** The time period for which the skim matrix applies, in minutes past midnight. For example, 0-1439 is the entire day, and 360-539 is 6:00 am to 8:59 am. For each variable/mode/path-type/vot-group combination, the skims should cover all minutes from 0 to 1439. An example is the last 5 lines above for toll, for 5 time periods. The last period 1110-179 spans midnight, and is 6:30 pm to 2:59 am.
- **length:** The "zone" system used for the matrix. "maxzone" uses Zone, "transitstop" uses TransitStopArea, "null" just returns a value of 0 instead of reading in a matrix (in which case the "file-type" and "name" columns should also be "null").
- **file-type**: The format of the file. Valid types are:
    - **text-ij:** Text files with a record for each I-J zone pair that can contain more than 1 skim variable. Column 1 of the text file is always the I zone and column 2 the J zone.
    - **transcad**: Native binary format written by TransCAD-requires a valid TransCAD license to be installed when running.
    - **cube**: Native binary format written by Cube. Requires a valid Cube license to be installed when running.
    - **emme**: Native binary format written to EMME databanks.
    - **visum-binary**: Native binary format written by Visum.
    - **hdf5**: HDF5 format, which can be written by various network packages or converters.
    - **omx**: Open Matrix format (modified HDF5), which can be written by various network packages or converters.
    - **bin:** A custom DaySim binary format, which is fastest to load, but requires pre-processing of the matrices.
    - **null:** Returns a value of 0. ("length" and "name" should also be set to "null")
- **name**: The matrix file name (plus the table name for HDF5 or OMX files). The directory path is assumed to be the same as for the Roster file. Note that different roster rows may refer to the same file—such as the first four rows in the example using "walkSkim.h5/1". In that case, the same matrix is used for multiple variables, but only read and stored in memory once. In the example, the toll matrices use the same table number, but different file names for the five time periods.

- **field**: The matrix number on the file for the particular variable. For "text-ij", it is the column number. For "cube" it is the Cube matrix number.
- **transpose**: This indicates that the origin and destination zones for the matrix should be "switched" and the transpose used. For example, the transpose of the AM peak transit matrices can be used to represent the PM peak period.
- **blend-variable**: This is the variable to use for "short distance blending". It is only relevant for the walk, bike and auto modes, but not for transit. In practice, this should always be set as "distance", which is what the DaySim code assumes.
- **blend-path-type**: This is the path type to use for "short distance blending". If it is set to "null", then DaySim assumes that the "blend-path-type" is the same as the "path-type" entry on the same row. It would be possible to use a different path type for the same mode if no "distance" matrix was available for "path type", but in almost all cases, the user should leave this as "null"
- **factor**: This allows one variable to be set as a factored version of another variable. In the example below, "walk" "time" is set to use the same matrix as "walk" "distance", but with a factor of 20, which assumes a walk speed of 20 minutes per mile. Similarly, "bike" "time" is set to use the same matrix as "bike" "distance", but with a factor of 6, which assumes a bike speed of 6 minutes per mile. If the entry is "null", the default factor is 1.0.
- **scaling**: This last setting causes some confusion, so deserves careful explanation. DaySim stores all matrix values in memory as 2-byte unsigned integers, which can take values between 0 and roughly 65,000. Those values are assumed to be hundredths of miles for distance, hundredths of minutes for times, hundredths of dollars (cents) for costs, and hundredths of boardings for transit boardings or transfers. So, the maximum skim values that can be stored in memory are roughly 650 miles, minutes, dollars, or boardings, which is sufficient for regional models. (Any values larger than this are capped at the maximum value.)

In most cases, the input matrices are in units of miles, minutes, or dollars. In that case, "scaling" should be set to TRUE, and DaySim will scale the values by 100 when storing them in memory, and then "unscale" them back to the original units when accessing them from memory. That is done for most of the walk, bike, and auto variables in the example, as well as transit times and boardings ("nboard").

In the example, the toll and fare matrices were already in units of cents rather than dollars, so no scaling is necessary, and "scaling" is FALSE for those variables. However, the DaySim code assumes that these costs are in dollars, so a factor of 0.1 is necessary to convert the cost in cents to a cost in dollars. In general, the rules are:

In summary, if the matrices are in units of miles, minutes or dollars, the proper setting of "scaling" is "TRUE" and "factor" is "null" (unless converting walk distance to walk time, or a similar conversion).

If the matrices are in cents, or they are already pre-scaled to hundredths of miles or hundredths of minutes, the proper setting of “scaling” is “FALSE” and “factor” is “0.01”.

Roster Combinations File

The RosterCombinations file is used together with the Roster file and tells which mode-path type combinations are valid in the Roster file. A “path type” is essentially a “submode” in DaySim, such as tolled versus non-tolled networks for auto modes or local bus versus light rail (plus optional bus) networks for transit. The roster combinations file is here: “1_Inputs\8_DaySim\06_Roster\roster_combinations.csv”.

The RosterCombinations file must be in .CSV format and is a matrix where the columns give the vaild mode labels for the Roster entries, and the rows give the valid “path type” labels for the Roster entries.

The format is presented in Table 69. A TRUE entry means that the mode/path type combination is valid for the matrix entries in the Roster, and a FALSE entry means that it is not valid. A TRUE entry does not mean that the path type is required, however, so using the file below would not require separate “no-tolls” matrices for the auto modes, or separate “ferry” matrices for transit – but it would allow them.

A few other things to note in this example…

There is a “park-and-ride” mode listed, but that does not require that the user provide skims for park and ride. (That is an option, but DaySim can also use its internal lot/path choice to create park and ride paths from the “sov” and “transit” skims.)

All of the path types are FALSE for the “school-bus” mode. Typically, the school bus mode uses the “hov3” skims in the models, so no separate “school-bus” skims are listed in the Roster.

The “other” mode is basically a placeholder, not currently used in the models. For BKR, we may add a “taxi-uber” mode, but that will use “hov2” skims, so all of the path types could remain FALSE.

Some DaySim users let the network software choose the best transit path, so only use a single path type. In that case, they often list the path type for all of the “transit” skims as “local-bus”, but it would also be possible to list them all as “full-network”, and change the RosterCombinations file so that only the “full-network” row is TRUE under transit, and all of the other rows are FALSE.

A user can add or substituted different path type names, and even different mode names, but that could require substantial corresponding changes to the DaySim code—essentially every call to PathTypeModel includes a mode label and path type label, and that combination must be TRUE in the RosterCombinations file.

TABLE 69: ROSTER COMBINATIONS FILE FORMAT

#	WALK	BIKE	SOV	HOV2	HOV3	TRANSIT	PARK-AND-RIDE	SCHOOL-BUS	OTHER

full-network	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE
no-tolls	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE
local-bus	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE
light-rail	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE
premium-bus	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE
commuter-rail	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE
ferry	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE

DaySim Configuration File

The configuration file is the main user input control file for DaySim. This file is created automatically before a DaySim run. As mentioned previously, for each feedback loop DaySim is run for four iterations. The first three iterations of DaySim run only long-term choice models (work and school location choice) to get stable shadow prices. The last iteration runs all models (long-term and short-term choice models). In all, each feedback loop generates two sets of properties files under the DaySim folder in a scenario directory: configuration.properties and configuration_shadow_price.properties.

Appendix F describes various settings available to configure a DaySim model run.

5.6 OUTPUTS

DaySim

The outputs generated from a DaySim run are stored inside a scenario folder here: “Scenarios\FC14_BASE\11_DaySim”. This section describes all necessary DaySim outputs, including the intermediate outputs generated by the buffer tool which are then used in DaySim.

Buffered Micro-zone File

The buffered micro-zone file is an output from the buffer tool and used in a DaySim run. This is a space-delimited delimited ASCII text format file (.dat) with one row of data per micro-zone and is the primary file used to maintain socioeconomic information. The file begins with several fields identifying the micro-zone and describing the physical location and size of the micro-zone. Then contains fields that describe the quantity of housing, school enrollment, and employment around the micro-zone using logistic distance decay curves with 1/8th mile and quarter mile inflection points. These two distance decay curves with 1/8th and quarter mile inflection points result in “buffer 1” and “buffer 2” variables respectively which are referred to in the file format table below. In addition, the micro-zone file contains information about urban form and the transportation system on and close to the micro-zone, including the proximity to transit stops and the price and supply of parking.

Table 70 presents a list of fields available in the buffered micro-zone file.

TABLE 70: BUFFERED MICRO-ZONE FILE

FIELD	DESCRIPTION
Id	Micro-zone ID number
xcoord_p	X coordinate – state plane feet
ycoord_p	Y coordinate – state plane feet
sqft_p	Area – square feet
taz_p	TAZ number
lutype_p	Park area (square feet)
hh_p	households on micro-zone
stugrd_p	grade school enrollment on micro-zone
stuhgh_p	high school enrollment on micro-zone
stuuni_p	university enrollment on micro-zone
empedu_p	educational employment on micro-zone
empfoo_p	food employment on micro-zone
empgov_p	government employment on micro-zone
empind_p	industrial employment on micro-zone
empmed_p	medical employment on micro-zone
empofc_p	office employment on micro-zone
empret_p	retail employment on micro-zone
empsvc_p	service employment on micro-zone
empoth_p	other employment on micro-zone
emptot_p	total employment on micro-zone
parkdy_p	offstreet daily parking on micro-zone
parkhr_p	offstreet hourly parking on micro-zone
Ppricdyp	offstreet daily parking price
Pprichrp	offstreet hourly parking price
hh_1	households within buffer 1

FIELD	DESCRIPTION
stugrd_1	grade school enrollment within buffer 1
stuhgh_1	high school enrollment within buffer 1
stuuni_1	university enrollment within buffer 1
empedu_1	educational employment within buffer 1
empfoo_1	food employment within buffer 1
empgov_1	government employment within buffer 1
empind_1	industrial employment within buffer 1
empmed_1	medical employment within buffer 1
empofc_1	office employment within buffer 1
empret_1	retail employment within buffer 1
empsvc_1	service employment within buffer 1
empoth_1	other employment within buffer 1
emptot_1	total employment within buffer 1
parkdy_1	offstreet daily parking within buffer 1
parkhr_1	offstreet hourly parking within buffer 1
ppricdy1	average offstreet daily parking price within buffer 1
pprichr1	average offstreet hourly parking price within buffer 1
nodes1_1	number of single link street nodes (dead ends) within buffer 1
nodes3_1	number of three-link street nodes (T-intersections) within buffer 1
nodes4_1	number of 4+ link street nodes (traditional 4-way +) within buffer 1
tstops_1	number of transit stops within buffer 1
nparks_1	number of open space parks within buffer 1
aparks_1	open space area in swuare feet within buffer 1
hh_2	households within buffer 2
stugrd_2	grade school enrollment within buffer 2
stuhgh_2	high school enrollment within buffer 2

FIELD	DESCRIPTION
stuuni_2	university enrollment within buffer 2
empedu_2	educational employment within buffer 2
empfoo_2	food employment within buffer 2
empgov_2	government employment within buffer 2
empind_2	industrial employment within buffer 2
empmed_2	medical employment within buffer 2
empofc_2	office employment within buffer 2
empret_2	retail employment within buffer 2
empsvc_2	service employment within buffer 2
empoth_2	other employment within buffer 2
emptot_2	total employment within buffer 2
parkdy_2	offstreet daily parking within buffer 2
parkhr_2	offstreet hourly parking within buffer 2
ppricdy2	average offstreet daily parking price within buffer 2
pprichr2	average offstreet hourly parking price within buffer 2
nodes1_2	number of single link street nodes (dead ends) within buffer 2
nodes3_2	number of three-link street nodes (T-intersections) within buffer 2
nodes4_2	number of 4+ link street nodes (traditional 4-way +) within buffer 2
tstops_2	number of transit stops within buffer 2
nparks_2	number of open space parks within buffer 2
aparks_2	open space area in swuare feet within buffer 2
dist_lbus	distance to nearest local bus stop from micro-zone
dist_ebus	distance to nearest express bus stop from micro-zone
dist_crt	distance to nearest commuter rail stop from micro-zone
dist_fry	distance to nearest ferry stop from micro-zone
dist_lrt	distance to nearest light rail stop from micro-zone

FIELD	DESCRIPTION
dist_park	distance to nearest park from micro-zone

Tour File (_tour.tsv)

This file has tour-level variables for all persons predicted by a DaySim run. The file is ASCII delimited with a record and is stored here: "Scenarios\FC14_BASE\11_DaySim_tour.tsv"

TABLE 71: TOUR FILE FORMAT

HEADER LABEL	VALID VALUES	DESCRIPTION, COMMENTS
id		Internal DaySim record ID
person_id		Internal DaySim record ID
person_day_id		Internal DaySim record ID
hhno	1 - 9999999	The household ID number
pno	1 – 99	The person sequence number within the household
day	1 – 99	The survey day sequence, if using multi-day survey data, or if DaySim were programmed to simulate multiple days per household (which is not a current feature).
tour	1 – 99	The tour sequence within the person-day
jtindex	0 – 99	Links to the sequence number of the tour in the JointTour file records for the HouseholdDay. (Only relevant for the H version of the models.)
parent	0 - 99	If it is a work-based subtour, the "tour" sequence number of the "parent" work tour, otherwise 0.
subtours	0 - 99	For home-based work tours, the number of work-based subtours made from the work activity of that tour.
pdpurp	1 - 9	The tour primary destination purpose (1=work, 2=school, 3=escort, 4=personal business (& medical), 5=shopping, 6=meal, 7=social (& recreation), 8=recreation (H version only) 9=medical (H version only)
tlvorig	0 - 1439	The time leaving the (sub)tour origin, in minutes after midnight (or hours*100+minute for estimation mode)

tardest	0 - 1439	The time arriving at the (sub)tour destination, in minutes after midnight (or hours*100+minute for estimation mode)
tlvdest	0 - 1439	The time leaving the (sub)tour destination, in minutes after midnight (or hours*100+minute for estimation mode)
tarorig	0 - 1439	The time arriving back at the (sub)tour origin, in minutes after midnight (or hours*100+minute for estimation mode)
toadtyp	1 - 5	Tour origin address type (1=home, 2=usual work location, 3=usual school location, 4=other location in region, 5=out of region/missing (survey data only)
tdadtyp	1 - 5	Tour destination address type (1=home, 2=usual work location, 3=usual school location, 4=other location in region, 5=out of region/missing (survey data only)
topcl	-1 - 9999999	Tour origin parcel ID. Must be a valid parcel ID present in the Parcel file.
totaz	-1 - 9999999	Tour origin zone ID. Must be a valid zone ID present in the Zone file.
tdpcl	-1 - 9999999	Tour destination parcel ID. Must be a valid parcel ID present in the Parcel file.
tdtaz	-1 - 9999999	Tour destination zone ID. Must be a valid zone ID present in the Zone file.
tmodetp	1 - 8	Tour main mode type (1=walk, 2=bike, 3=sov, 4=hov 2, 5=hov 3+, 6=walk to transit, 7=park and ride, 8=school bus, 9=TNC, 10=other – survey only)
tpathtp	1 - 8	Tour main mode path type (1=full network, 2=no-toll network, 3=local bus, 4=light rail, 5=premium bus, 6=commuter rail, 7=ferry)
tautotime	-1 - 9999999	The one-way auto travel time between the origin and destination (a real number of minutes)
tautocost	-1 - 9999999	The one-way auto toll cost between the origin and destination (a real number of dollars)
tautodist	-1 - 9999999	The one-way auto travel distance between the origin and destination (a real number of miles)
tripsh1	1 - 99	The number of trips segments on the half tour to the destination.

tripsh2	1 - 99	The number of trips segments on the half tour from the destination.
phtindx1	0 - 99	Links to the sequence number of the first half tour in the PartialHalfTour file records for the HouseholdDay. (Only relevant for the H version of the models.)
phtindx2	0 - 99	Links to the sequence number of the second half tour in the PartialHalfTour file records for the HouseholdDay. (Only relevant for the H version of the models.)
fhtindx1	0 - 99	Links to the sequence number of the first half tour in the FullHalfTour file records for the HouseholdDay. (Only relevant for the H version of the models.)
fhtindx2	0 - 99	Links to the sequence number of the second half tour in the FullHalfTour file records for the HouseholdDay. (Only relevant for the H version of the models.)
toexpfac	Real >= 0	The expansion factor for the tour – a non-negative real number. (Is set equal to hhexpfac in application mode)

Trip File

This file is an output from a DaySim run and has trip-level variables for all persons predicted by a DaySim run. The file is ASCII delimited with a record and is stored here: "Scenarios\FC14_BASE\11_DaySim_trip.tsv"

TABLE 72: TRIP FILE FORMAT

HEADER LABEL	VALID VALUES	DESCRIPTION, COMMENTS
Id		Internal DaySim record ID
Tour_id		Internal DaySim record ID
hhno	1 - 9999999	The household ID number. Values must be unique positive integers, and should be in ascending order. Must be present in the Household file.
pno	1 – 99	The person sequence number within the household. Values must be unique positive integers within a household, and match the hhno/pno combinations in the Person file

day	1 – 99	The survey day sequence, if using multi-day survey data, or if DaySim were programmed to simulate multiple days per household (which is not a current feature).
tour	1 – 99	The tour sequence within the person-day. Must match a tour present for the person-day in the Tour file.
half	1 –2	The half tour (1=to the destination, 2=from the destination)
tseg	1 - 99	The trip sequence number within the half tour.
tsvid	1 - 99	Links to a travel survey trip ID (not relevant in application mode)
opurp	0 – 10	The purpose at the trip origin (0=home, 1=work, 2=school, 3=escort, 4=personal business (& medical), 5=shopping, 6=meal, 7=social (& recreation), 8=recreation (H version only) 9=medical (H version only), 10=change mode at a park and ride lot
dpurp	0 – 10	The purpose at the trip destination (0=home, 1=work, 2=school, 3=escort, 4=personal business (& medical), 5=shopping, 6=meal, 7=social (& recreation), 8=recreation (H version only) 9=medical (H version only), 10=change mode at a park and ride lot
oadtyp	1 – 6	Trip origin address type (1=home, 2=usual work location, 3=usual school location, 4=other location in region, 5=out of region/missing (survey data only), 6=inserted change mode location for park and ride
dadtyp	1 – 6	Trip dest. address type (1=home, 2=usual work location, 3=usual school location, 4=other location in region, 5=out of region/missing (survey data only), 6=inserted change mode location for park and ride
opcl	-1 - 9999999	Trip origin parcel ID. Must be a valid parcel ID present in the Parcel file.
otaz	-1 - 9999999	Trip origin zone ID. Must be a valid zone ID present in the Zone file.
dpcl	-1 - 9999999	Trip destination parcel ID. Must be a valid parcel ID present in the Parcel file.
dtaz	-1 - 9999999	Trip destination zone ID. Must be a valid zone ID present in the Zone file.
mode	1 - 8	Trip mode (1=walk, 2=bike, 3=sov, 4=hov 2, 5=hov 3+, 6=walk to transit, 7=park and ride, 8=school bus, 9=TNC, 10=other – survey only)
pathtype	1 - 8	Trip path type (1=full network, 2=no-toll network, 3=local bus, 4=light rail, 5=premium bus, 6=commuter rail, 7=ferry)

dorp	0 - 999	For auto trips, 1=driver, 2=passenger; for transit trips, is set to the total walk access+egress time, in integer minutes
deptm	0 – 1439	The trip departure time, in minutes after midnight (or hours*100+minute for estimation mode)
arrtm	0 – 1439	The trip arrival time, in minutes after midnight (or hours*100+minute for estimation mode)
endacttm	0 – 1439	The end time of the destination activity, in minutes after midnight (or hours*100+minute for estimation mode)
travtime	-1 - 9999999	The travel time by the trip mode and path type (a real number of minutes)
travcost	-1 - 9999999	The travel cost by the trip mode and path type (a real number of dollars)
travdist	-1 - 9999999	The network distance between the trip origin and destination (a real number of miles, SOV distance used for transit trips)
trexpfac	Real >= 0	The expansion factor for the trip – a non-negative real number. (Is set equal to hhexpfac in application mode)

Household and Household Day Files

These files are outputs from a DaySim run. The household and household day output files append the model predicted information into the household input files. Household output file is in the same format as the input file in Table 60. The household file is here: "Scenarios\FC14_BASE\11_DaySim_household.tsv"

The household day output file is here: "Scenarios\FC14_BASE\11_DaySim_household_day.tsv". Table 73 presents the format of the file.

TABLE 73: HOUSEHOLD DAY FILE FORMAT

HEADER LABEL	VALID VALUES	DESCRIPTION, COMMENTS
id		Internal DaySim record ID
hhno	1 - 9999999	The household ID number. Values must be unique positive integers, and should be in ascending order. Must be present in the Household file.

day	1 – 99	The survey day sequence, if using multi-day survey data, or if DaySim were programmed to simulate multiple days per household (which is not a current feature).
dow	1 - 7	The day of the week, which is relevant for survey data, but is not currently used in the DaySim models.
jttours	0 - 99	The number of fully joint tour records output for the household. (Is set only in the H version of the DaySim models.)
phtours	0 - 99	The number of partially joint half tour records output for the household. (Is set only in the H version of the DaySim models.)
fhtours	0 - 99	The number of fully joint half tour records output for the household. (Is set only in the H version of the DaySim models.)
hdexpfac	Real >= 0	The expansion factor for the household-day – a non-negative real number. (Is set equal to hhexpfac in application mode)

Person and Person Day Files

The person output file is in the same format as the input person file in Table 61. The person file is here: "Scenarios\FC14_BASE\11_DaySim_person.tsv"

The person-day output file has the person-day-level variables from a DaySim run. The file is here: "Scenarios\FC14_BASE\11_DaySim_person_day.tsv". The format of the file is shown in Table 74.

TABLE 74: PERSON DAY FILE FORMAT

HEADER LABEL	VALID VALUES	DESCRIPTION, COMMENTS
id		Internal DaySim record ID
person_id		Internal DaySim record ID
household_day_id		Internal DaySim record ID
hhno	1 - 9999999	The household ID number. Values must be unique positive integers, and should be in ascending order. Must be present in the Household file.
pno	1 – 99	The person sequence number within the household. Values must be unique positive integers within a household, and match the hhno/pno combinations in the Person file

day	1 – 99	The survey day sequence, if using multi-day survey data, or if DaySim were programmed to simulate multiple days per household (which is not a current feature).
beghom	0 – 1	A flag if the survey diary day begins at home. (Not currently relevant for application mode, where all days are simulated to begin at home.)
endhom	0 – 1	A flag if the survey diary day ends at home. (Not currently relevant for application mode, where all days are simulated to end at home.)
hbtours	0 – 99	The total number of home-based tour records predicted for the person-day.
wbtours	0 – 99	The total number of work-based subtour records predicted for the person-day.
uwtours	0 – 99	The total number of home-based work tours predicted to go to the usual workplace in the person-day
wktours	0 – 99	The number of home-based work tours predicted in the person-day
sctours	0 – 99	The number of home-based school tours predicted in the person-day
estours	0 – 99	The number of home-based escort tours predicted in the person-day
pbtours	0 – 99	The number of home-based personal business tours predicted in the person-day (also includes medical tours in the Default models)
shtours	0 – 99	The number of home-based shopping tours predicted in the person-day
mltours	0 – 99	The number of home-based meal tours predicted in the person-day
sotours	0 – 99	The number of home-based social tours predicted in the person-day (also includes recreational tours in the Default models)

retours	0 – 99	The number of home-based recreation tours predicted in the person-day. (Is only predicted by the H version of the models.)
metours	0 – 99	The number of home-based medical tours predicted in the person-day. . (Is only predicted by the H version of the models.)
wkstops	0 – 99	The number of home-based work stops predicted in the person-day
scstops	0 – 99	The number of home-based school stops predicted in the person-day
esstops	0 – 99	The number of home-based escort stops predicted in the person-day
pbstops	0 – 99	The number of home-based personal business stops predicted in the person-day (also includes medical stops in the Default models)
shstops	0 – 99	The number of home-based shopping stops predicted in the person-day
mlstops	0 – 99	The number of home-based meal stops predicted in the person-day
sostops	0 – 99	The number of home-based social stops predicted in the person-day (also includes recreational stops in the Default models)
restops	0 – 99	The number of home-based recreation stops predicted in the person-day. (Is only predicted by the H version of the models.)
mestops	0 – 99	The number of home-based medical stops predicted in the person-day. (Is only predicted by the H version of the models.)
wkathome	0 – 1439	The number of minutes spent working at home during the day. (Is only predicted by the H version of the models.)
pdexpfac	Real >= 0	The expansion factor for the household-day – a non-negative real number. (Is set equal to hhexpfac in application mode)

Network Skims

The model produces three sets of skims: highway, transit, and non-motorized. The skims are under “Scenarios\FC14_BASE\01_Skims”.

Highway

The highway skims are for two time periods (PK: peak and OK: off-peak) and three modes (D1: drive alone, S2: shared-ride 2, and S3: shared-ride 3+). In all, six highway skims are output as {Scenario_Name}_SKM_{tod}_{mode}.mat, with each containing attributes as shown in Table 75.

TABLE 75: HIGHWAY SKIM ATTRIBUTES

SKIM INDEX	ATTRIBUTE	DESCRIPTION
1	GENTIME_0Veh	Generalized time for zone pair for low value-of-time group
2	TIME_0Veh	Congested travel time (mins) for zone pair for low value-of-time group
3	DIST_0Veh	Travel distance (miles) for zone pair for low value-of-time group
4	COST_0Veh	Travel cost for zone pair for low value-of-time group
5	GENTIME_1Veh	Generalized time for zone pair for medium value-of-time group
6	TIME_1Veh	Congested travel time (mins) for zone pair for medium value-of-time group
7	DIST_1Veh	Travel distance (miles) for zone pair for medium value-of-time group
8	COST_1Veh	Travel cost for zone pair for medium value-of-time group
9	GENTIME_2Veh	Generalized time for zone pair for high value-of-time group
10	TIME_2Veh	Congested travel time (mins) for zone pair for high value-of-time group

11	DIST_2Veh	Travel distance (miles) for zone pair for high value-of-time group
12	COST_2Veh	Travel cost for zone pair for high value-of-time group

Transit

The transit skims are for two time periods (PK: peak and OK: off-peak), two sub-modes (B: bus and R: rail), and two access modes (W: walk and D: drive). In all eight transit skims are output as { scenario_name}_SKM_{tod}_T{access_mode}{mode}.mat, with each containing attributes as shown in Table 76Table 75.

TABLE 76: TRANSIT SKIM ATTRIBUTES

SKIM INDEX	SKIM ATTRIBUTE	DESCRIPTION
1	IVTT	In-vehicle travel time (mins)
2	DRV_P	
3	DRVDIST_P	
4	WLK_P	
5	WLK_A	Walk access (mins)
6	WLK_X	
7	IWAIT	Initial wait time (mins)
8	XWAIT	Transfer wait time (mins)
9	FARE	Fare (dollars)
10	BRDS	Number of boardings

Non-motorized

The non-motorized (bike and walk) skims are generated at micro-zone (MAZ) level. Though, walk skims are also produced at zonal (TAZ) level. The bike skim is {scenario_name}_MAZ_SKM_BIKE.mat and contains attributes as shown in Table 77.

TABLE 77: BIKE SKIM (MAZ) ATTRIBUTES

SKIM INDEX	SKIM ATTRIBUTE	DESCRIPTION
1	DIST_BIKE	Bike distance (miles) for OD pair

The MAZ and TAZ walk skims are WALK_SKIM_MAZ_MAZ_SORTED.TXT and {scenario_name}_TAZ_SKM_WALK.mat respectively. Available attributes are presented in Table 78 and Table 79.

TABLE 78: WALK SKIM (MAZ) ATTRIBUTES

FIELD	DESCRIPTION
Record_id	Record number
From_node_id	Origin node id
To_node_id	Destination node id
Distance	Distance (miles)

Note: node ids are nearest nodes to micro-zones. The file "maz_node_2014.dat" provides nearest node to a micro-zone. The file is in the following format: id (mazid), node_id (nearest node id).

TABLE 79: WALK SKIM (TAZ) ATTRIBUTES

SKIM INDEX	SKIM ATTRIBUTE	DESCRIPTION
1	TIME_WALK	Walk travel time (mins) for OD pair
2	DIST_WALK	Walk distance (miles) for OD pair

Assignment Results

Assignment results are output in "Scenarios\FC14_BASE\09_Assignment".

Highway

The highway assignment results are output in a CUBE network format "FC14_BASE_LOADEDNETWORK_DETAIL.NET"), as well as in a database file format

("FC14_BASE_LOADEDNETWORK_DETAIL.DBF"). Both outputs contain attributes as shown in Table 80.

TABLE 80: ATTRIBUTES IN HIGHWAY ASSIGNMENT RESULT

FIELD	DESCRIPTION
A	A node
B	B node
SHAPE_LENGT	Link length (feet)
DISTANCE	Distance (miles)
CAPCLASS	Capacity class
LANES	Number of lanes
NAME	Street name
ROUTE	Route
TERRAIN	Terrain type (F or R)
JURISDICTIO	Jurisdiction
SCREENLINE	Screen line id
SPEED	Speed (mph)
AREATYP	Area type (R-rural, U-urban, SU-suburban)
FACTYP	Facility type 0: local 1: Freeway 2: Highway 3: Expressway 4: Arterial 5: Collector 6: Local 7: Ramp-Freeway-Freeway

	8: Ramp-Slip 9: Ramp-Loop 10: Connector1 11: Connector2
AUX	Presence of auxiliary lane (1-yes; 0-no)
USE	Type of use
TOLL	Toll (cents?)
IMPROVED	Improvement id
A01_VOL	AB volume in AM peak hour
TOT_A01_VOL	Total (AB+BA) volume in AM peak hour
A03_VOL	AB volume in AM period
TOT_A03_VOL	Total (AB+BA) volume in AM period
M07_VOL	AB volume in MD period
TOT_M07_VOL	Total (AB+BA) volume in MD period
P01_VOL	AB volume in PM peak hour
TOT_P01_VOL	Total (AB+BA) volume in PM peak hour
P03_VOL	AB volume in PM period
TOT_P03_VOL	Total (AB+BA) volume in PM period
E11_VOL	AB volume in EV period
TOT_E11_VOL	Total (AB+BA) volume in EV period
D24_VOL	Daily AB volume
TOT_D24_VOL	Daily total (AB+BA) volume
A01_PAS_VOL	AB passenger car volume in AM peak hour

TOT_A01_PAS	Total (AB+BA) passenger car volume in AM peak hour
A03_PAS_VOL	AB passenger car volume in AM period
TOT_A03_PAS	Total (AB+BA) passenger car volume in AM period
M07_PAS_VOL	AB passenger car volume in MD period
TOT_M07_PAS	Total (AB+BA) passenger car volume in MD period
P01_PAS_VOL	AB passenger car volume in PM peak hour
TOT_P01_PAS	Total (AB+BA) passenger car volume in PM peak hour
P03_PAS_VOL	AB passenger car volume in PM period
TOT_P03_PAS	Total (AB+BA) passenger car volume in PM period
E11_PAS_VOL	AB passenger car volume in EV period
TOT_E11_PAS	Total (AB+BA) passenger car volume in EV period
D24_PAS_VOL	Daily AB passenger car volume
TOT_D24_PAS	Daily total (AB+BA) passenger car volume
A01_XX_VOL	AB external volume in AM peak hour
TOT_A01_XX_	Total (AB+BA) external volume in AM peak hour
A03_XX_VOL	AB external volume in AM period
TOT_A03_XX_	Total (AB+BA) external volume in AM period
M07_XX_VOL	AB external volume in MD period
TOT_M07_XX_	Total (AB+BA) external volume in MD period

P01_XX_VOL	AB external volume in PM peak hour
TOT_P01_XX_	Total (AB+BA) external volume in PM peak hour
P03_XX_VOL	AB external volume in PM period
TOT_P03_XX_	Total (AB+BA) external volume in PM period
E11_XX_VOL	AB external volume in EV period
TOT_E11_XX_	Total (AB+BA) external volume in EV period
D24_XX_VOL	Daily AB external volume
TOT_D24_XX_	Daily total (AB+BA) external volume
A01_TS_VOL	AB small truck volume in AM peak hour
TOT_A01_TS_	Total (AB+BA) small truck volume in AM peak hour
A03_TS_VOL	AB small truck volume in AM period
TOT_A03_TS_	Total (AB+BA) small truck volume in AM period
M07_TS_VOL	AB small truck volume in MD period
TOT_M07_TS_	Total (AB+BA) small truck volume in MD period
P01_TS_VOL	AB small truck volume in PM peak hour
TOT_P01_TS_	Total (AB+BA) small truck volume in PM peak hour
P03_TS_VOL	AB small truck volume in PM period
TOT_P03_TS_	Total (AB+BA) small truck volume in PM period
E11_TS_VOL	AB small truck volume in EV period

TOT_E11_TS_	Total (AB+BA) small truck volume in EV period
D24_TS_VOL	Daily AB small truck volume
TOT_D24_TS_	Daily total (AB+BA) small truck volume
A01_MED_VOL	AB medium truck volume in AM peak hour
TOT_A01_MED	Total (AB+BA) medium truck volume in AM peak hour
A03_MED_VOL	AB medium truck volume in AM period
TOT_A03_MED	Total (AB+BA) medium truck volume in AM period
M07_MED_VOL	AB medium truck volume in MD period
TOT_M07_MED	Total (AB+BA) medium truck volume in MD period
P01_MED_VOL	AB medium truck volume in PM peak hour
TOT_P01_MED	Total (AB+BA) medium truck volume in PM peak hour
P03_MED_VOL	AB medium truck volume in PM period
TOT_P03_MED	Total (AB+BA) medium truck volume in PM period
E11_MED_VOL	AB medium truck volume in EV period
TOT_E11_MED	Total (AB+BA) medium truck volume in EV period
D24_MED_VOL	Daily AB medium truck volume
TOT_D24_MED	Daily total (AB+BA) medium truck volume
A01_HVY_VOL	AB heavy truck volume in AM peak hour
TOT_A01_HVY	Total (AB+BA) heavy truck volume in AM peak hour

A03_HVY_VOL	AB heavy truck volume in AM period
TOT_A03_HVY	Total (AB+BA) heavy truck volume in AM period
M07_HVY_VOL	AB heavy truck volume in MD period
TOT_M07_HVY	Total (AB+BA) heavy truck volume in MD period
P01_HVY_VOL	AB heavy truck volume in PM peak hour
TOT_P01_HVY	Total (AB+BA) heavy truck volume in PM peak hour
P03_HVY_VOL	AB heavy truck volume in PM period
TOT_P03_HVY	Total (AB+BA) heavy truck volume in PM period
E11_HVY_VOL	AB heavy truck volume in EV period
TOT_E11_HVY	Total (AB+BA) heavy truck volume in EV period
D24_HVY_VOL	Daily AB heavy truck volume
TOT_D24_HVY	Daily total (AB+BA) heavy truck volume
A01_TRK_VOL	AB truck volume in AM peak hour
TOT_A01_TRK	Total (AB+BA) truck volume in AM peak hour
A03_TRK_VOL	AB truck volume in AM period
TOT_A03_TRK	Total (AB+BA) truck volume in AM period
M07_TRK_VOL	AB truck volume in MD period
TOT_M07_TRK	Total (AB+BA) truck volume in MD period
P01_TRK_VOL	AB truck volume in PM peak hour
TOT_P01_TRK	Total (AB+BA) truck volume in PM peak hour

P03_TRK_VOL	AB truck volume in PM period
TOT_P03_TRK	Total (AB+BA) truck volume in PM period
E11_TRK_VOL	AB truck volume in EV period
TOT_E11_TRK	Total (AB+BA) truck volume in EV period
D24_TRK_VOL	Daily AB truck volume
TOT_D24_TRK	Daily total (AB+BA) truck volume
A01_ASG_SP	AB speed in AM peak hour
A03_ASG_SP	AB speed in AM period
M07_ASG_SP	AB speed in AM period
P01_ASG_SP	AB speed in PM peak hour
P03_ASG_SP	AB speed in PM period
E11_ASG_SP	AB speed in EV period
D24_ASG_SP	Average daily AB speed
AIRBASIN	
LOS_AM	Level of service in AM period
LOS_MD	Level of service in MD period
LOS_PM	Level of service in PM period
LOS_EV	Level of service in EV period
LOS_AM1HR	Level of service in AM peak hour
LOS_PM1HR	Level of service in PM peak hour
TSM	
EJ	
A03_DA	AB drive alone volume in AM peak hour

TOT_A03_DA	Total (AB+BA) drive alone volume in AM peak hour
A03_DA	AB drive alone volume in AM period
TOT_A03_DA	Total (AB+BA) drive alone volume in AM period
M07_DA	AB drive alone volume in MD period
TOT_M07_DA	Total (AB+BA) drive alone volume in MD period
P01_DA	AB drive alone volume in PM peak hour
TOT_P01_DA	Total (AB+BA) drive alone volume in PM peak hour
P03_DA	AB drive alone volume in PM period
TOT_P03_DA	Total (AB+BA) drive alone volume in PM period
E11_DA	AB drive alone volume in EV period
TOT_E11_DA	Total (AB+BA) drive alone volume in EV period
D24_DA	Daily AB drive alone volume
TOT_D24_DA	Daily total (AB+BA) drive alone volume
A03_S2	AB shared-ride 2 volume in AM peak hour
TOT_A03_S2	Total (AB+BA) shared-ride 2 volume in AM peak hour
A03_S2	AB shared-ride 2 volume in AM period
TOT_A03_S2	Total (AB+BA) shared-ride 2 volume in AM period
M07_S2	AB shared-ride 2 volume in MD period

TOT_M07_S2	Total (AB+BA) shared-ride 2 volume in MD period
P01_S2	AB shared-ride 2 volume in PM peak hour
TOT_P01_S2	Total (AB+BA) shared-ride 2 volume in PM peak hour
P03_S2	AB shared-ride 2 volume in PM period
TOT_P03_S2	Total (AB+BA) shared-ride 2 volume in PM period
E11_S2	AB shared-ride 2 volume in EV period
TOT_E11_S2	Total (AB+BA) shared-ride 2 volume in EV period
D24_S2	Daily AB shared-ride 2 volume
TOT_D24_S2	Daily total (AB+BA) shared-ride 2 volume
A03_S3	AB shared-ride 3+ volume in AM peak hour
TOT_A03_S3	Total (AB+BA) shared-ride 3+ volume in AM peak hour
A03_S3	AB shared-ride 3+ volume in AM period
TOT_A03_S3	Total (AB+BA) shared-ride 3+ volume in AM period
M07_S3	AB shared-ride 3+ volume in MD period
TOT_M07_S3	Total (AB+BA) shared-ride 3+ volume in MD period
P01_S3	AB shared-ride 3+ volume in PM peak hour
TOT_P01_S3	Total (AB+BA) shared-ride 3+ volume in PM peak hour
P03_S3	AB shared-ride 3+ volume in PM period

TOT_P03_S3	Total (AB+BA) shared-ride 3+ volume in PM period
E11_S3	AB shared-ride 3+ volume in EV period
TOT_E11_S3	Total (AB+BA) shared-ride 3+ volume in EV period
D24_S3	Daily AB shared-ride 3+ volume
TOT_D24_S3	Daily total (AB+BA) shared-ride 3+ volume

Transit

The transit assignment results are available by two time periods (PK: peak and OK: off-peak), two sub-modes (B: bus and R: rail), and two access modes (W: walk and D: drive). The outputs are produced both in CUBE network format ({scenario_name}_VOL_{tod}_T{access_mode}{submode}.NET), as well as in database file format {scenario_name}_VOL_{tod}_T{access_mode}{submode}.DBF). The outputs contain attributes as shown in Table 81.

TABLE 81: ATTRIBUTES IN TRANSIT ASSIGNMENT RESULTS

FIELD	DESCRIPTION
A	Node A
B	Node B
MODE	Transit sub-mode id 1: local bus (FAX and Clovis) 2: local bus (FCRTA)
OPERATOR	Transit operator
NAME	Transit line name
LONGNAME	Transit line long name
DIST	Distance (miles)
TIME	Transit travel time
LINKSEQ	Link sequence

HEADWAY_2	Headway
STOPA	Is node A stop (1: yes, 0: no)
STOPB	Is node B stop (1: yes, 0: no)
VOL	Transit volume
ONA	On-boarding at node A
OFFA	Off-boarding at node A
ONB	On-boarding at node B
OFFB	Off-boarding at node B
REV_VOL	Revised transit volume
REV_ONA	Revised on-boarding at node A
REV_OFFA	Revised off-boarding at node A
REV_ONB	Revised on-boarding at node B
REV_OFFB	Revised off-boarding at node B

Non-motorized

The non-motorized assignment results are output in database file format: FC14_BASE_LOADEDNETWORK_BIKE.DBF and FC14_BASE_LOADEDNETWORK_WALK.DBF. Attributes of bike and walk assignment outputs are presented in Table 82 and Table 83.

TABLE 82: ATTRIBUTES IN BIKE ASSIGNMENT RESULTS

FIELD	DESCRIPTION
A	Node A
B	Node B
DISTANCE	Distance (miles)

CAPCLASS	Capacity class
LANES	Number of lanes
NAME	Street name
ROUTE	Route number
TERRAIN	Terrain
JURISDICTIO	Jurisdiction
SCREENLINE	Screen line id
SPEED	Posted speed (mph)
AREATYP	Area type
FACTYP	Facility type 0: local 1: Freeway 2: Highway 3: Expressway 4: Arterial 5: Collector 6: Local 7: Ramp-Freeway-Freeway 8: Ramp-Slip 9: Ramp-Loop 10: Connector1 11: Connector2
AUX	Presence of auxiliary lane
USE	Use type
TOLL	Toll (cents?)
IMPROVED	Improvement id

BIKE_FACTY	Bike facility type : Shared Roadway (No Bikeway Designation). 1: Class I Bikeway (Bike Path). Provides a completely separated right of way for the exclusive use of bicycles and pedestrians with crossflow by motorists minimized. 2: Class II Bikeway (Bike Lane). Provides a striped lane for one-way bike travel on a street or highway. 3: Class III Bikeway (Bike Route). Provides for shared use with pedestrian or motor vehicle traffic. 4: Class IV Separated Bikeway (cycle tracks). On-street bicycle facilities that include a vertical physical barrier between the bikeway and moving traffic. 5: Separate highway overcrossings 6: Unpaved Multipurpose Trails 9: Freeways and Ramps (bicycling not permitted)
BIKE_VOL	Bike volume

TABLE 83: ATTRIBUTES IN WALK ASSIGNMENT RESULTS

FIELD	DESCRIPTION
A	Node A
B	Node B
DISTANCE	Distance (miles)
CAPCLASS	Capacity class
LANES	Number of lanes
NAME	Street name

ROUTE	Route number
TERRAIN	Terrain
JURISDICTIO	Jurisdiction
SCREENLINE	Screen line id
SPEED	Posted speed (mph)
AREATYP	Area type
FACTYP	Facility type 0: local 1: Freeway 2: Highway 3: Expressway 4: Arterial 5: Collector 6: Local 7: Ramp-Freeway-Freeway 8: Ramp-Slip 9: Ramp-Loop 10: Connector1 11: Connector2
AUX	Presence of auxiliary lane
USE	Use type
TOLL	Toll (cents?)
IMPROVED	Improvement id
BIKE_FACTY	Bike facility type : Shared Roadway (No Bikeway Designation). 1: Class I Bikeway (Bike Path). Provides a completely separated right of way for the

	exclusive use of bicycles and pedestrians with crossflow by motorists minimized. 2: Class II Bikeway (Bike Lane). Provides a striped lane for one-way bike travel on a street or highway. 3: Class III Bikeway (Bike Route). Provides for shared use with pedestrian or motor vehicle traffic. 4: Class IV Separated Bikeway (cycle tracks). On-street bicycle facilities that include a vertical physical barrier between the bikeway and moving traffic. 5: Separate highway overcrossings 6: Unpaved Multipurpose Trails 9: Freeways and Ramps (bicycling not permitted)
WALK_VOL	Pedestrian volume

APPENDIX A. SYNTHETIC POPULATION

This appendix describes the setup and process to generate synthetic population for the Fresno activity-based model. The synthetic population is generated using PopSynIII that work with a MySQL database and is, for most part, automated in R.

The objective of this task is to update the existing MTC PopSynIII setup to work for Fresno ABM. All the data processing scripts are written in R including processes to build geographic crosswalks, download Census data, build controls, process the Public Use Microdata Sample (PUMS), uploading the final dataset to MySQL server and generate validation summaries and plots. The following sections of this memo describe instructions on setting up a PopSynIII run, details of R scripts, and validation results.

SOFTWARE REQUIREMENTS

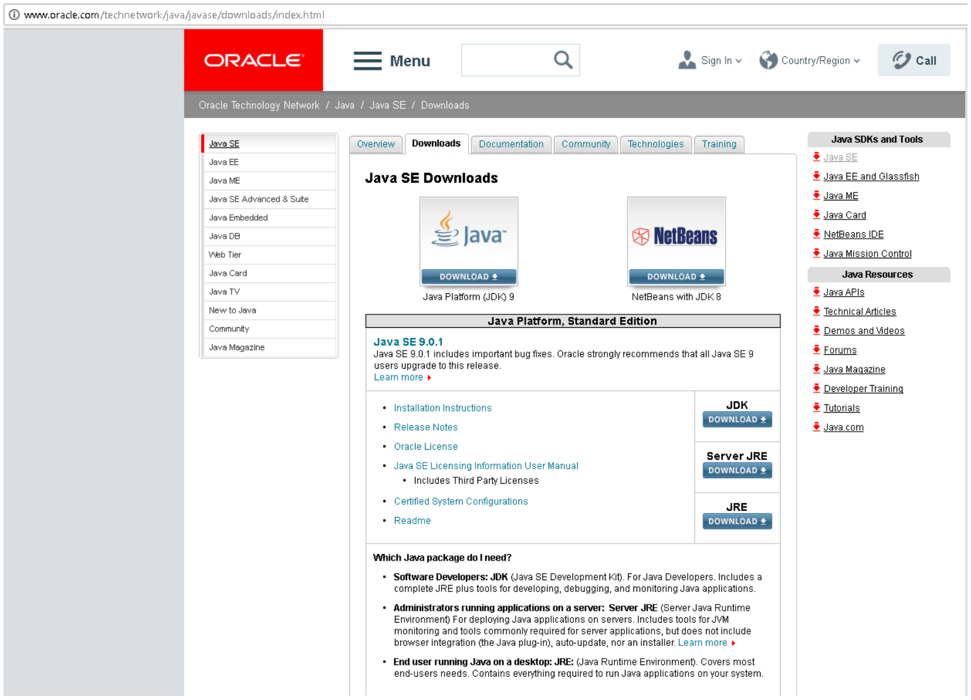
The instructions below are for a machine with Windows operating systems (Windows 7). To setup PopSynIII, following software are required:

- Java
- MySQL
- R

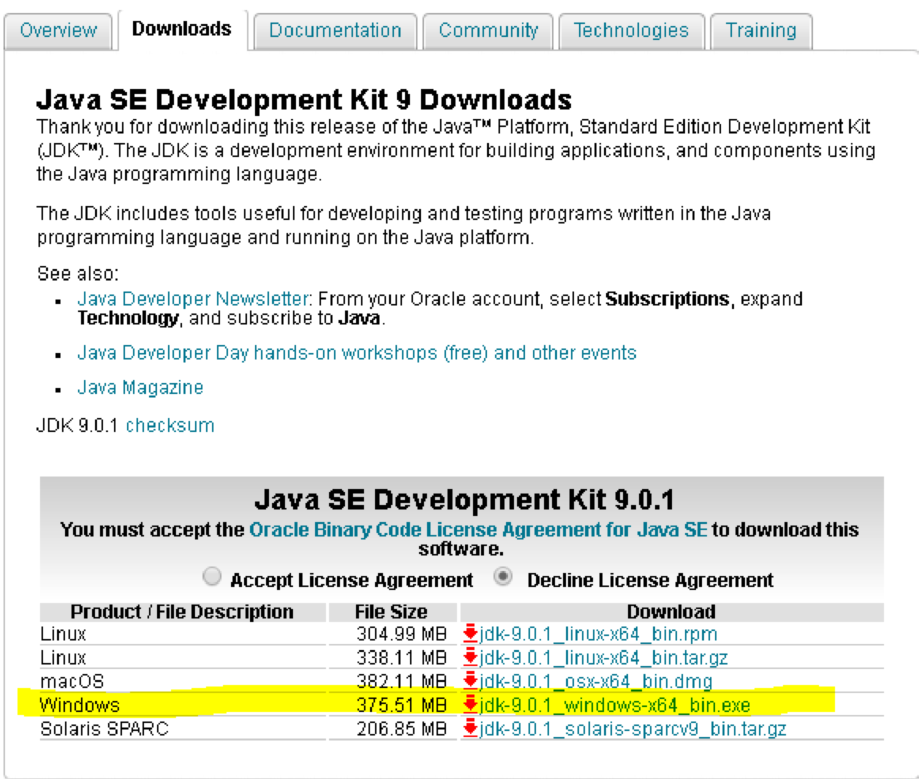
The subsequent sub-sections provide installation steps for the above software.

Java

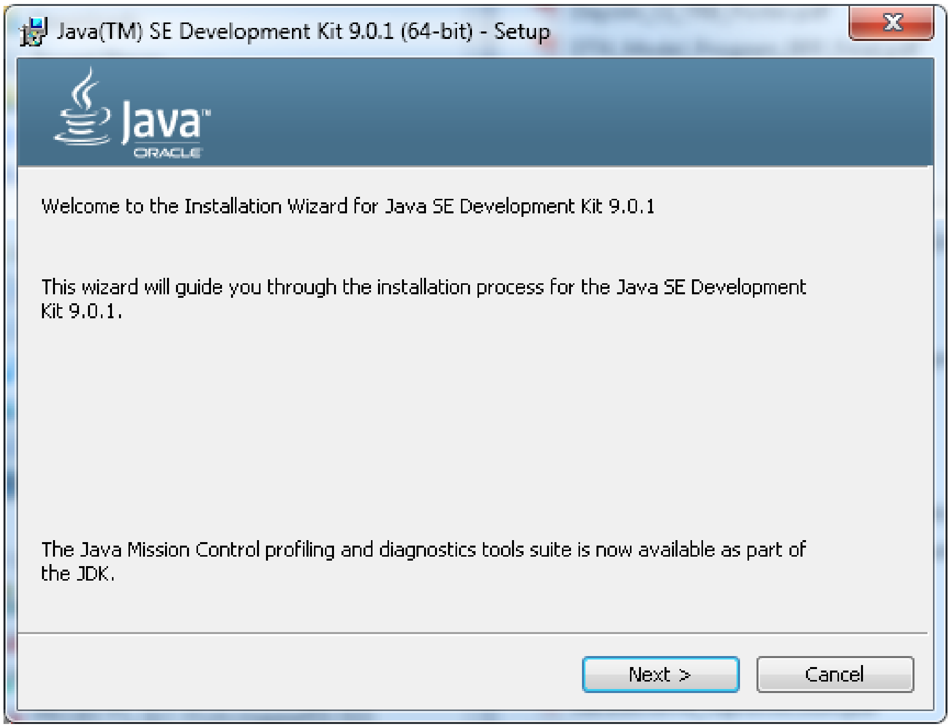
Go to the following link to download latest Java:
http://www.oracle.com/technetwork/java/javase/downloads/index.html



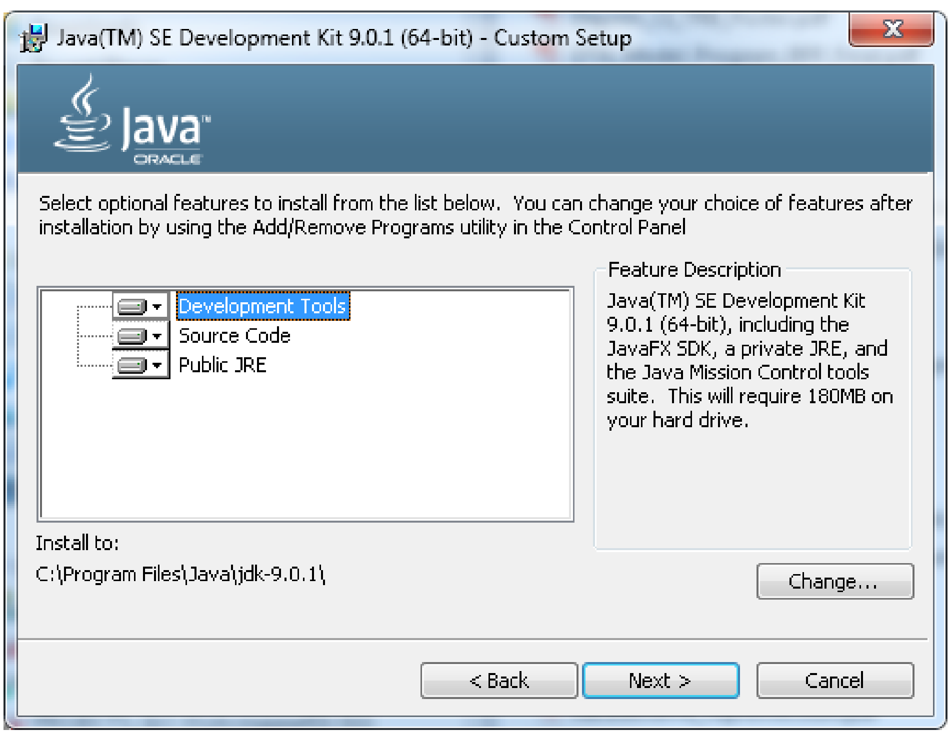
As of the writing, Java SE Development Kit 9.0.1 is the latest Java product. Click on “Java Platform (JDK) 9” under Java SE Downloads.



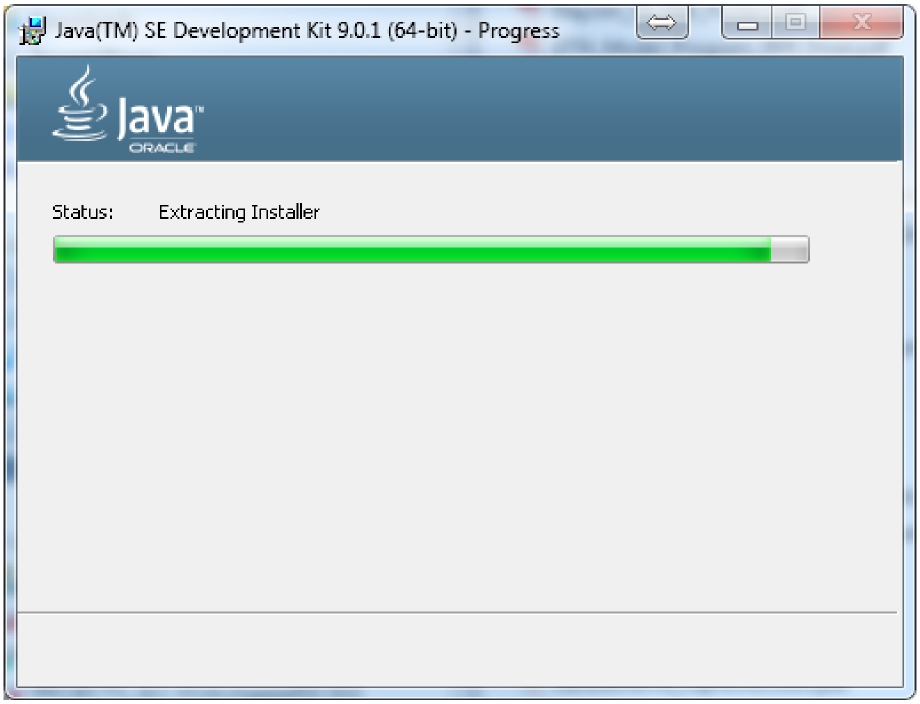
Download the Windows version (highlighted above). Once the download is finishes, go to the downloaded folder and double click on the executable (jdk-9.0.1_windows-x64_bin.exe). This will bring up a setup window.



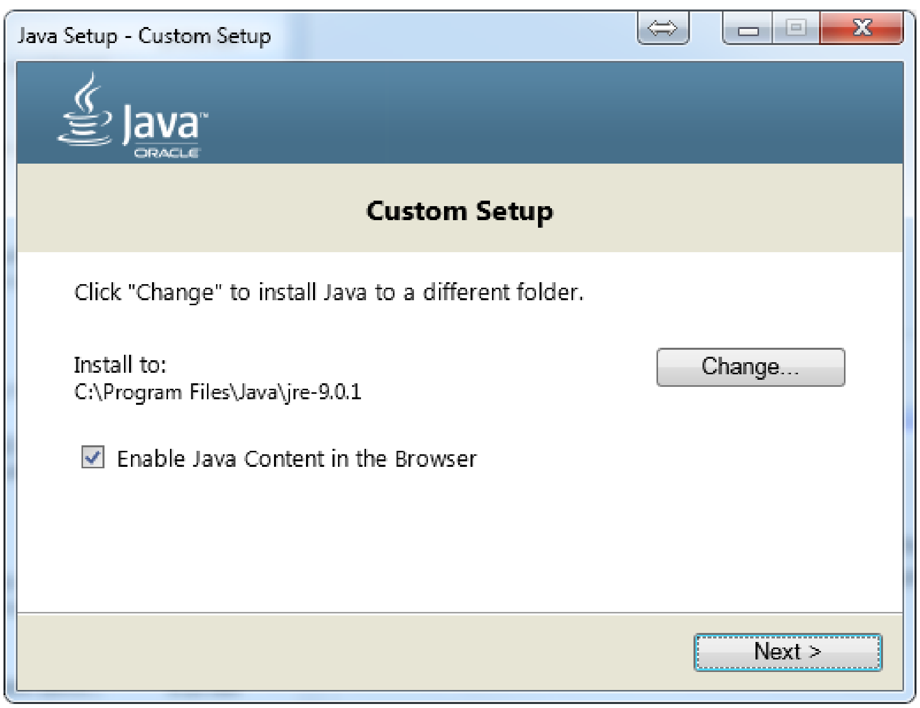
Click “Next” to go to custom setup.



Choose a new installation directory if you want to install in a different location. Otherwise, keep the default and click “Next”.

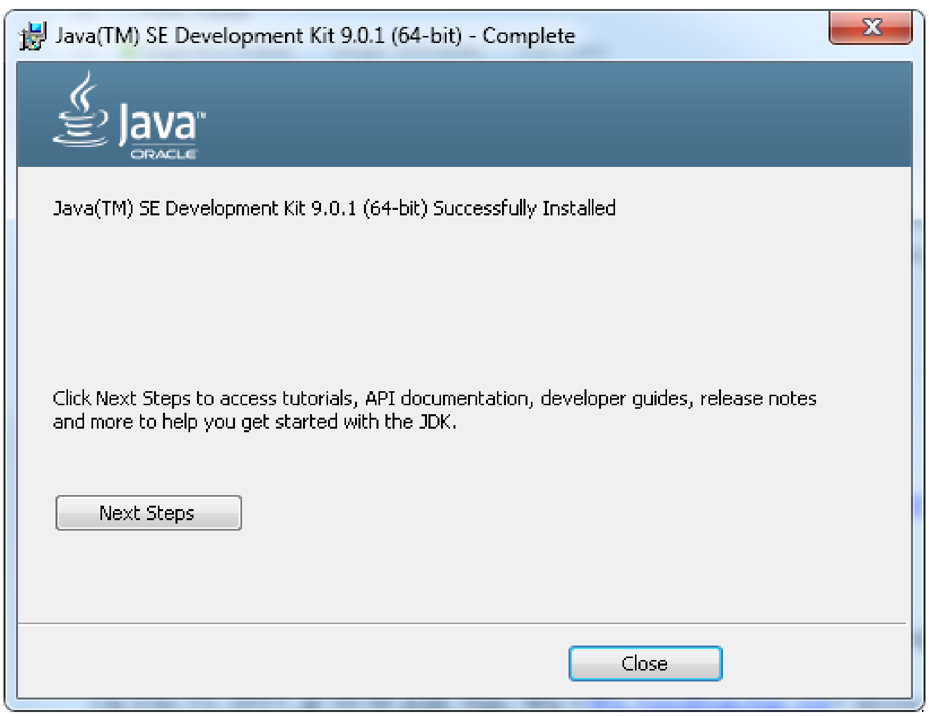


This will start extraction of the installer.



Choose a new installation directory if you want to install in a different location. Otherwise, keep the default and click “Next”.

The installation will begin, and a progress can be monitored on the Setup window.

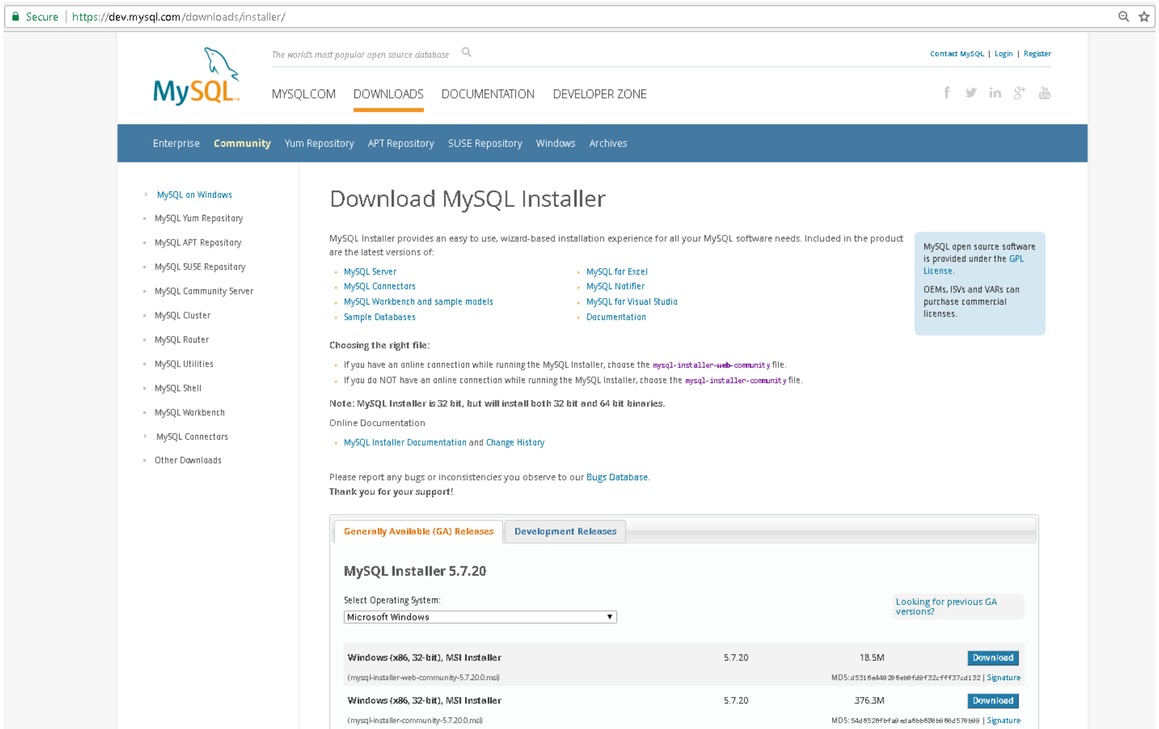


Once complete, click “Close” to finish the installation.

MySQL

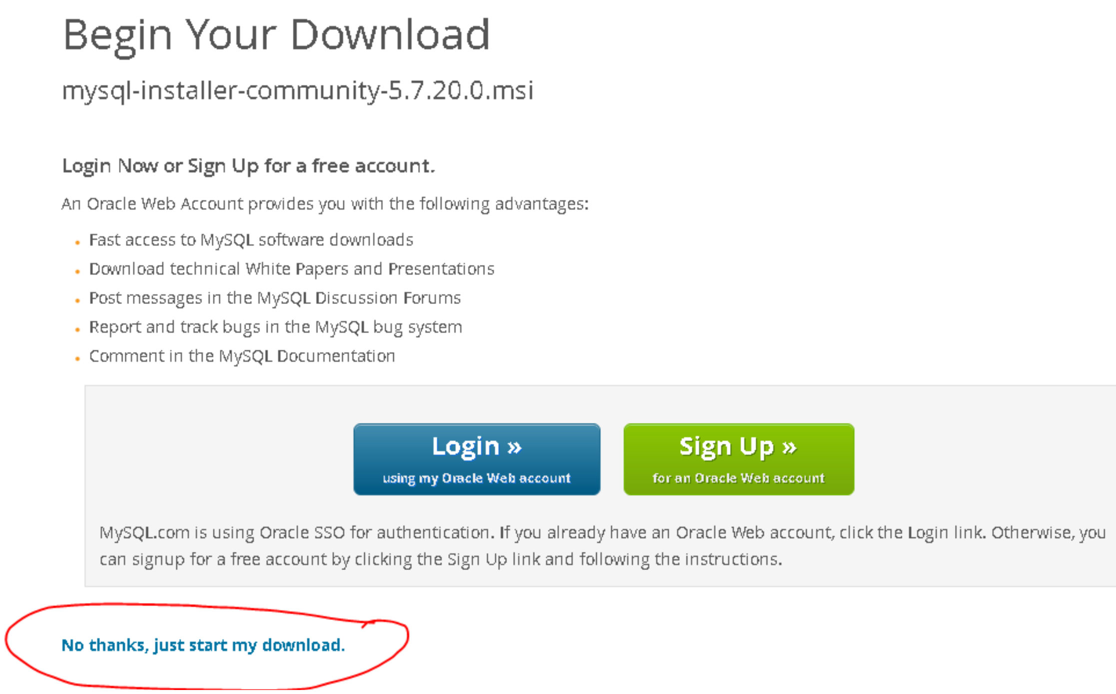
Go to the following web link to download MySQL Workbench:
https://dev.mysql.com/downloads/installer/

The webpage looks similar to this:

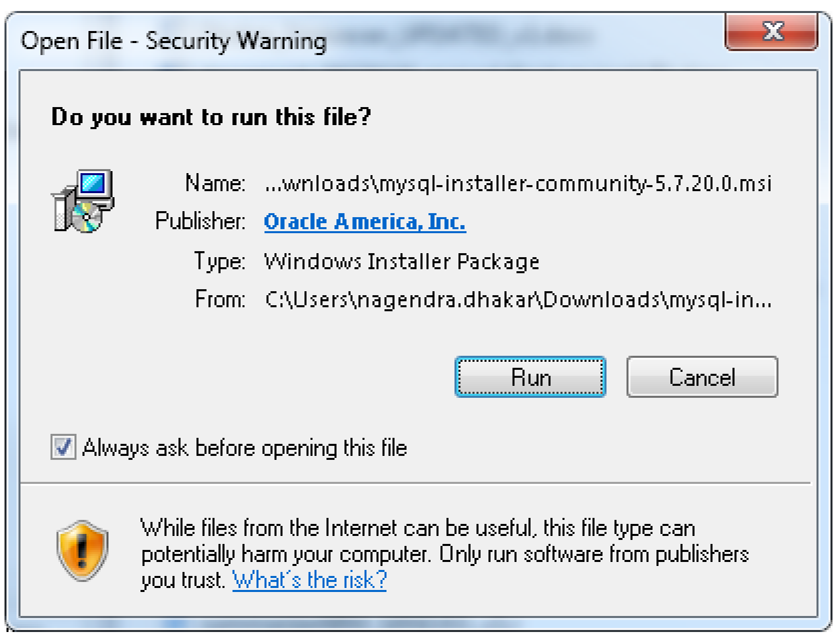


Bottom of the page provides options to download MySQL installer 5.7.20 for Microsoft Windows. Choose the 2nd option to download. This will download the full installer on your disk instead of an installer that would need an online connection while installing (1st option).

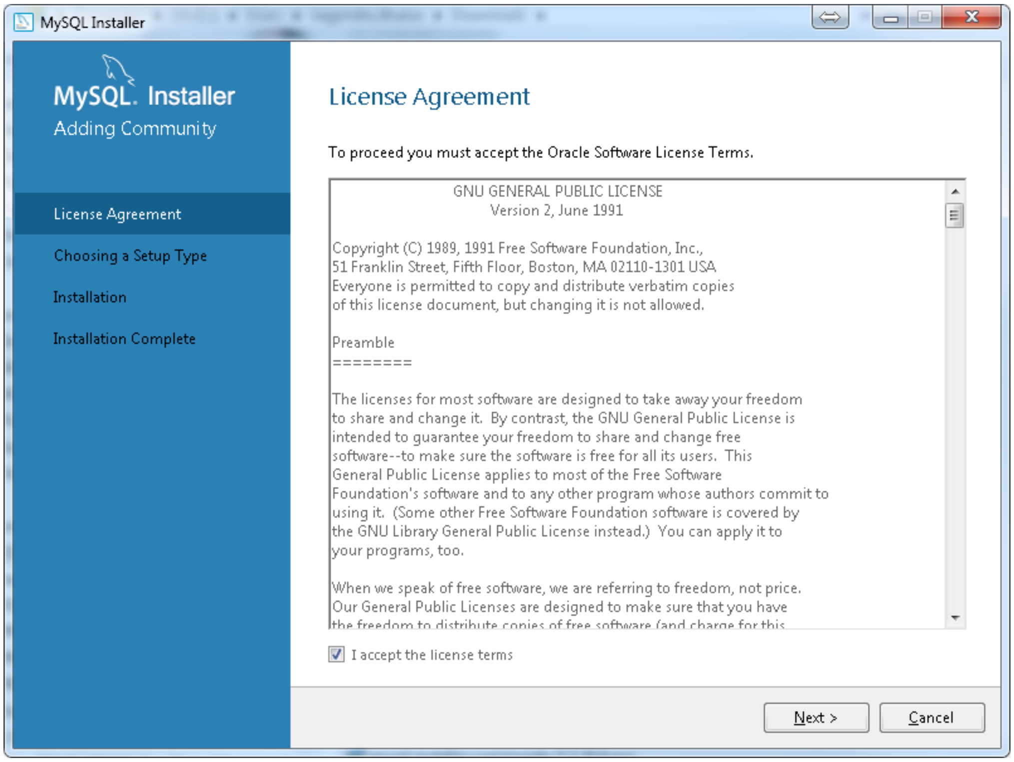
A click on "Download" brings a new page asking to login and signup to begin download. You can choose not to do either and directly download by clicking on a link at the bottom: "No thanks, just start my download".



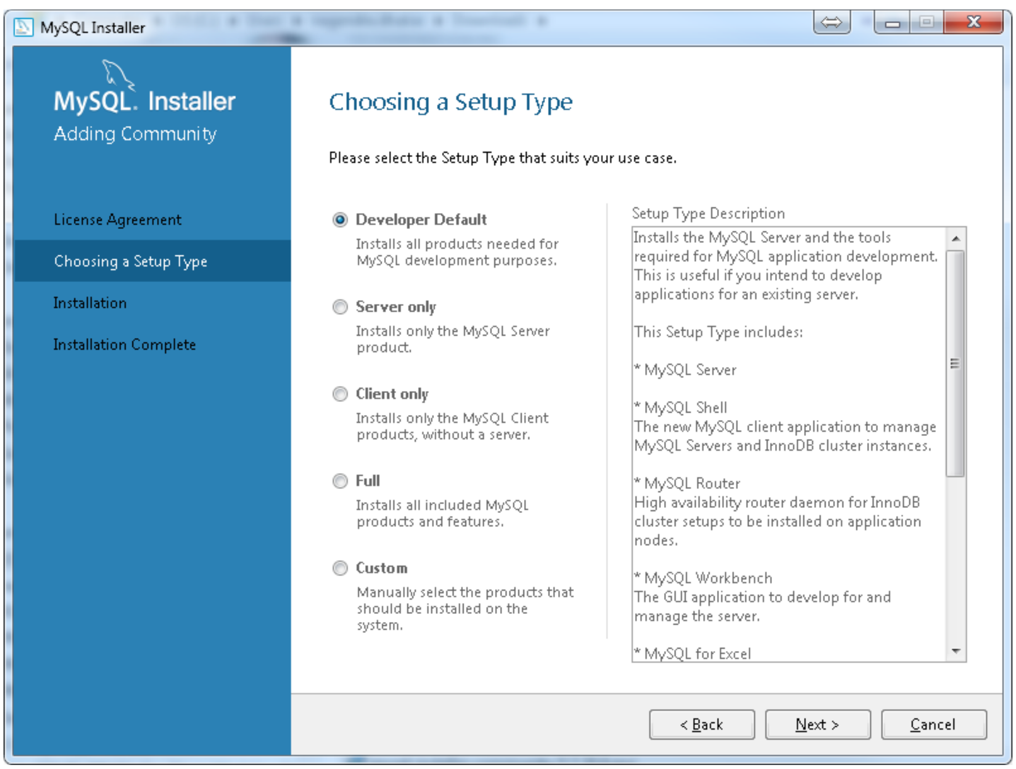
When download is finished, go to your download folder and double click on the MySQL Workbench installer: "mysql-installer-community-5.7.20.0.msi".



Click "Run" and it will bring up a Setup Wizard.

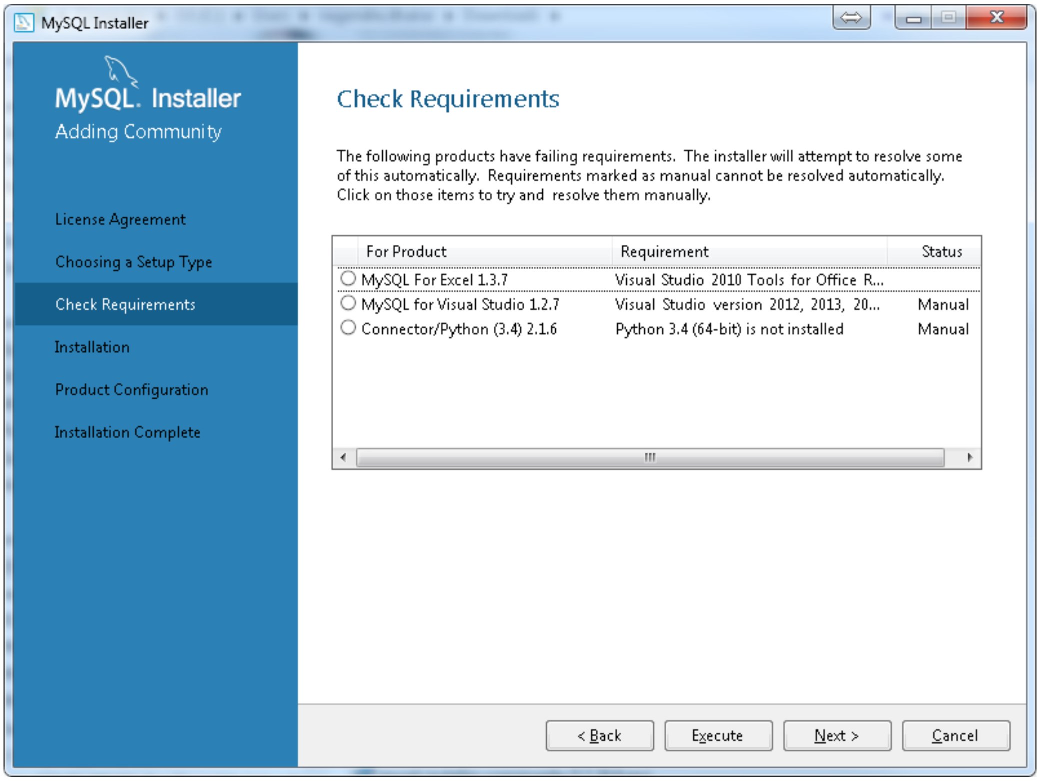


Check “I accept the license terms” and click on “Next”. The next page asks you to choose a setup Type.

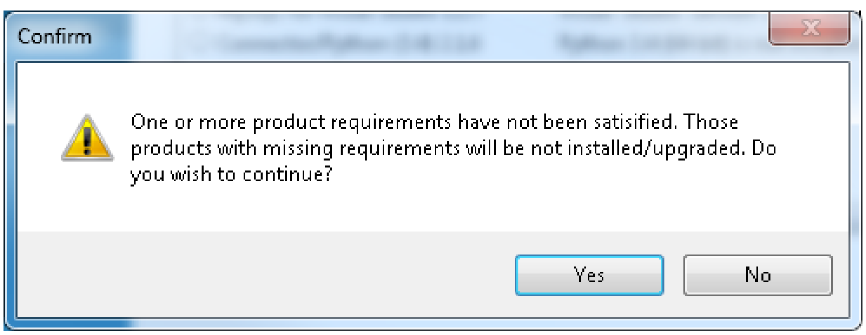


Choose “Developer Default” as it includes both MySQL Server and MySQL Workbench. These two are needed to run PopSynIII setup. Click “Next”.

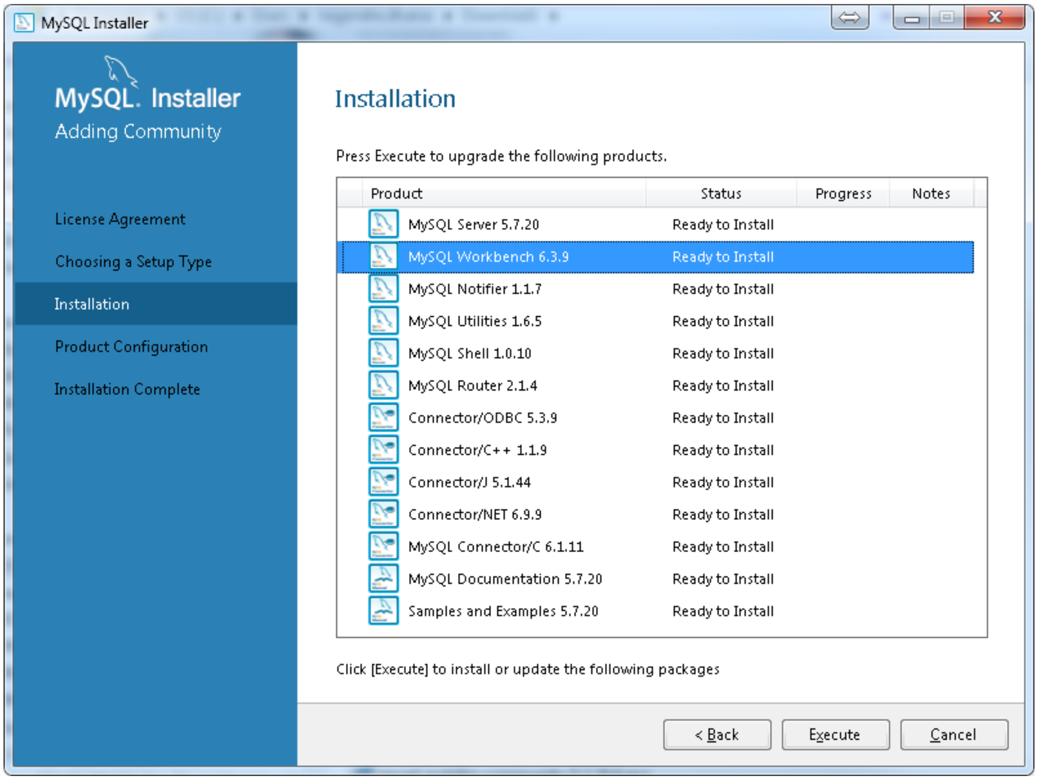
The next page will list products that fail requirements. You can ignore these as long as they do not include the two-necessary product (MySQL Server and MySQL Workbench).



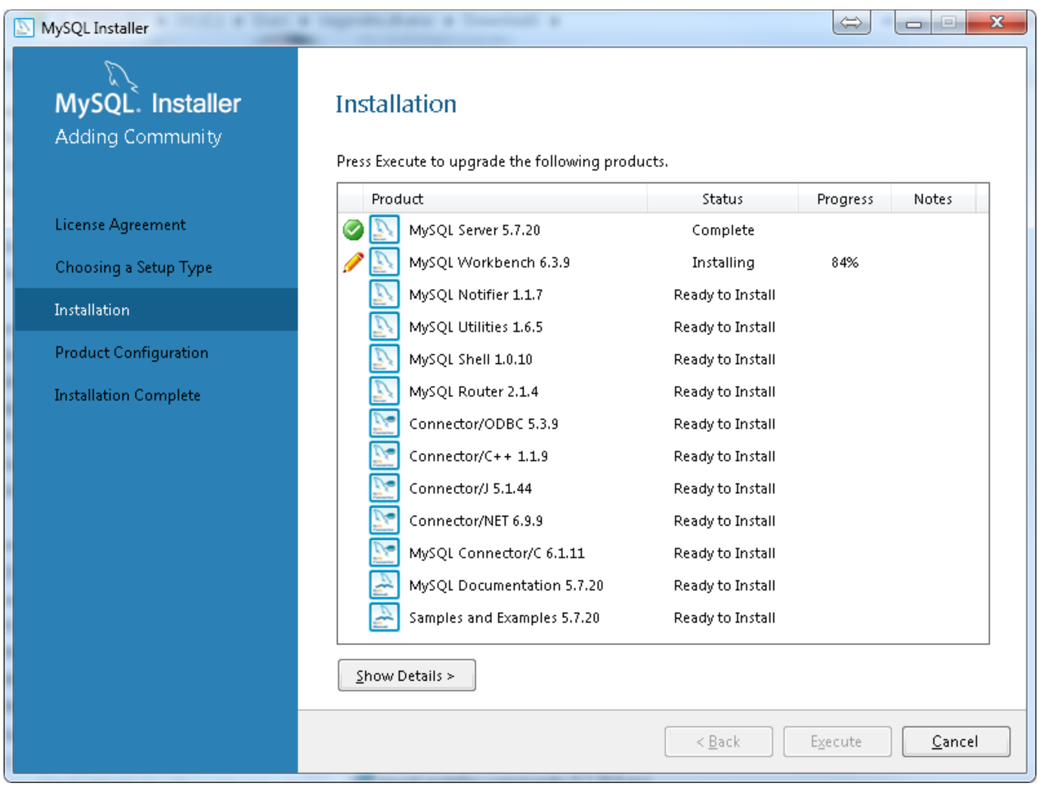
Ignore the list and click “Next”. This will bring up a confirmation window.



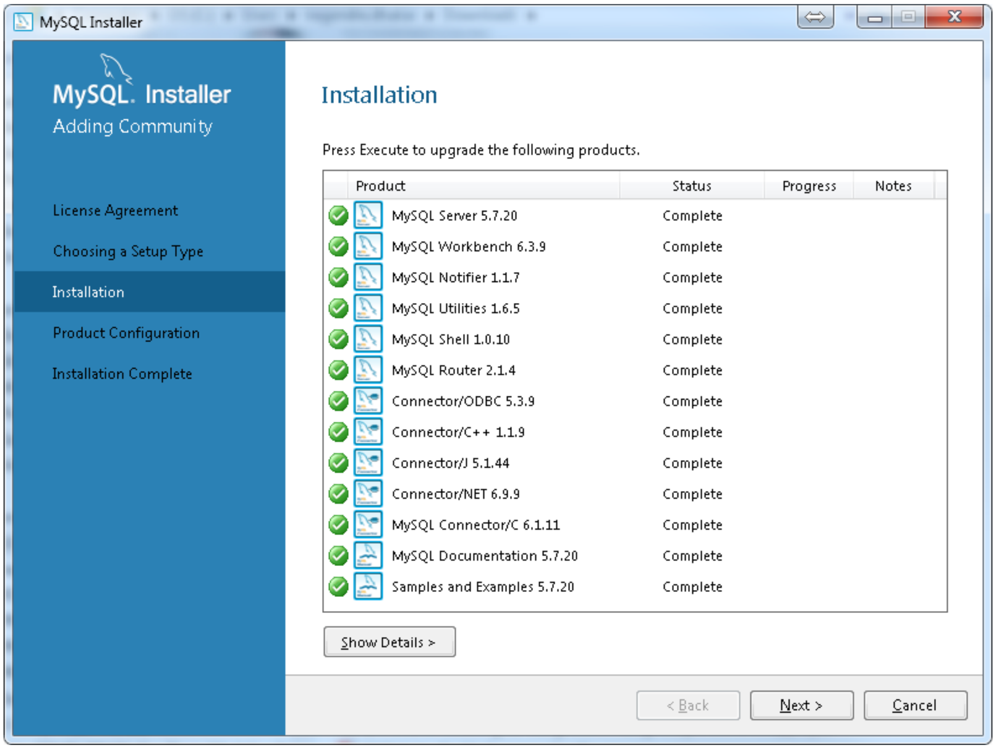
Click “Yes” to bring “Installation” window.



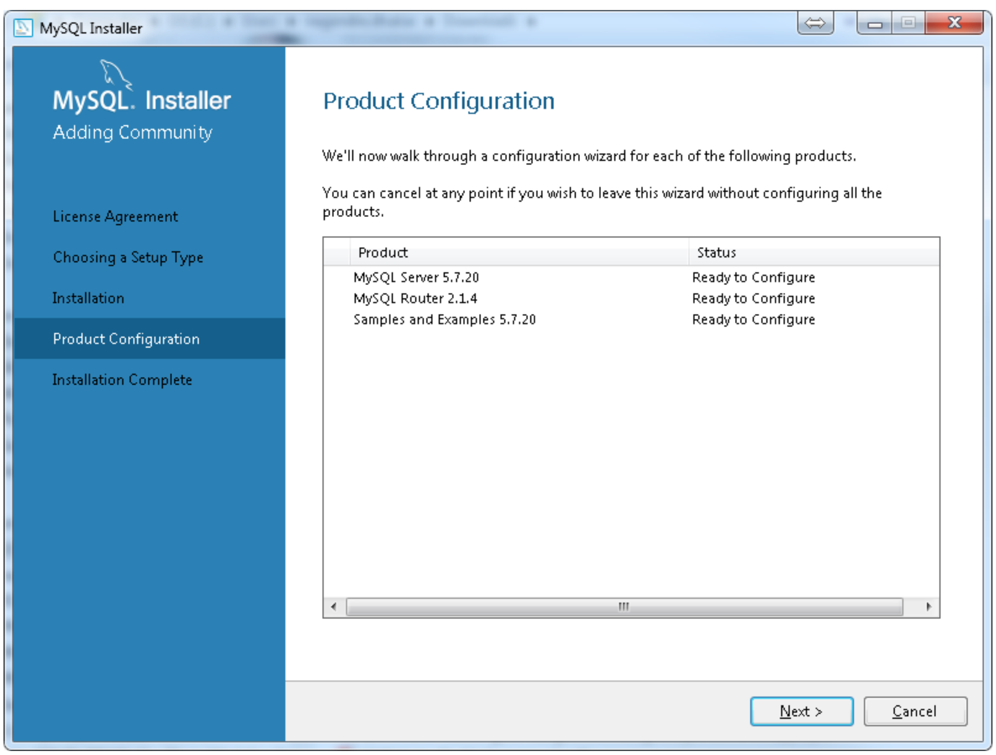
The installation window lists products that are ready to installed. Click “Execute”.



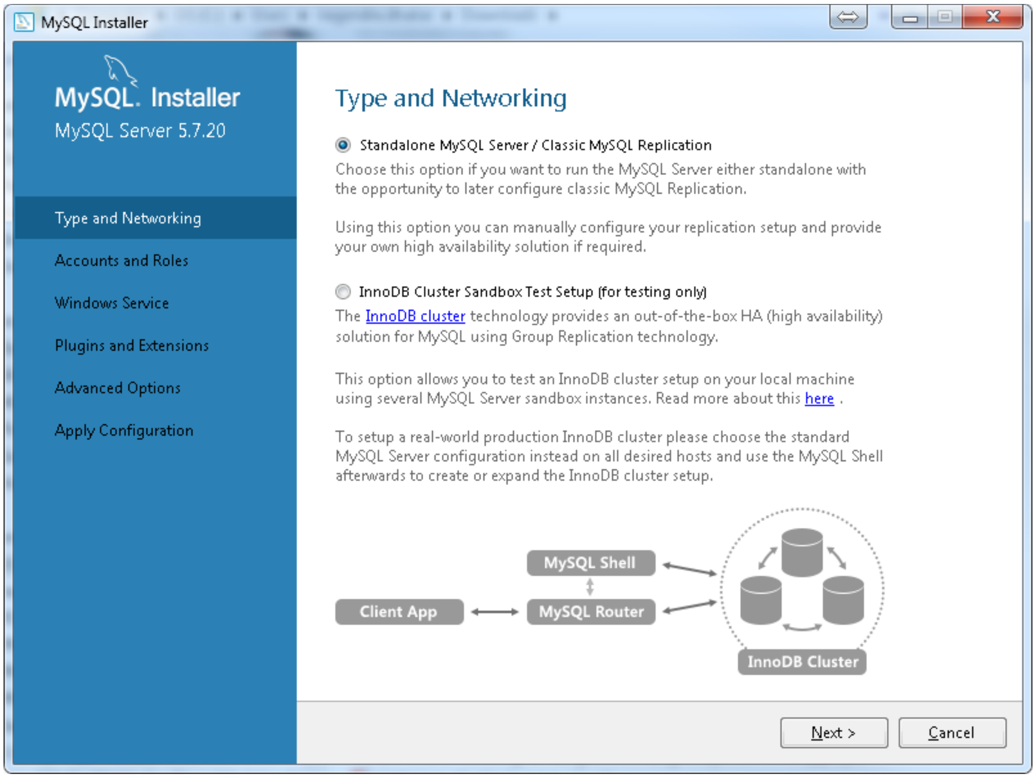
For next few minutes, it will installed the products in the list. You can see the products that are already installed and the one is under progress. Once all products are installed, a green mark will appear in-front of all products.



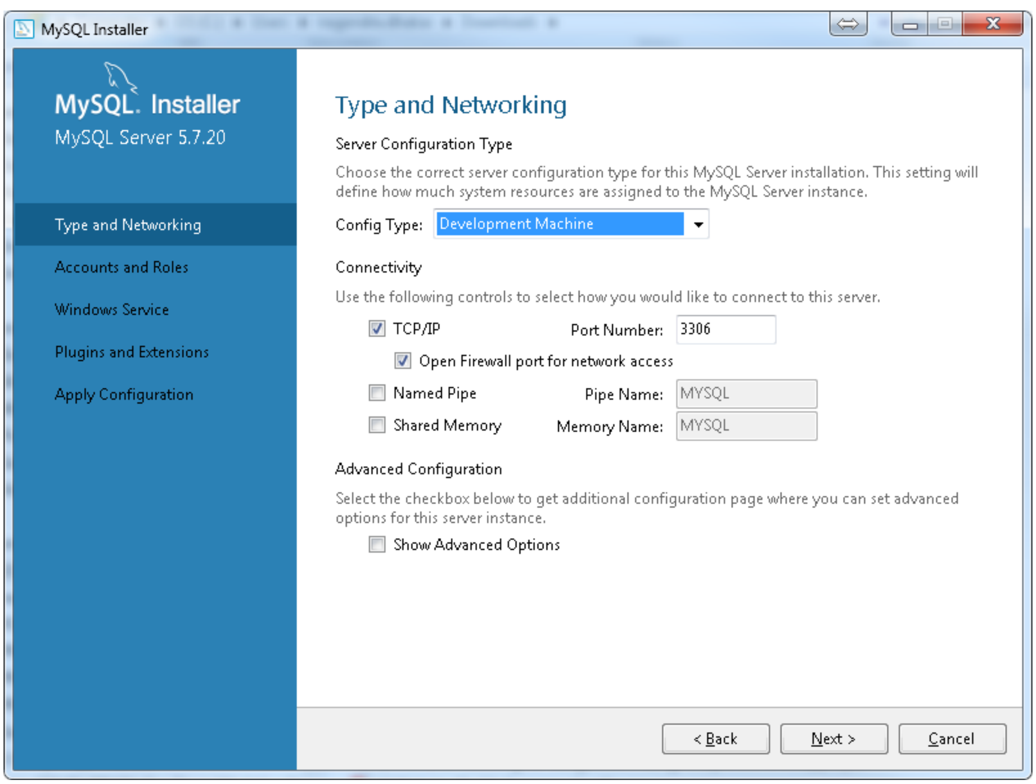
Click “Next” to bring “Product Configuration”.



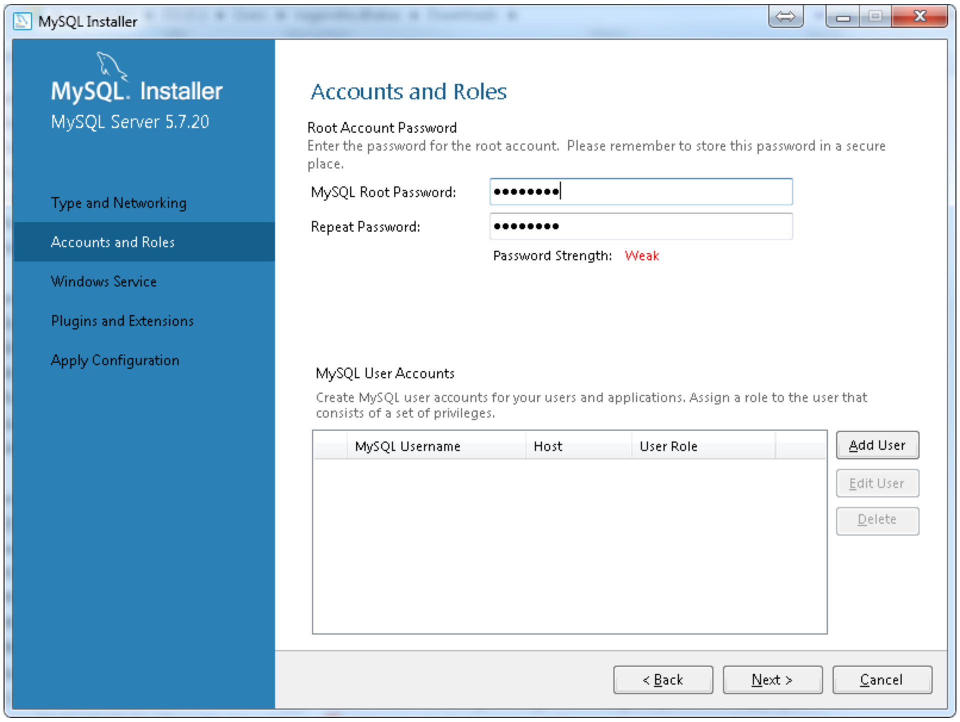
Click "Next"



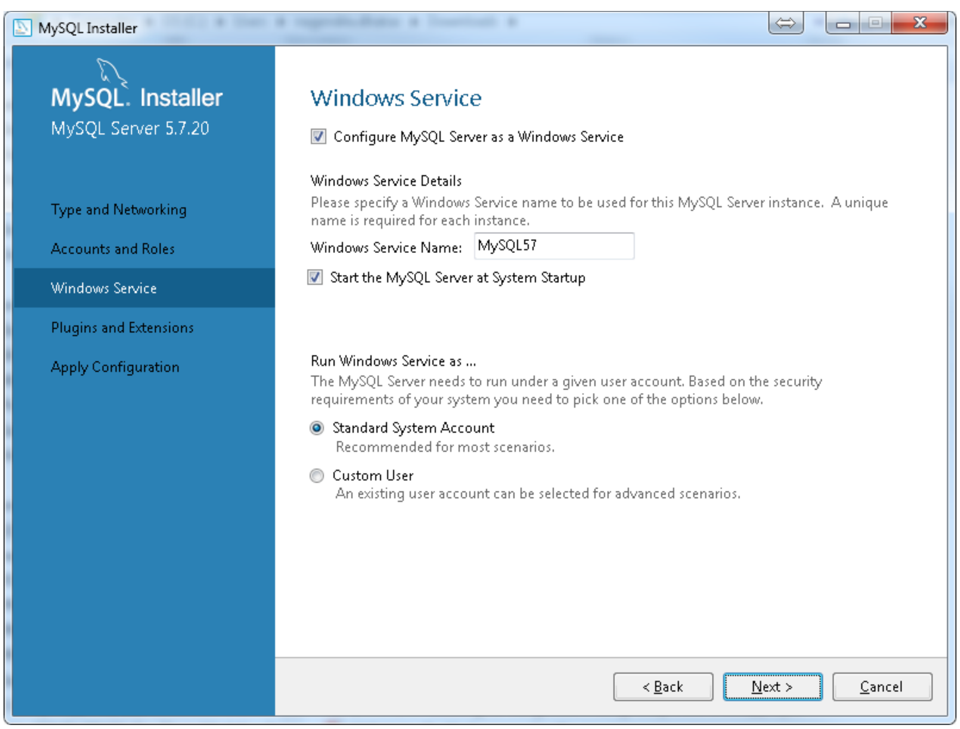
Choose "Standalone MySQL Server/Classic MySQL Replication" and click "Next".



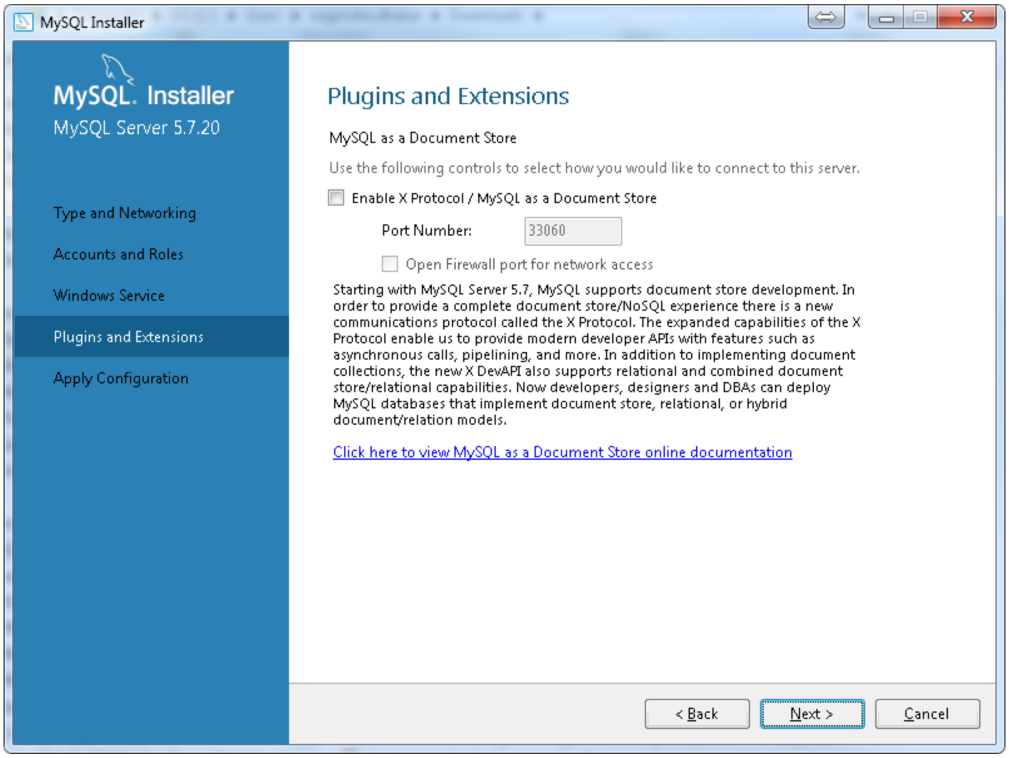
Next page asks you to select “config type”. The available options are based on how much system resources are assigned to the MySQL Server instance. You can choose “Development Machine”. Keep other settings as default. Click “Next”.



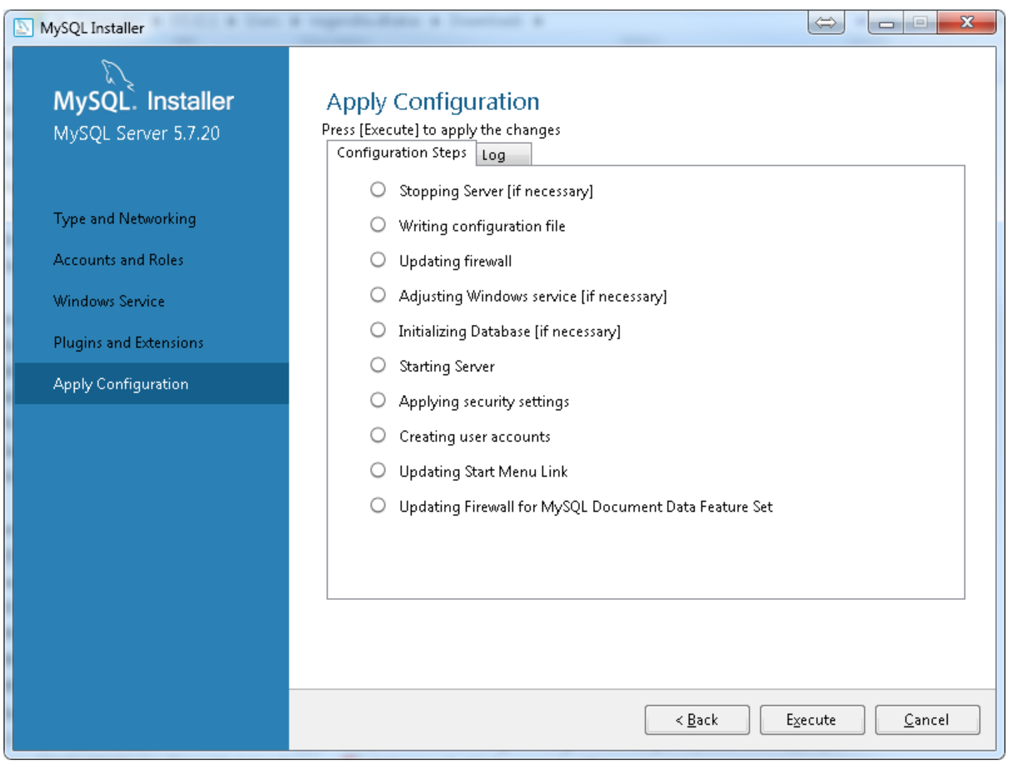
Enter a new password of your choice for the root account. This password will be required while connecting to MySQL Server through R Scripts. Click “Next”.



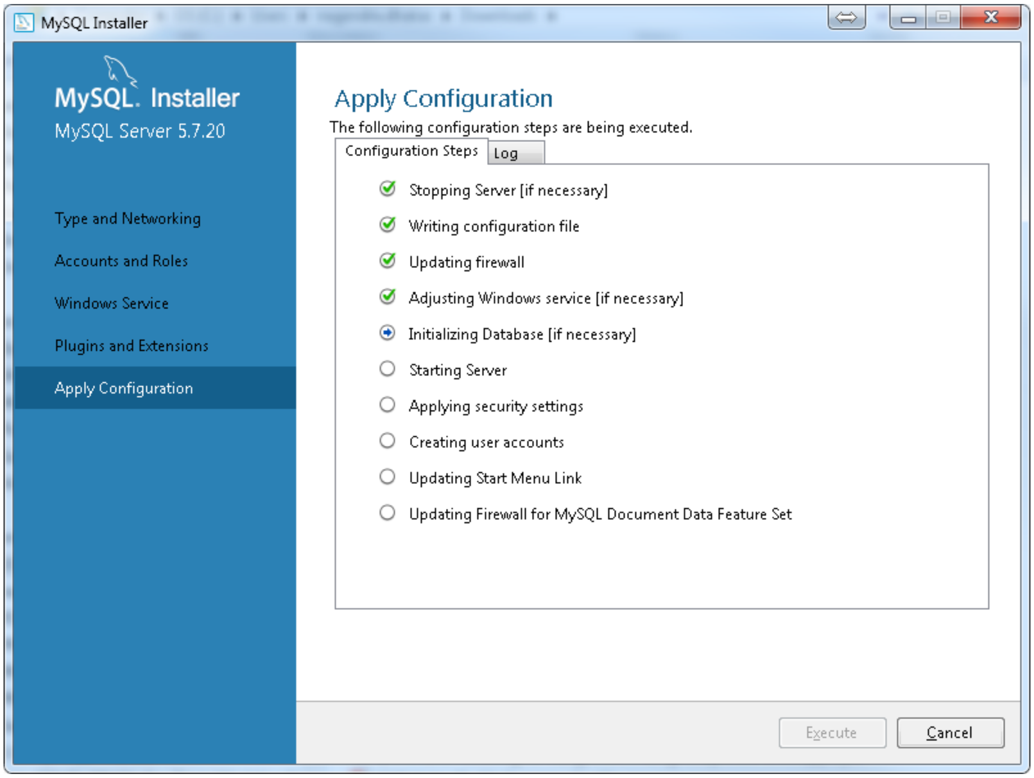
Keep the default settings as show in the above picture. Click “Next”.



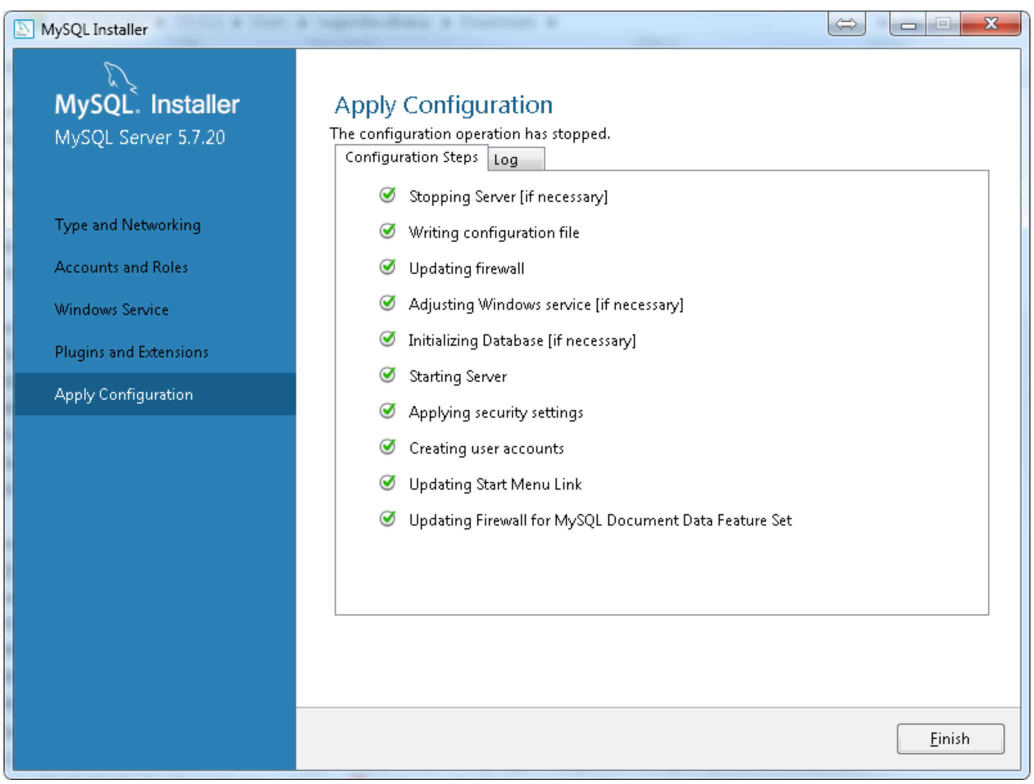
Keep default settings as show in the above picture and click “Next”.



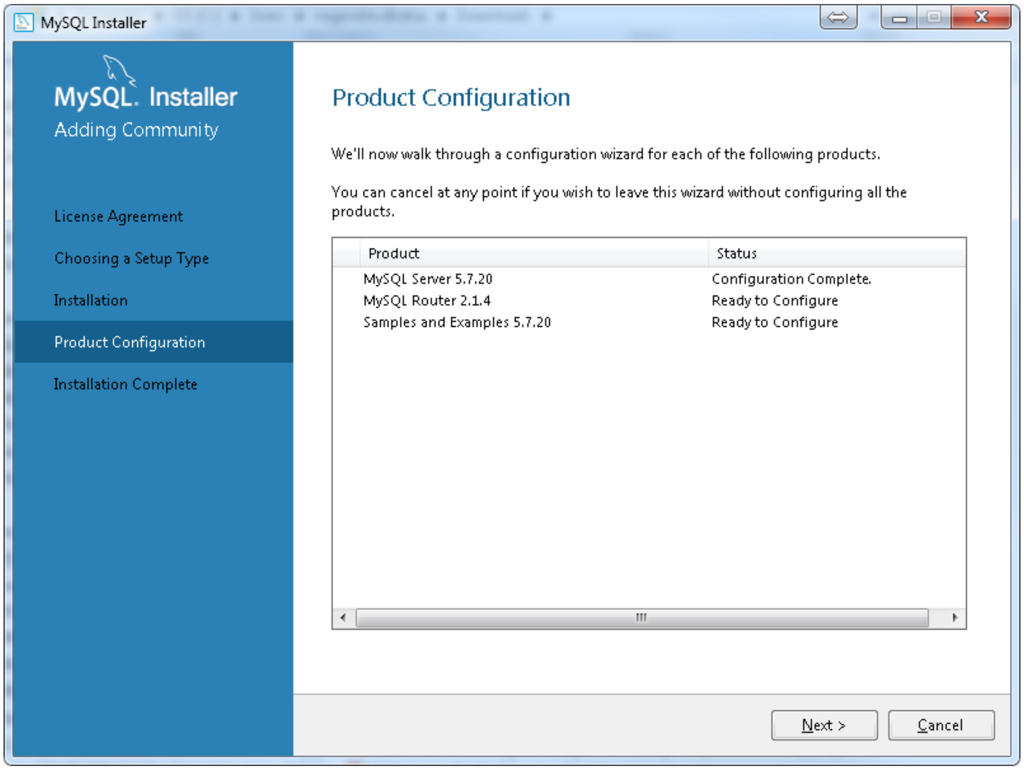
Next page lists configuration steps that will be applied. Click “Execute” to apply these steps.



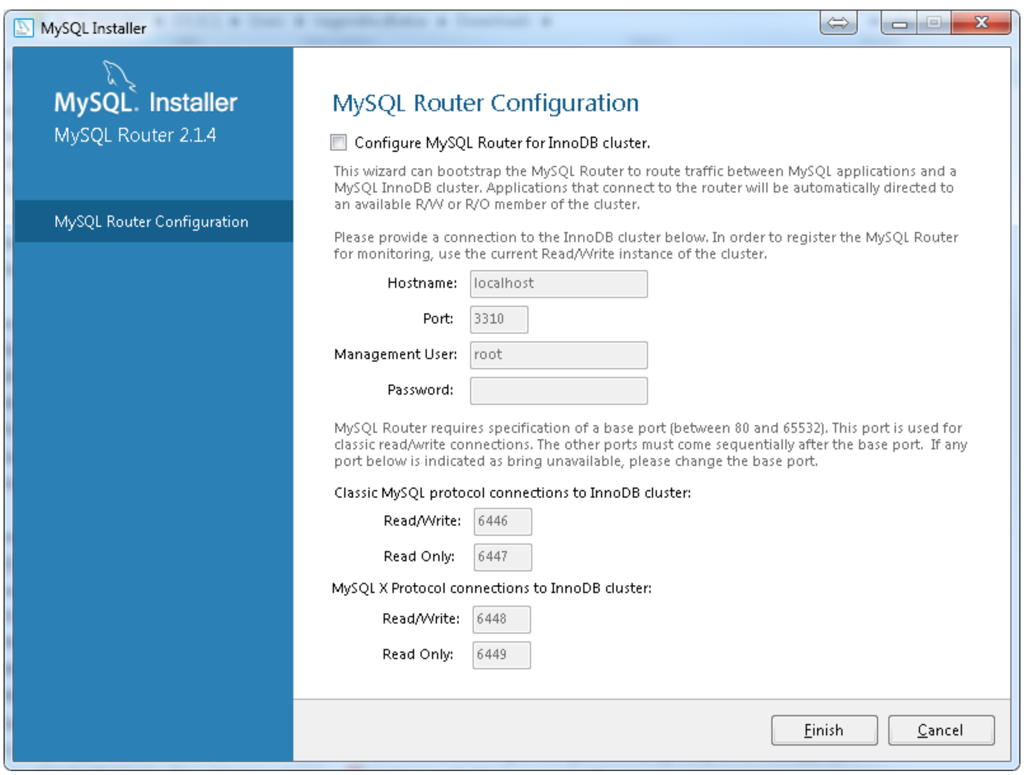
For next few minutes, it will apply those configuration settings. The window will show progress.



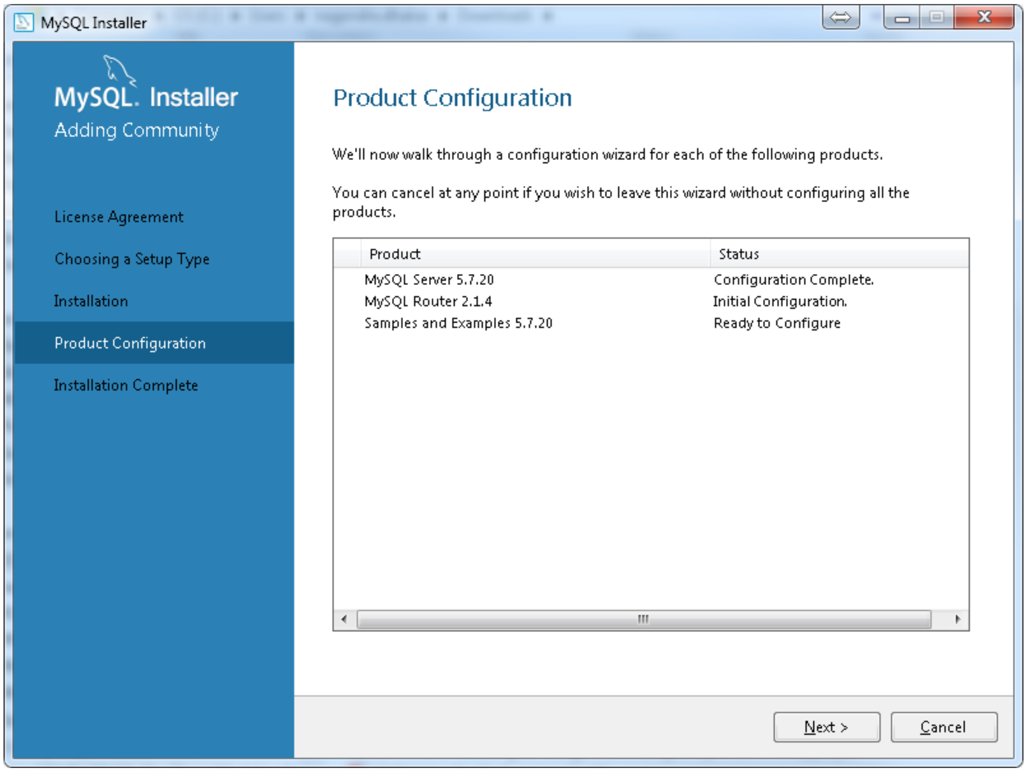
After all steps are applied. A “finish” button will appear. Click “finish” to go to “Product Configuration”.



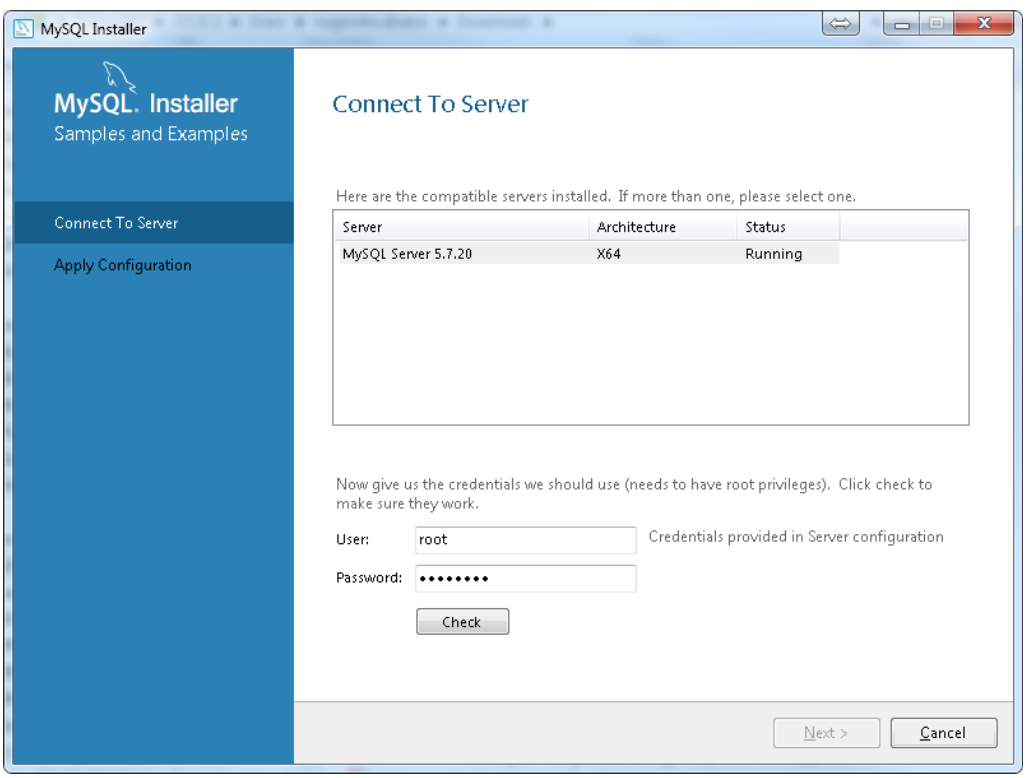
Click “Next”.



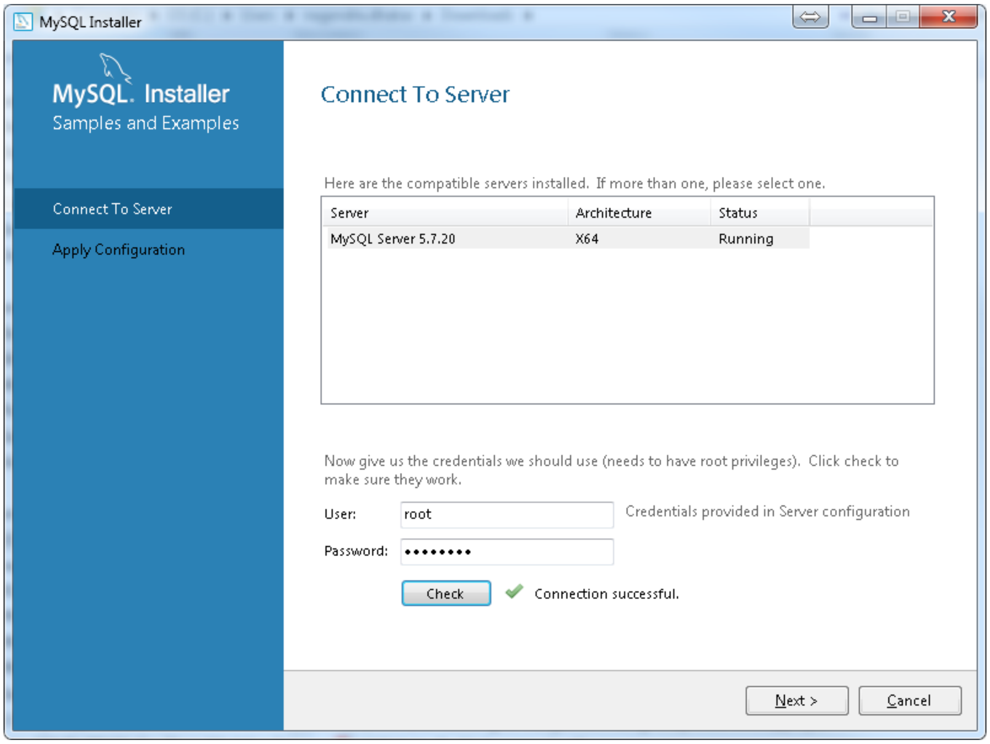
Keep default settings as shown above and click "Finish' to complete MySQL Router configuration.



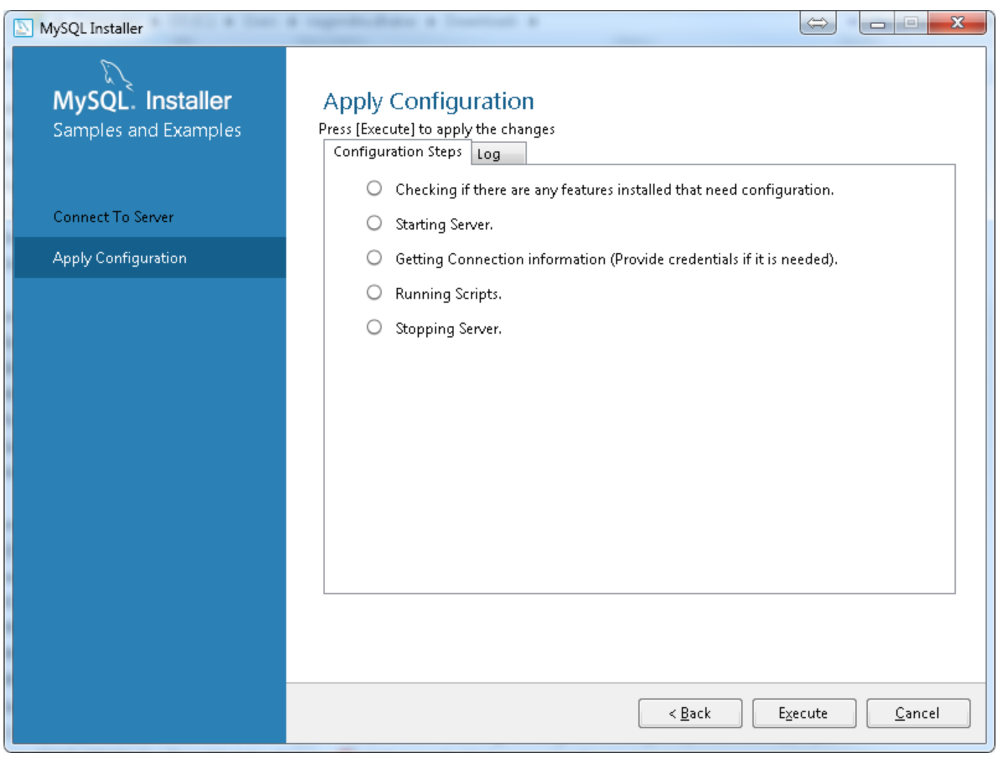
Click "Next".



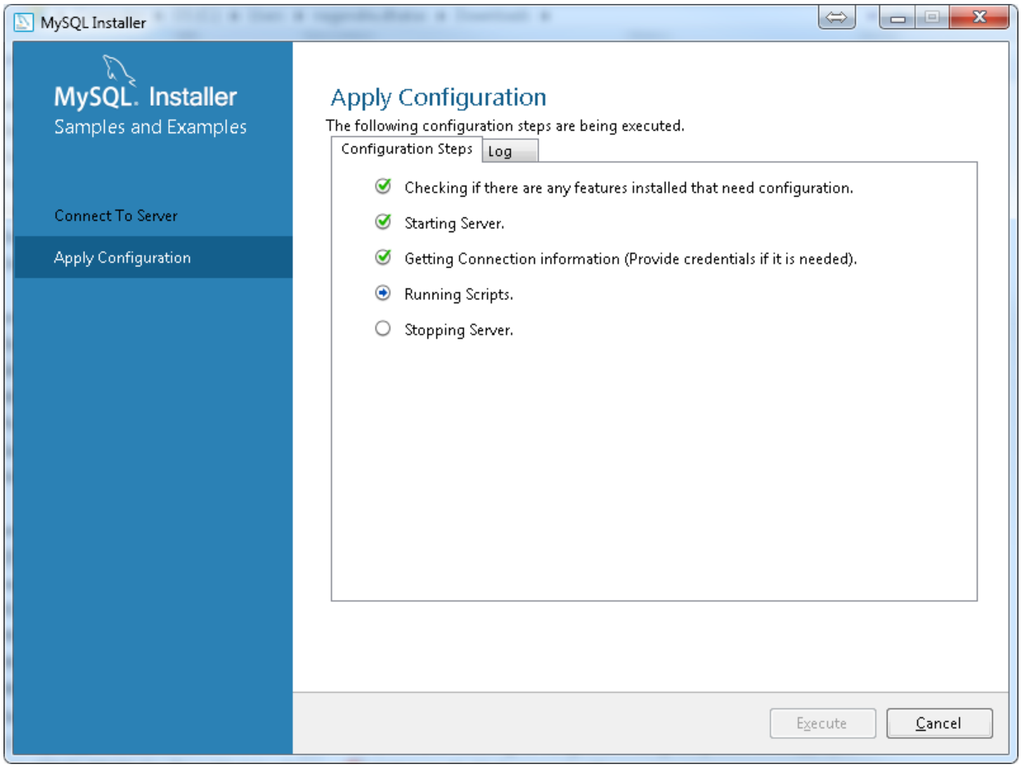
Next page let you test connection to a server. By default it would have root password. Click on “Check” to see if credential works.



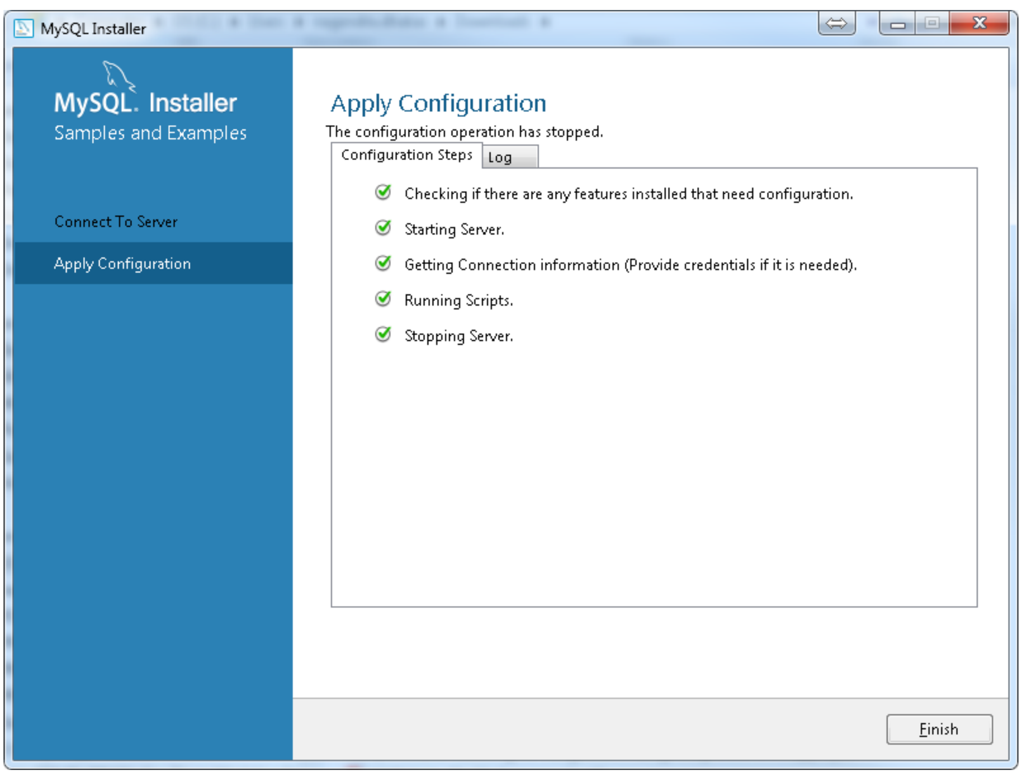
A successful connection will mean you can connect to the server with these credentials. Click “Next”



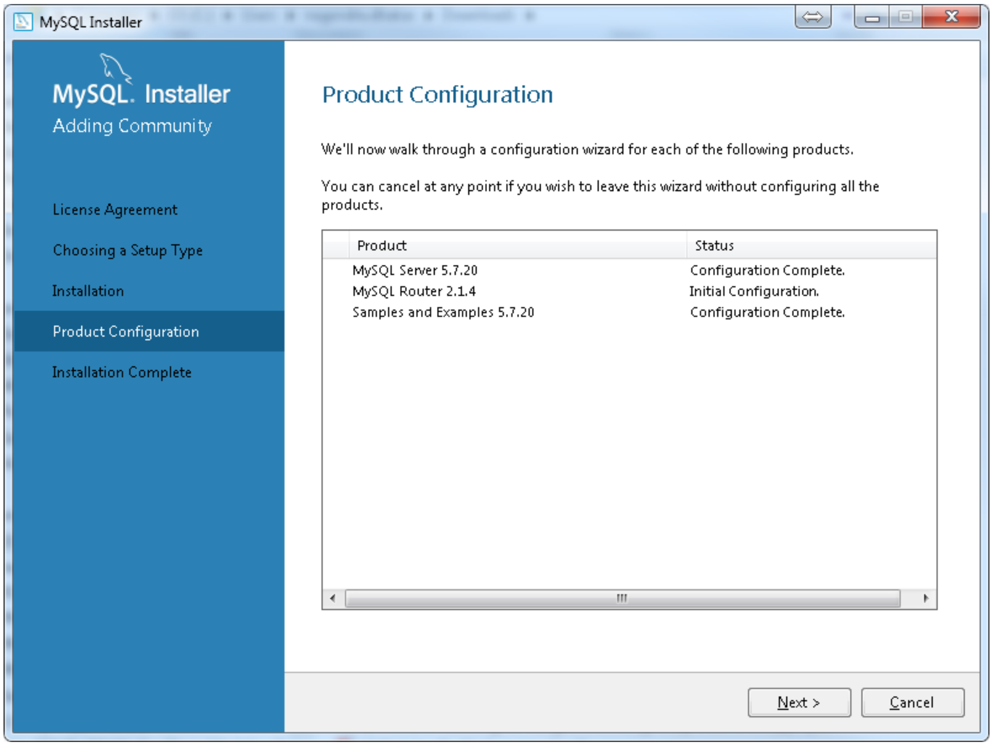
Next page, lists a set of configuration steps to make sure that the setup is working as needed. Click “Execute”.



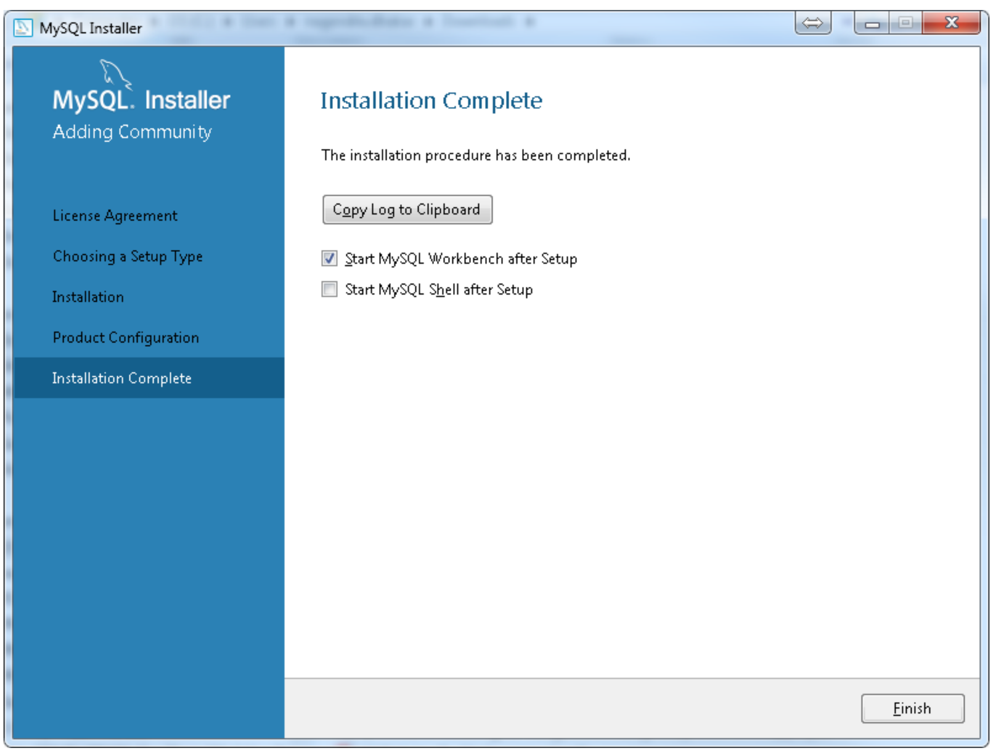
For next few seconds, the installer will run those steps.



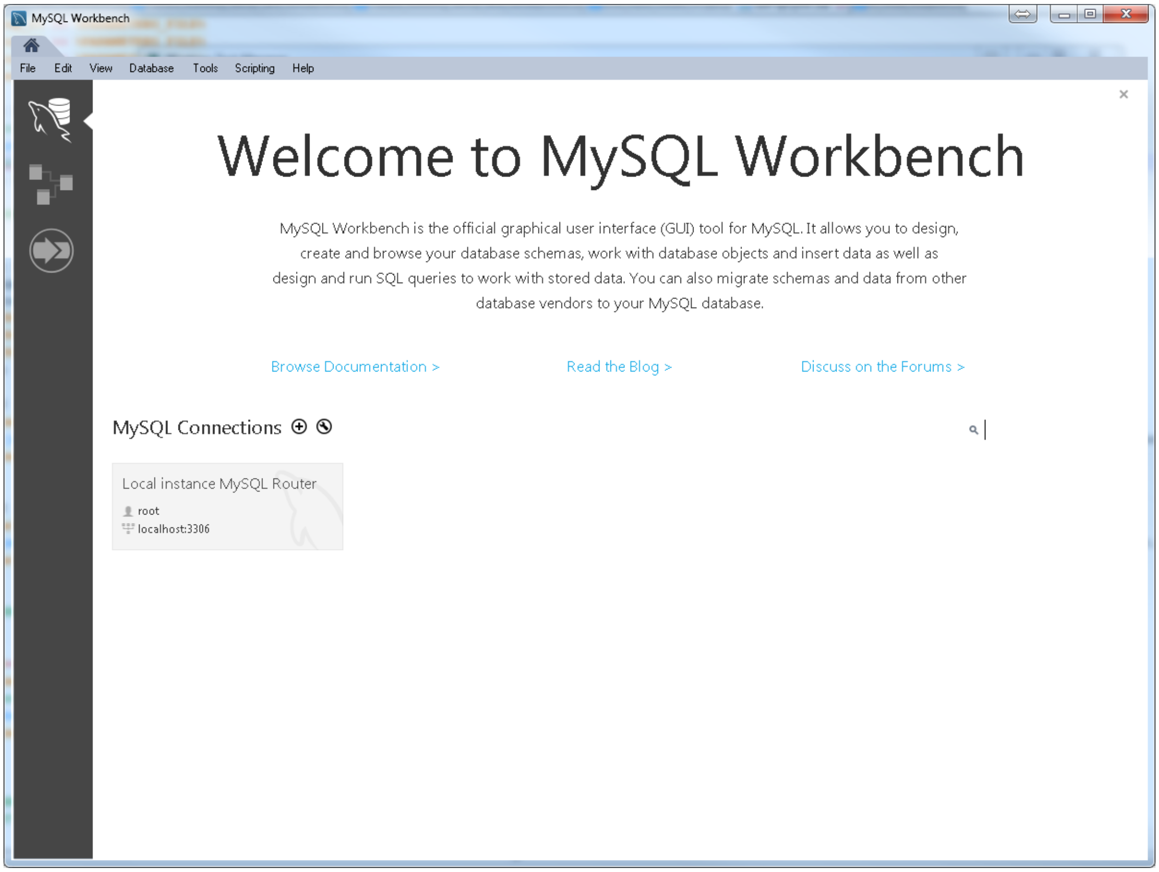
Once all the steps are applied successfully, a “Finish” command button will appear. Click “Finish”.



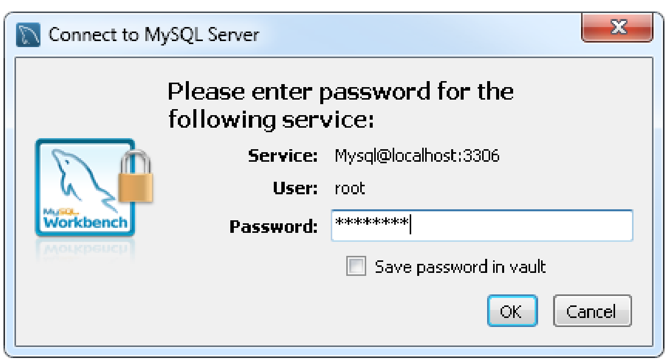
Next page, lists status of configuring different products. Click “Next”.



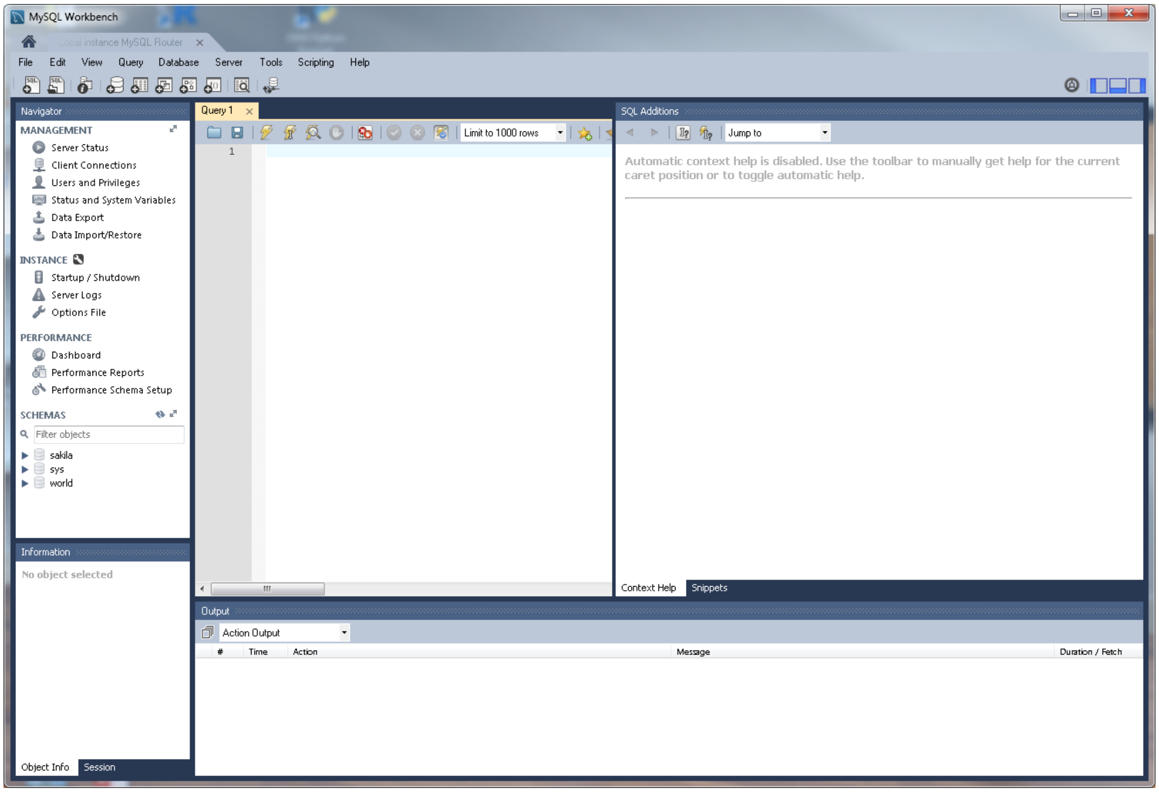
Next page will confirm completion of the MySQL installation. Uncheck “Start MySQL Shell after Setup” and keep “Start MySQL Workbench after Setup” checked. Click “Finish”.



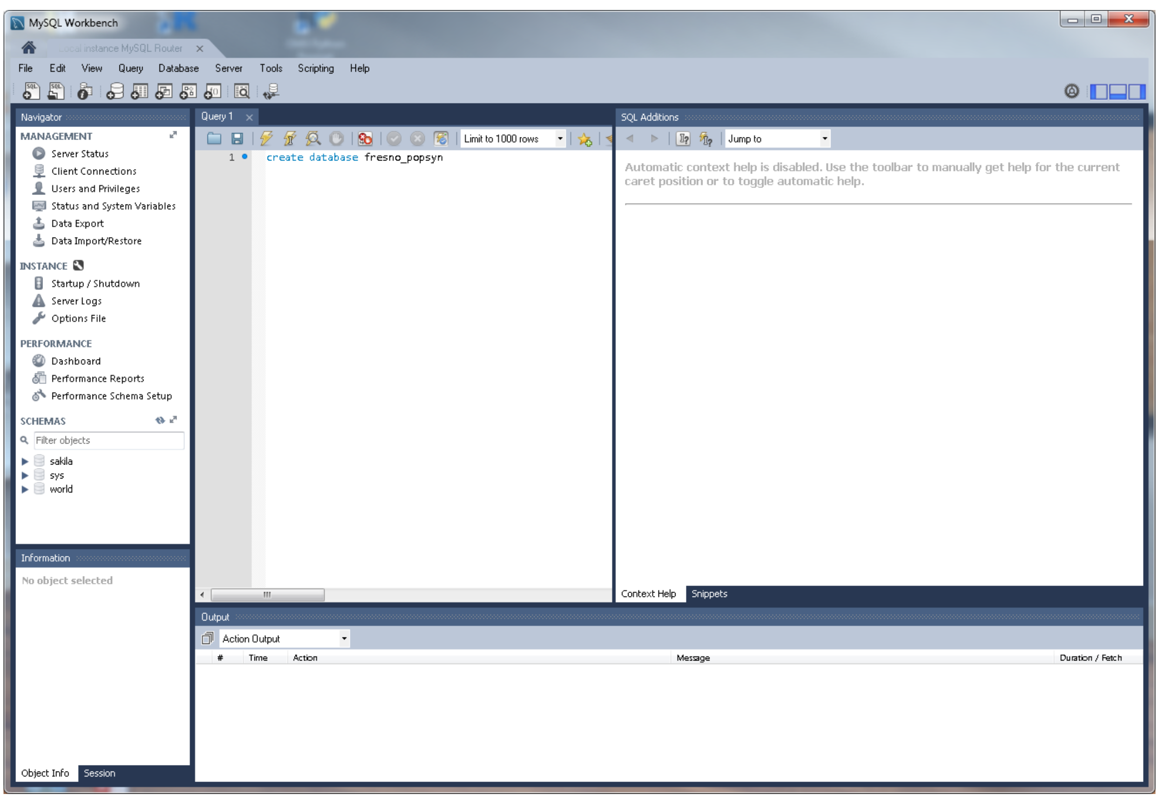
It will bring MySQL Workbench home page. You would see a local instance of MySQL Router. Double click on “Local Instance MySQL Router”. It will bring up a window asking for password for the root account.



Enter password to the root account and click “OK”. This will bring up a local instance window.

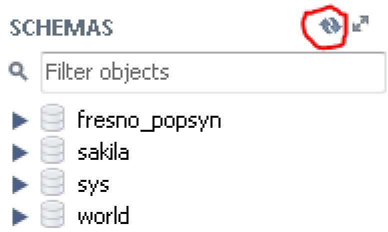


Now, we need to create a new database (“fresno_popsyn”) for popsyn.

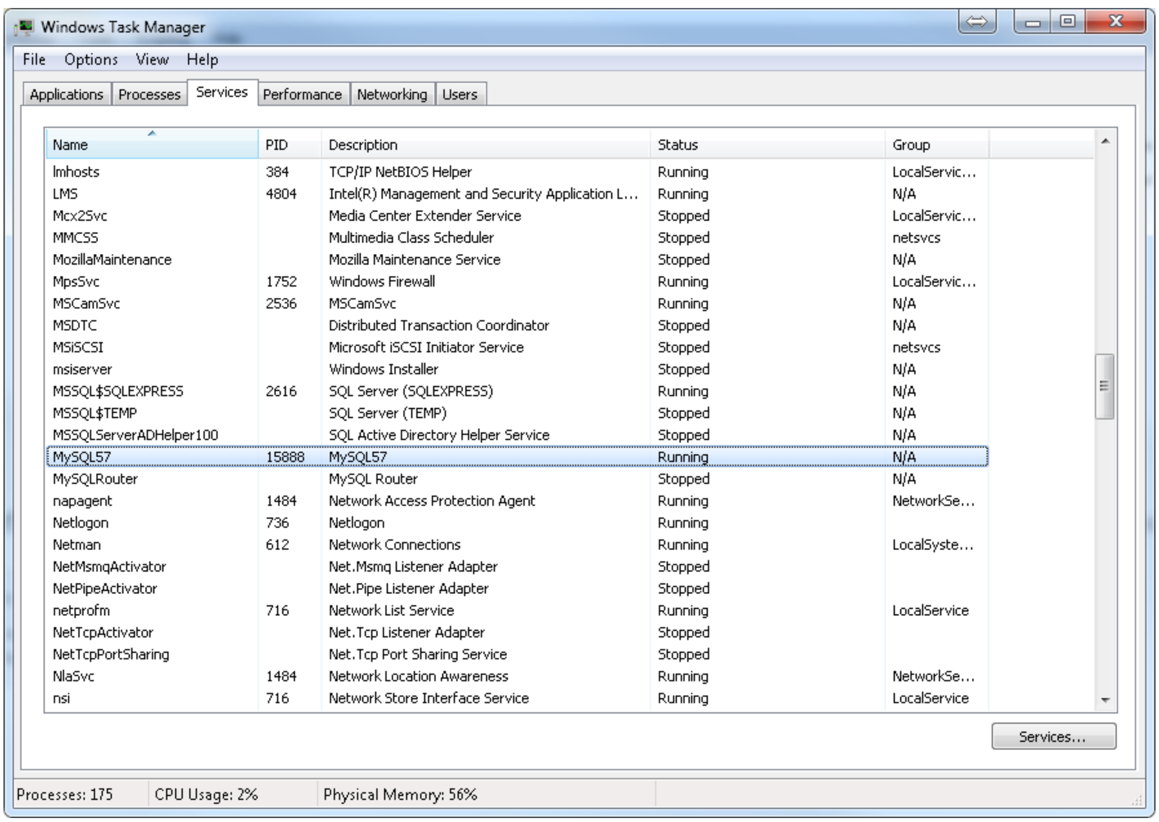


Write “create database fresno_popsyn” in the window “Query1” and click the thunder button on the ribbon (see below).

This will create a new empty database with name “fresno_popsyn”. Click refresh button for SCHEMAS to see the newly created database (see below).

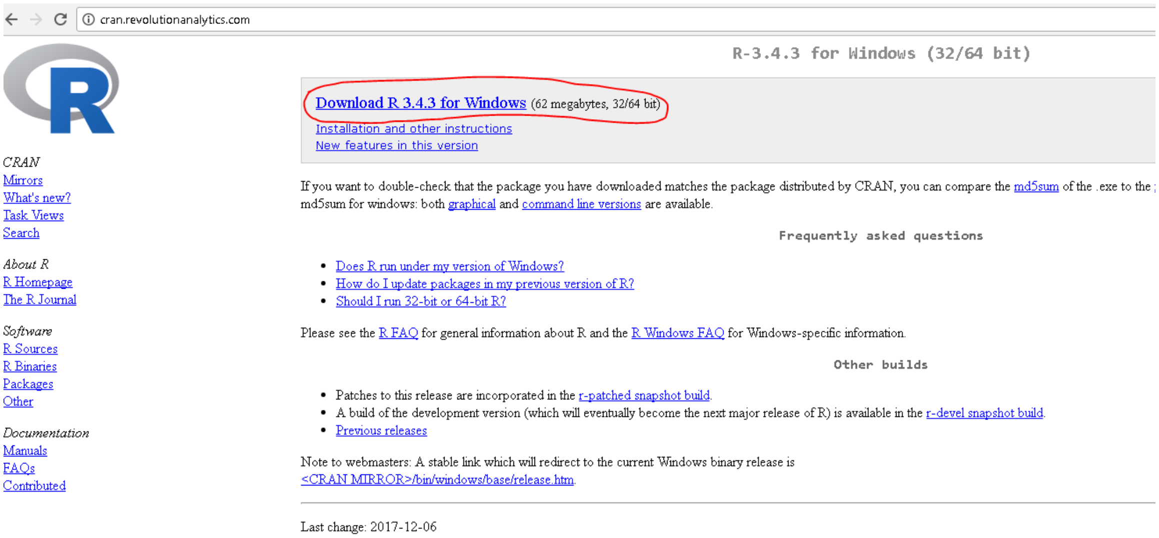


To confirm, open Windows Task Manager and go to “Services”. Confirm that “MySQL57” is running (Status should be “Running”). If status is “Stopped” then right click on the service and click “Start Service. This will change the status to “Running”.

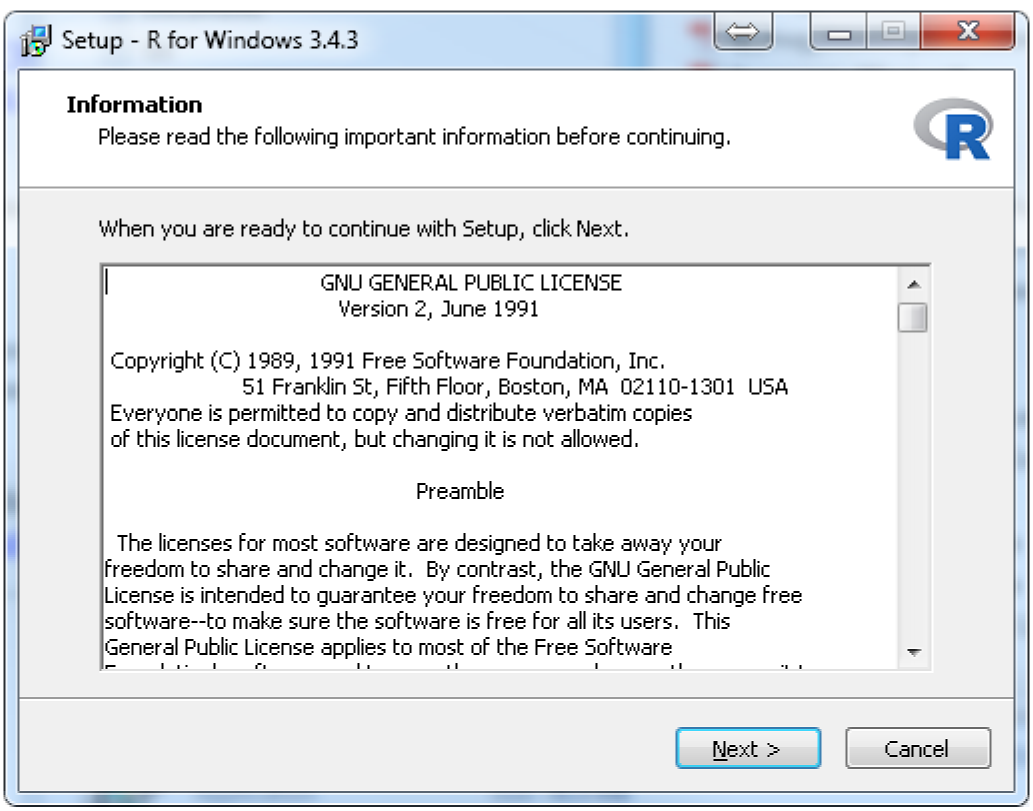


5.7 R

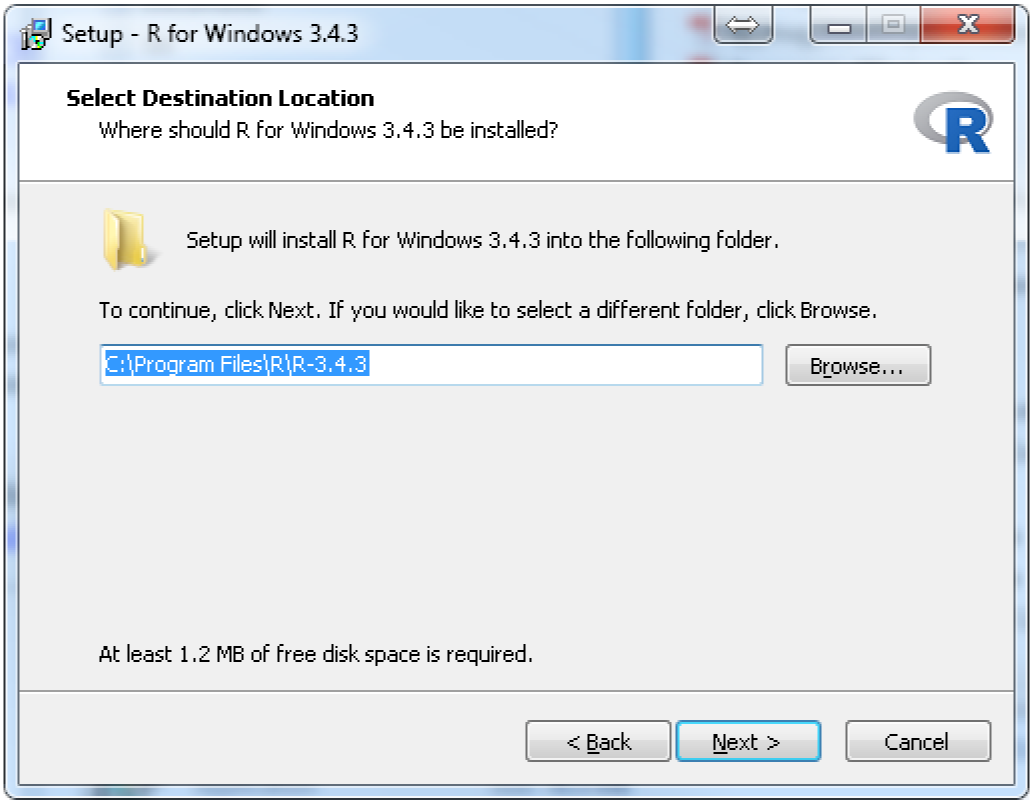
Go to the following page to download latest R (R 3.2.0 or later):
http://cran.revolutionanalytics.com/



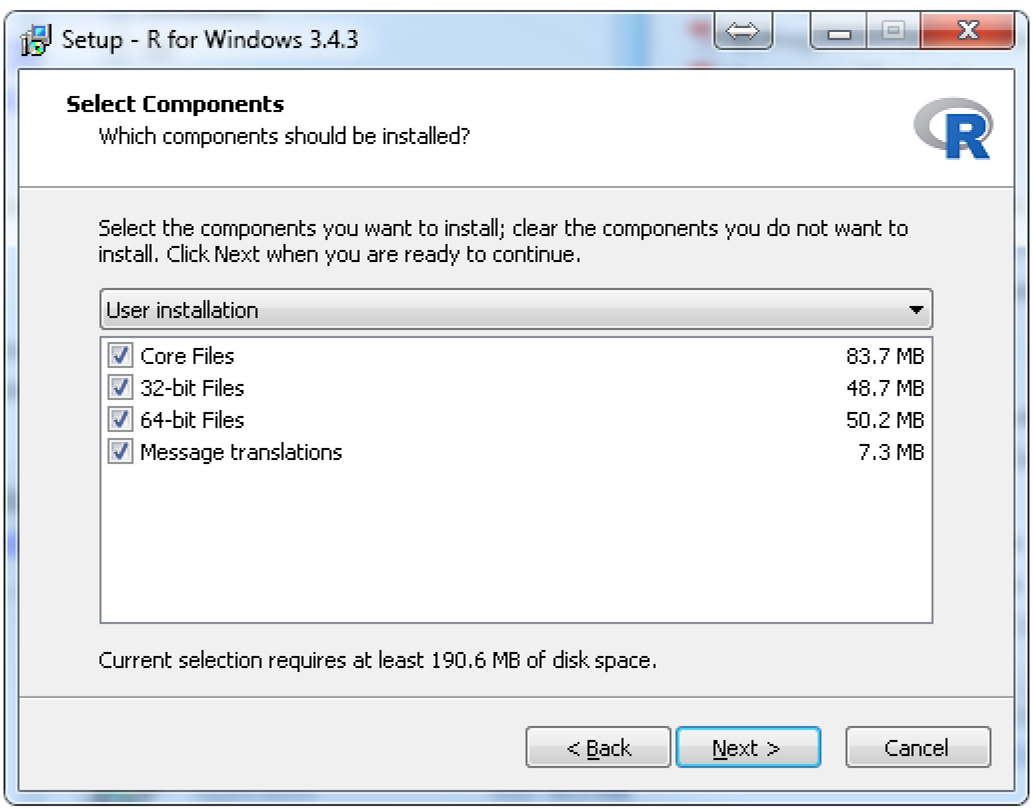
Click on “Download R.3.4.3 for Windows” and download the windows executable (R-3.4.3-win.exe) to a desired location on your machine. After the download is complete, go to the download folder and double click on the EXE file (R-3.4.3-win.exe).



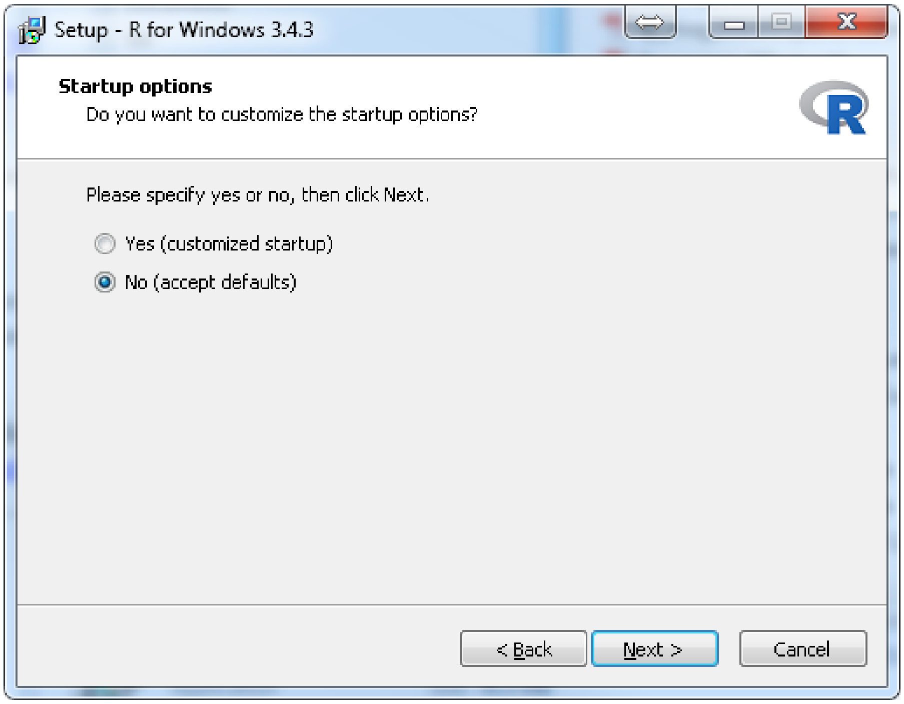
Click “Next”.



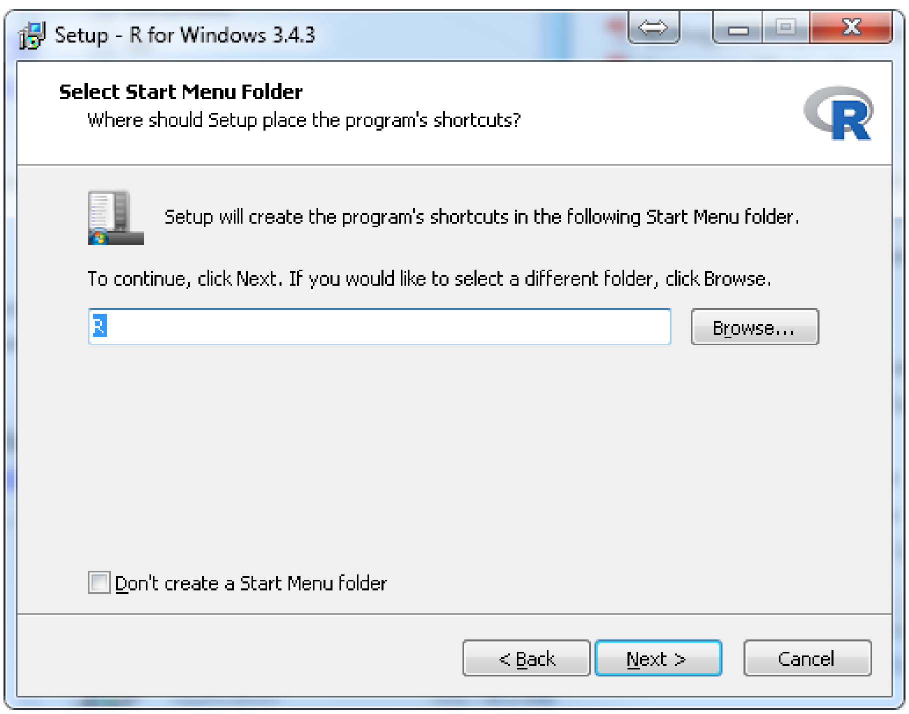
Choose a different folder if you want to install in a different location, otherwise keep the default and click “Next”.



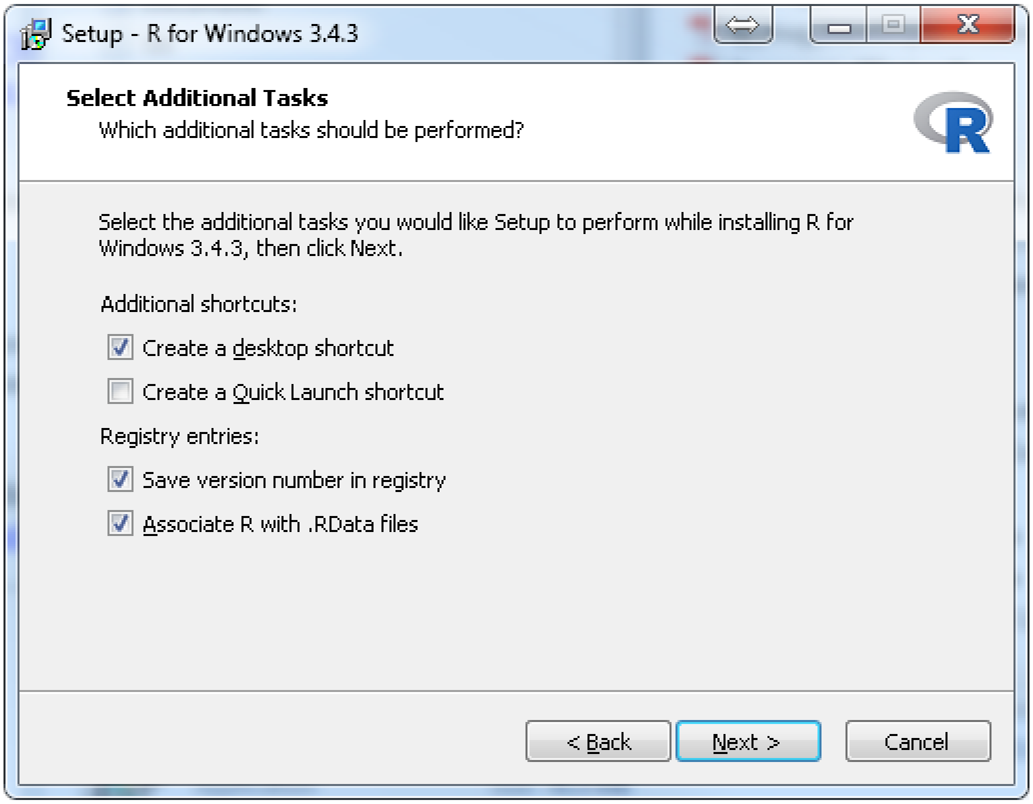
Keep the default options (as shown above) and click “Next”.



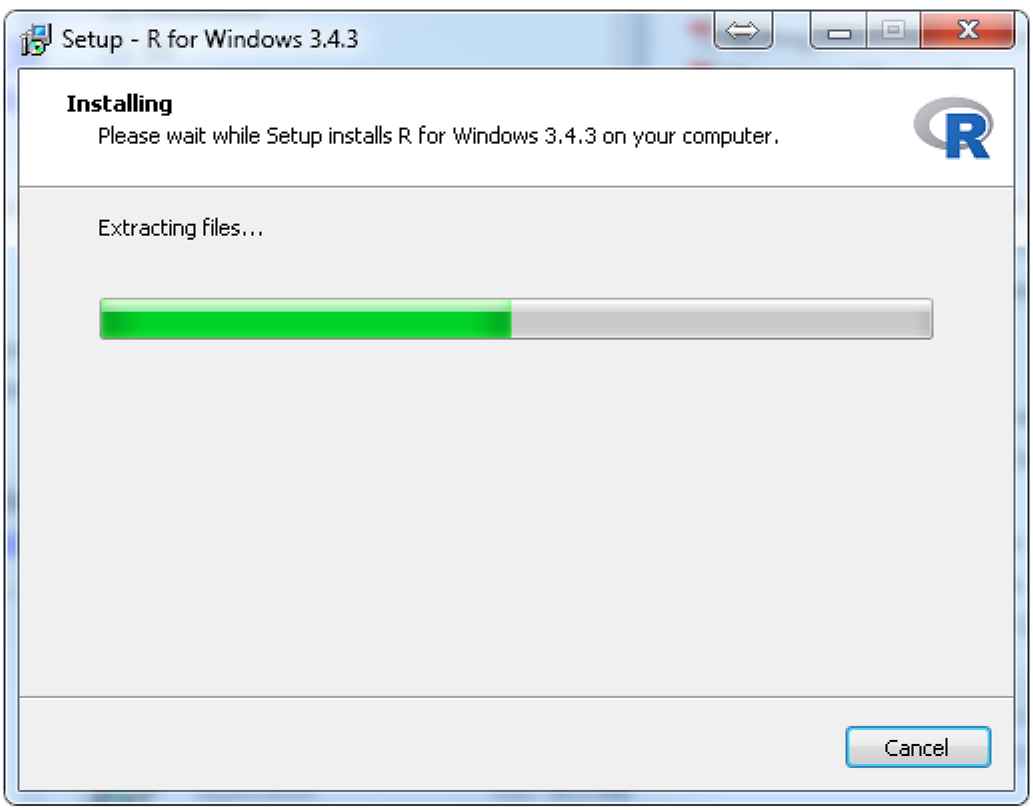
Accept defaults as startup options (check “No (accept defaults)”) and click “Next”.

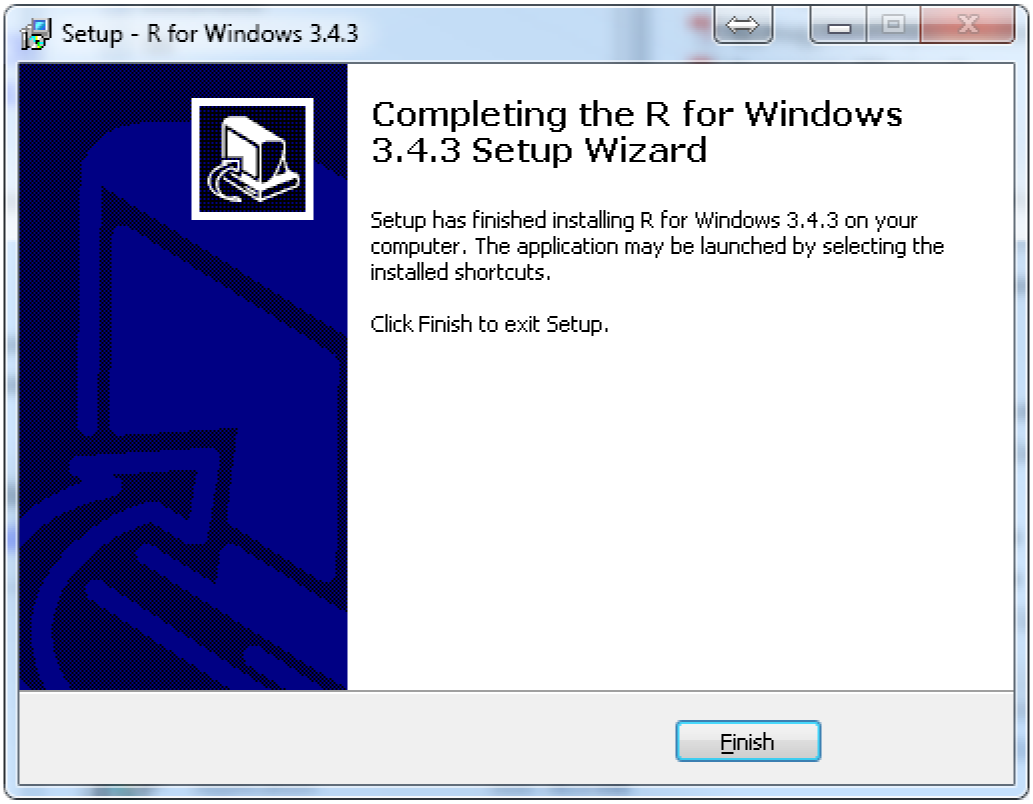


Choose a different folder if desired, otherwise keep the default and click “Next”.



Choose options as desired and click “Next”. This will start the installation.





Once setup is finished, click “Finish” to complete the installation.

SETUP POPSYNIII

Setup Directory

Figure 48 presents the directory structure for the PopSynIII setup. The *data* directory holds all the input data. The Census data that is downloaded from Web is also stored in an appropriate sub-directory within the data directory. The *runtime* directory houses the core PopSynIII JARs and associated libraries. The *scripts* directory has all the scripts called during various steps. All the outputs are exported to the *outputs* directory. The *Validation* folder contains all the configuration files required by the validation scripts. All the validation summaries and plots are outputted to this folder. The details of the validation procedure will be discussed later.

The *format_popsyn* folder contains python script that reads household and person files generated in the outputs folder and formats them into a format expected in DaySim. These formatted household and person files are input to DaySim.

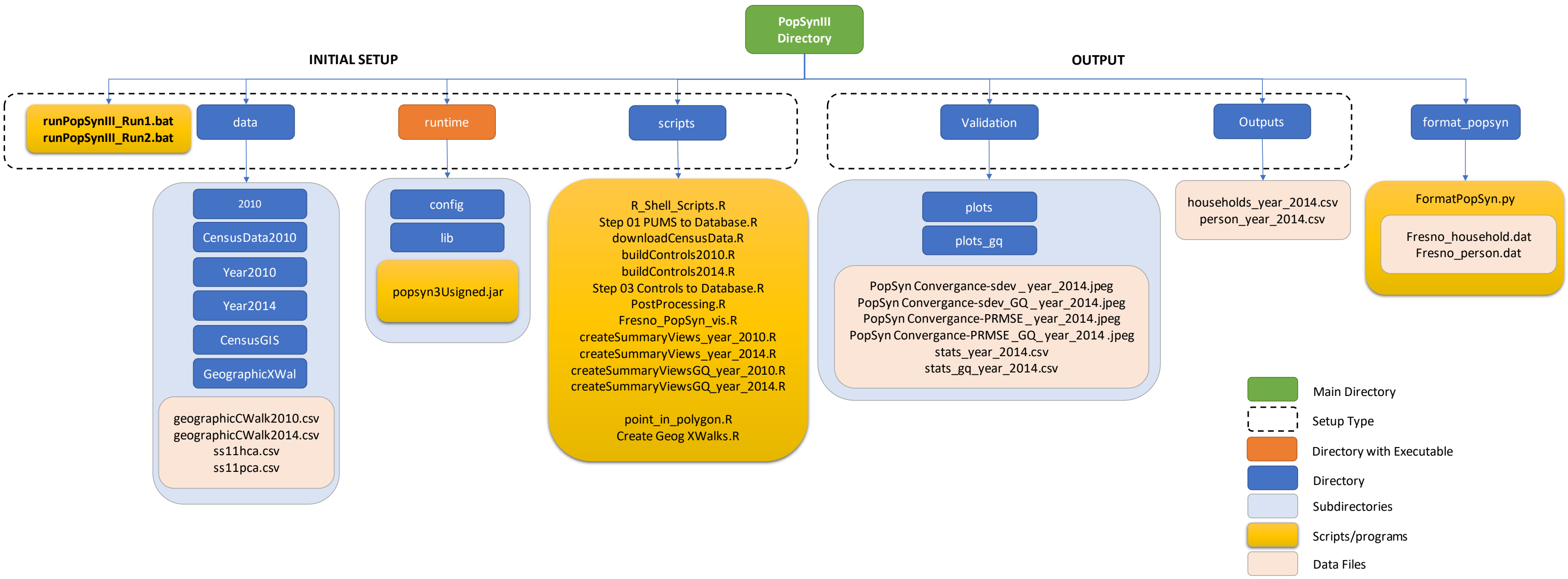


Figure 48: POPSYN III DIRECTORY STRUCTURE

System Setup

Instructions for setting up the system to run PopSyn III are included in the section "software requirements".

Model Setup

User need to update the following files:

- Main Batch Files (runPopSynIII_Run1.bat and runPopSynIII_Run2.bat)
- Settings File (settings_< POPSYN_Year >.xml and settingsGQ_<POPSYN_Year>.xml)
- MySQL password file (mysql.csv)

The following subsections provides more details about specific settings/parameters that need to be updated. First a brief description of the setting is provided, followed by a note (**<u>IMPORTANT</u>**) about what needs to be done.

Main Batch File (runPopSynIII_Run1.bat and runPopSynIII_Run2.bat)

These are the primary files that are used to run the PopSynIII software. Both batch files have same settings except mentioned.

The first group of settings are the general MySQL settings which include name of the MySQL server, name of the database, user name, etc. and path to the file which stores the passwords to each MySQL user account.

<u>IMPORTANT 1</u>: Here user needs to **update** MYSQL_PWD_FILE as per the setup.

```
SET WORKING_DIR=%CD%

:: MySQL
SET MYSQLSERVER=localhost
SET DATABASE=fresno_popsyn
SET DB_USER=root
SET MYSQL_PWD_FILE="C:\Projects\FresnoCOG\_tasks\PopSynIII\runtime\config\mysql.csv"
```

The next group relates to R which includes path to the R executable and names of R scripts for running different steps. `R_MAIN_SCRIPT` is the shell script which calls the other data processing scripts. `R_POST_PROCESS` is the post processing R script and `R_VALIDATION` is the R script which creates validation summaries. This setup expects that all the scripts are housed within the "\scripts" directory inside the PopSynIII working directory. Details of data processing using R scripts is discussed in the next section.

<u>IMPORTANT 2</u>: Here user needs to **update** RSCRIPT_64_PATH as per the setup.

```
:: R
SET RSCRIPT_64_PATH=C:\Progra~1\R\R-3.4.3\bin
SET R_MAIN_SCRIPT=%CD%\scripts\R_Shell_Script.R
SET R_POST_PROCESS=%CD%\scripts\postProcessing.R
SET R_VALIDATION=%CD%\scripts\Fresno_Popsyn_vis.R
```

The control building script downloads Census data from Web using Census API feature. The user needs to generate a Census API key from the Census website and store it in a CSV file. The `CENSUS_API_KEY_FILE` stores the path to the Census API key file. The file comes with the setup and is under "[working_directory]\runtime\config".

<u>IMPORTANT 3</u>: Here user needs to **update** CENSUS_API_KEY_FILE path to reflect the new setup.

```
:: Census API
SET CENSUS_API_KEY_FILE="C:\Projects\FresnoCOG\_tasks\PopSynIII\runtime\config\census_api_key.csv"

:: Parameters file [Do not change file name]
SET PARAMETERS_FILE="%CD%\runtime\config\parameters.csv"
```

The next set of parameters are synthetic population year (PopSyn_YEAR) and switches to control which steps to run and which ones to skip. The steps include `Run_HH_PopSyn`, `Run_GQ_PopSyn`, `Run_PostProcessing` and `Run_Validation`. The setup also produces synthetic non-institutional group quarters (GQ) population. This is achieved by running PopSynIII twice, first for the general population and then for the group quarter population.

The user can opt to run PopSynIII for either the general population (Run_HH_PopSyn), the GQ population (Run_HH_PopSyn), or both. The post processing (Run_PostProcessing) and validation scripts (Run_Validation) are run for each population group depending upon whether the user opted to run PopSynIII for that population group or not.

<u>IMPORTANT 4</u>: Here user **does not need to update** anything as these settings are already set in the two main batch files. **Do not change** the PARAMETERS_FILE.

```
:: SET PopSyn_YEAR as year_2010 or year_2015
:: -----------------------------------------------------------------
SET PopSyn_YEAR=year_2010

:: SET switch to run HH and GQ PopSyn runs [YES/NO]
:: -----------------------------------------------------------------
SET Run_HH_PopSyn=NO
SET Run_GQ_PopSyn=NO
SET Run_PostProcessing=NO
SET Run_Validation=NO
```

The initial data processing scripts build controls and seed samples for both the general and GQ population. Three switches are used to control the running of various data processing steps. Three settings are available: Run_Step_1 (PUMS to Database), Run_Step_2 (Build Controls), and Run_Step_3 (Controls to Database). The control building step has an option to download the Census data which is controlled by the downloadCensus. If user has run the setup before by setting downloadCensus=TRUE, then this setting can be set to FALSE as census data needs to be downloaded only once.

IMPORTANT 5: Here user **does not need to update** anything as these settings are already set in the two main batch files.

```
:: Switches [YES/NO] for running data processing steps
:: 1. Step 01 PUMS to Database, 2. Step 02 Build Controls, 3. Step 03 Controls to Database
:: -----------------------------------------------------------------
SET Run_Step_1=YES
SET Run_Step_2=YES
SET Run_Step_3=YES

:: Need to download Census data only once for building controls, can be read for next time [set to TRUE/FALSE]
:: 2010 controls should be build before building 2015 controls
SET downloadCensus=TRUE
```

All the input files are stamped by the population synthesis year. The setup expects a separate settings file for general and GQ runs and files should be named in the following format:

IMPORTANT 6: User **does not need to update** anything here.

```
:: PopSyn settings file
SET settingsFile="settings_%PopSyn_YEAR%.xml"
SET GQsettingsFile="settingsGQ_%PopSyn_YEAR%.xml"
```

The final group of settings in the runPopSynIII.bat file relates to calls to various scripts and PopSynIII core software.

IMPORTANT 7: Here user needs to **update** the Java path by setting the JAVA_64_PATH parameter.

```
:: ################################################################################
:: ### RUN POPSYN

ECHO %startTime%%Time%: Running population synthesizer...
SET JAVA_64_PATH=C:\Progra~1\Java\jdk-9.0.1
SET CLASSPATH=runtime\config
SET CLASSPATH=%CLASSPATH%;runtime\*
SET CLASSPATH=%CLASSPATH%;runtime\lib\*
SET CLASSPATH=%CLASSPATH%;runtime\lib\JPFF-3.2.2\JPPF-3.2.2-admin-ui\lib\*
SET LIBPATH=runtime\lib
```

The remaining settings/procedures in the batch file are automatically generated based on the settings discussed above.

Settings File (settings_< PopSyn_Year >.xml and settingsGQ_< PopSyn _Year>.xml)

To set up a PopSynIII run, the user must configure the XML settings file as well. User needs to review only the database settings that are beginning of the XML settings files. For the database, as described in Model Setup, the user would need to set up a MySQL server and create appropriate database (fresno_popsyn) to be used by PopSynIII.

A fully configured settings file for each population synthesis year has been added to the deliverable package. The user must update the settings file to add/drop controls from a PopSynIII run. There are separate settings file for general (household) and GQ runs and files are be named in the following format: settings_< PopSyn _Year >.xml and settingsGQ_< PopSyn _Year>.xml. Where, PopSyn _Year is a setting in run PopSyn III.bat which defines the run year (2010 or 2014). The files are under “runtime\config”.

IMPORTANT: User needs to update <password> (line 8) in the two settings file. Other database settings are standard.

```xml
<?xml version="1.0" encoding="UTF-8"?>

<targets>
  <database>
    <server>localhost</server>
    <type>MYSQL</type>
    <user>root</user>
    <password>nagendra</password>
    <dbName>fresno_popsyn</dbName>
    <useJppf>false</useJppf>
  </database>
```

MySQL Password File

The user needs to provide username and password to access MySQL database in a configuration file as well: “PopSynIII\runtime\config\mysql.csv”.

RUN POPSYNIII

To run PopSynIII, a batch file needs to be run using a command line. To do that, go to the working directory and enter the name of the batch file on the command line and hit enter.

```
Administrator: C:\Windows\system32\cmd.exe
C:\Projects\FresnoCOG\_tasks\PopSynIII>runPopSynIII.bat
```

To generate synthetic population for year 2014, user needs to run two batch files sequentially:

1. Run1 – For Year 2010 (runPopSynIII_Run1.bat)
2. Run 2 – For year 2014 (runPopSynIII_Run2.bat)

POPULATION SYNTHESIS PROCEDURE

The PopSynIII is configured to run using a batch file and the order of execution of different steps are as follows – PUMS data processing, building controls, data upload, PopSynIII run, post processing and finally, validation.

R & MySQL environment

The PopSynIII work with a MySQL database. In the existing setup, the main batch file makes system calls to various R scripts.

Data Processing in R

The data processing steps required to prepare inputs for PopSynIII in the right format are implemented via R scripts. Separate R scripts handle data processing after the PopSyn runs are completed and creates validation summaries and plots. The runPopSynIII.bat file makes system calls to run these R scripts. A brief description of each R script follows:

R_Shell_Script.R

Since each R script uses different R packages and inputs, a shell R script was written to install all the required libraries. This shell script also reads various user inputs and parameters from the main batch file. The main batch file writes all the user settings to a CSV file and then calls the shell R script. The shell R scripts reads the CSV parameters file and calls other R scripts as per the user choice selection. The parameters file (parameters.csv) is read by other scripts as well to read in user parameters and settings.

Step 01 PUMS to Database.R

This script is called by the R_Shell_Script.R script if the user sets the `Run_Step_1` setting to "YES". This script reads the 2007_11 ACS PUMS household and person sample and creates appropriate fields to be used by the PopSynIII run. The filenames of the input households and person sample can be changed in the R_Shell_Script.R script. Another input is the geographic crosswalk file which is used to get the list of relevant PUMAs in the modeling region. There is some hard-coded information in the scripts which relates to occupation categories etc. Separate samples are created for the general and GQ population. The final samples are uploaded to the MySQL data base specified in the batch file. Following tables are created in the MySQL database corresponding to the Census year (2007-2011):

1. household_table_200_2011, gqhousehold_table_2007_2011
2. person_table_2007_2011, gqperson_table_2007_2011

buildControls%Year%.R

There is a separate buildControls R script for each population synthesis year (2010 and 2015). The shell script calls the appropriates buildControls script as per user's population synthesis year choice if `Run_Step_2` was set to "YES". The objective of this script is to build MAZ, TAZ and County level controls. The main inputs are block group or tract level Census data and the existing MAZ and TAZ level data which mainly contains the number of households for each geography. The script expects all the Census data to be housed in the "/CensusData%Year%" sub-directory. Census data is downloaded from the web if the input files do not exist or if the user sets the `downloadCensus` setting to "YES". Census data is downloaded using the **downloadCensusData.R** script which creates the appropriate URL to fetch data via the Census

API. Census data which is not available for download via the Census API is downloaded once and is always read from the data directory.

Once the Census data has been downloaded at the block group and tract level, the next step is to allocate the control data and scale it to match the households at each of the relevant PopSynIII geographies. Block group level data is allocated to MAZs to create MAZ level controls and tract level data is allocated to TAZs to create TAZ level controls. MAZ level data can also aggregated to TAZ geography to create TAZ level controls. The allocations are done proportional to the number of households in each lower geography. The allocation step requires two geographic crosswalks: MAZ-block group and TAZ-tract. The crosswalk creation is described in the next sub-section. The allocated data was scaled to match the total number of households for each geography as specified in the existing MAZ and TAZ level data. PopSynIII requires integer controls; therefore, the allocated data was rounded off to the nearest integer. To ensure that the resulting distribution of households sums to the total households, the difference between scaled households and target households was added to the majority category for each household variables. Similar checks were made for the population variables as well.

Census data is not available for year 2014. For this year, the corresponding script reads the year 2010 controls file and scale them in proportion to the number of households in each geography as specified in the MAZ level data for those years. This requires that data processing for year 2010 be done before 2014.

Create Geog XWalks.R

As described in the previous section, this script creates the geographic crosswalks between MAZ, TAZ and Census 2000 and Census 2010 geographies. The script overlays the MAZ and TAZ centroids layer over the Block Groups and Census Tract polygon shape files. The MAZ and TAZ centroids were exported as a DBF file from the Fresno Cube highway network file. Next, a geographic file was created using the following projection - NAD 83, California Zone 6, Feet. The coordinates corresponding to MAZs and TAZs were copied to separate files to be read by the crosswalk creation script. The Longitude and Latitude values had to be divided by 1,000,000 to bring to the same format as Census projection (NAD 83). Census shape files block groups and tracts were downloaded for both Census 2000 and Census 2010 boundaries[21].

For creating geographic crosswalks, it was assumed that each MAZ lies within a block group and correspondence can be established if the MAZ centroid lies within a block group polygon. Similar assumptions were made for TAZ centroids and Census Tracts. It was observed that the TAZ and block group boundaries frequently intersect. Therefore, a clean correspondence cannot be established between block groups and TAZs. However, PUMAs fit perfectly within the

[21] https://www.census.gov/cgi-bin/geo/shapefiles/index.php

County boundaries inside the MTC modeling region. Therefore, two sets of crosswalks were produced for both Census 2000 and Census 2010 geographies:

1. MAZ – Block Group
2. TAZ – Tract – PUMA – County

Step 03 Controls to Database.R

This script is run when the `Run_Step_3` setting is "YES" in the runPopSynIII.bat file. This script attaches geographic information to the MAZ level control and upload all the control files to MySQL database. This script also takes the PopSynIII year as an input and uploads MAZ, TAZ and County controls only for that year. This script also reads the MAZ level GQ data to create controls for the GQ PopSynIII run. It was found that there are very few military and university student GQ records in many PUMAs, which causes problems for the balancing and drawing algorithms in PopSynIII. To address this problem, all the PUMA within a County were combined and the meta geography was defined as the whole modeling region for the purposes of GQ synthetic population generation.

The following tables are uploaded to the MySQL database by this script:

1. control_totals_maz_year_%Year%
2. control_totals_taz_year_%Year%
3. control_totals_meta_year_%Year%
4. control_totals_maz_gq_year_%Year%
5. control_totals_meta_gq_year_%Year%

postprocessing.R

This script is called directly by the runPopSynIII.bat file at the completion of the PopSynIII runs. The objective of this script is to download the final PopSynIII outputs from the MySQL database, expand the data and write Comma Separated Value (CSV) household and person files that can be read by CT-RAMP. The PopSynIII outputs are expanded based on the final weight that was assigned to that record such that, each record in the final household file represent a single household, and each record in the person file represent a single person. A unique household identifier is added to both the household and person file. These final output files are exported as CSVs to the "/outputs" sub-directory in the working directory. If the user opted to run PopSynIII for both general and GQ population, the post processing combines the household and persons records from both the runs into a single file.

The remaining R scripts relates to creation of validation summaries and plots and will be discussed in next section.

Validation Procedures

One of the most critical steps in population synthesis is validating the final synthetic population. Validation can give us clues about inconsistencies among controls, data processing errors or misspecification of any settings. The validation procedure implemented for Fresno PopSynIII produces a validation summary, validation chart, frequency plots and expansion factor distribution. Each of these are described briefly below:

Validation summary statistics

At a regional level, for each control, total number of records (household/person) desired by the control, the total number of records synthesized, the difference between the synthesized totals and the control totals and the percentage difference are reported.

Statistics that inform us of convergence at a more disaggregate level are also computed – please note that these statistics are being computed for the geography at which the controls are specified i.e. MAZ, TAZ or Meta as the case might be. The following three statistics are computed as a part of this exercise:

(1) the average percentage difference between the control totals and the synthesized totals,

(2) the standard deviation of the percentage difference – this measure informs us of how much dispersion from the average exists, and

(3) the percentage root mean square error (RMSE) - an indicator of the proximity of synthesized and control totals.

The number of geographies for which the control is non-zero (N) are also reported.

Validation Chart

The validation chart is a visualization of the disaggregate summary statistics – mean percentage difference, STDEV and RMSE of percentage differences. 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.

Frequency Distribution Plots

These are simply frequency distribution plots of differences between control and synthesized values across the geography at which the controls were specified.

Expansion factor Distribution

While a synthetic population may match the controls well, it is important to know how uniform the household weights are, and how different they are from the initial weights. The closer the final weights are to the initial PUMS weight, higher is the chance of matching the distribution of uncontrolled variables. An expansion factor is computed for each record in the PUMS data as total final weight/initial weight. A distribution plot of these expansion factors is created for each PUMA. A good synthetic population would have most of these expansion factors as close to one as possible.

Validation R Scripts

The validation procedures were also written in R. The first step involves creation of summary views in MySQL for each geography for the controls specified at that geography. Depending on the year that was specified in the runPopSynIII.bat file and if options to run general/GQ/both PopSynIII runs were selected, appropriate **createSummaryViews_year_%Year%.R** or **createSummaryViewsGQ_year_%Year%.R** script is called. Next, the **Fresno_Popsyn_vis.R** scripts reads in the appropriate summary views and control tables from the MySQL database and creates all the summaries and plots in the "/Validation" directory. The **Fresno_Popsyn_vis.R** script also reads columnMap.csv as input from the "/Validation" directory. This specifies the controls which should be included in the validation process and their field names in summary views and controls tables.

5.8 VALIDATION

This section first presents the performance of the existing PopSynIII setup and validation of the recommended controls for 2010. A base year (year 2014) PopSyn run was completed with the initial setup, which included the following controls at various geographies:

1. MAZ
    a. Total number of households
    b. Number of household by household size groups [1, 2, 3, 4+]
2. TAZ
    a. Number of households by presence of children
    b. Number of households by number of workers [0, 1, 2, 3+]
    c. Number of households by income groups [0-$30K, $30K-$60K, $60K-$100K, $100K+]
    d. Number of persons by age groups [0-18, 19-35, 36-64, 65+]
3. County
    a. Number of persons by occupation

Table 84 describes the source of data for each control. Data available at block groups was used to construct MAZ level controls and data available at tract level was used to build TAZ level controls.

Table 84: Control Data source description - 2010

CONTROL NAME	CONTROL GEOGRAPHY	DATA SOURCE	TABLES/VARIABLES
Total households	MAZ	Fresno MAZ Data	
HH Size 1	MAZ	2010 SF1 Block Group	["H0130001" to "H0130008"]
HH Size 2	MAZ	2010 SF1 Block Group	["H0130001" to "H0130008"]
HH Size 3	MAZ	2010 SF1 Block Group	["H0130001" to "H0130008"]
HH Size 4+	MAZ	2010 SF1 Block Group	["H0130001" to "H0130008"]
HH Income (-Inf, $30K)	TAZ	2011-15 ACS Tract	["B19001_001E" to "B19001_017E"]
HH Income [$30K, $60K)	TAZ	2011-15 ACS Tract	["B19001_001E" to "B19001_017E"]
HH Income [$60K, $100K)	TAZ	2011-15 ACS Tract	["B19001_001E" to "B19001_017E"]
HH Income [$100K, +Inf)	TAZ	2011-15 ACS Tract	["B19001_001E" to "B19001_017E"]
HH Workers 0	TAZ	2011-15 ACS Tract	["B08202_001E" to "B08202_005E"]
HH Workers 1	TAZ	2011-15 ACS Tract	["B08202_001E" to "B08202_005E"]
HH Workers 2	TAZ	2011-15 ACS Tract	["B08202_001E" to "B08202_005E"]

HH Workers 3+	TAZ	2011-15 ACS Tract	["B08202_001E" to "B08202_005E"]
HH w/ children	TAZ	2010 SF1 Tract	["PCT0160001","PCT0160004", "PCT0160010", "PCT0160016"]
HH w/o children	TAZ	2010 SF1 Tract	["PCT0160001","PCT0160004", "PCT0160010", "PCT0160016"]
Population age 0 to 19 years	TAZ	2010 SF1 Block Group	["P0120001" to "P0120049"]
Population age 20 to 34 years	TAZ	2010 SF1 Block Group	["P0120001" to "P0120049"]
Population age 35 to 64 years	TAZ	2010 SF1 Block Group	["P0120001" to "P0120049"]
Population age 65plus years	TAZ	2010 SF1 Block Group	["P0120001" to "P0120049"]
Occupation: Management	County	2007-11 ACS PUMS	[socp00, socp10]
Occupation: Professional	County	2007-11 ACS PUMS	[socp00, socp10]
Occupation: Services	County	2007-11 ACS PUMS	[socp00, socp10]
Occupation: Retail	County	2007-11 ACS PUMS	[socp00, socp10]
Occupation: Manual	County	2007-11 ACS PUMS	[socp00, socp10]
Occupation: Military	County	2007-11 ACS PUMS	[socp00, socp10]

We address the inconsistencies among the controls by appropriately scaling the controls to match the total number of households to the MAZ level total number of households control for each geography.

The validation results from the run for 2014 is presented in Figure 49.

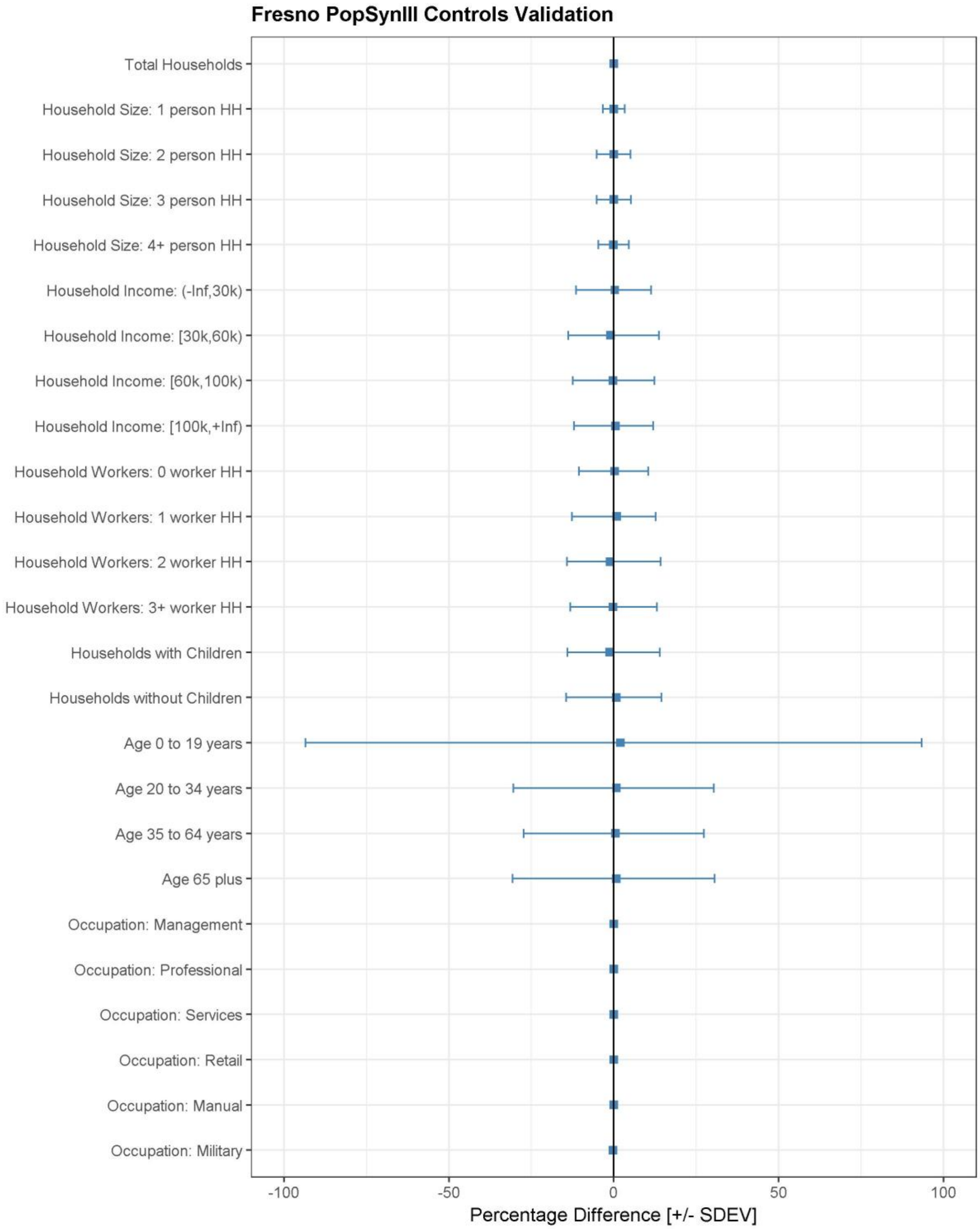


Figure 49: POPSYNIII VALIDATION [2014]

The validation results indicate that PopSynIII does reasonably well in matching the controls overall, as can be observed by the close to zero mean percentage differences across all controls. The standard deviation is also quite low for all the controls. However, some deviation can be observed for the person age group controls.

RUN POPSYNIII FOR A FUTURE YEAR

This section provides instructions for running the current PopSynIII setup to generate synthetic population for a future year. Here are sequential steps:

- Run PopSynIII Run 1
- Update 2014 Inputs
- Update Control Distribution
- Run PopSynIII Run 2

The below are more detailed instructions for each of the four steps:

Run PopSynIII Run 1

First, user needs to run the first main batch file – “runPopSynIII_Run1.bat”. This will generate controls for year 2010.

Update 2014 Inputs

Now, user needs to update following files for households and group quarter population for the future year:

- PopSynIII\year_2014\mazData_2014.csv
- PopSynIII\year_2014\gq_maz_2014.csv
- PopSynIII\year_2014\meta_controls_gq.csv

Update Control Distribution

User can either use the current census distribution or provide forecasted distribution of household/people in different categories.

- If user chooses to use the current Census distribution, then no action is required here – just skip this step.
- If user chooses to provide forecasted distribution (for one or multiple socio-demographics), then depending on the geography of the available distribution that can be done either in the census data or in the 2010 controls (built by running “runPopSynIII_Run1.bat”).

- The census data contains distributions at block and tract level. If distribution in the census data are updated then user needs to re-run PopSynIII Run 1 as well. The downloaded census data is here: PopSynIII\data\CensusData2010\Downloads
- The 2010 controls are at maz, taz and country level. Table 84 provides a list of controls and corresponding census data and control geography. The 2010 controls are here: PopSynIII\data\year_2010.

Run PopSynIII Run 2

Before running the second main batch file (runPopSynIII_Run2.bat), user needs to set maximum expansion factor to a higher number. This needs to be done in "PopSynIII\runtime\config\settings_year_2014.xml" and "PopSynIII\runtime\config\settingsGQ_year_2014.xml".

```
<pumsData>
    <idField>unique_hh_id</idField>
    <pumaFieldName>PUMA</pumaFieldName>
    <metaFieldName>mtc_county_id</metaFieldName>
    <tazFieldName>taz</tazFieldName>
    <mazFieldName>maz</mazFieldName>
    <weightField>WGTP</weightField>
    <hhTable>household_table_2007_2011</hhTable>
    <persTable>person_table_2007_2011</persTable>
    <pumsHhTable>household_table_2007_2011</pumsHhTable>
    <pumsHhIdField>unique_hh_id</pumsHhIdField>
    <pumsPersTable>person_table_2007_2011</pumsPersTable>
    <maxExpansionFactor>5</maxExpansionFactor>
    <synpopOutputHhTableName>synpop_hh_2015</synpopOutputHhTableName>
    <synpopOutputPersTableName>synpop_person_2015</synpopOutputPersTableName>
    <outputHhAttributes>serialno, np, hincp, ten, bld, hh_workers_from_esr, hh
    <outputPersAttributes>sporder, agep, sex, occupation, wkhp, esr, schg, emp
</pumsData>
```

Now, user needs to run the second batch file "runPopSynIII_Run2.bat" and generate synthetic population under "outputs".

Note that the generated synthetic population files need to be converted into DaySim format using "FormatPopSyn.py"

APPENDIX B. GENERATE EXTERNAL AUTO TRIP TABLES

Inputs and scripts used to generate external auto trip tables for the base year are included under “external_autotrips” in “1_Inputs\Support”. Note that for a successful completion, the statewide model should be run on Cube 6.1.0.

The following are sequential steps to generate external auto trip tables for a scenario year:

1. Obtain statewide model for the scenario year. This process runs highway assignments, therefore the files needed for running the final assignment are required.
2. Create Fresno subarea network for four time periods from statewide model networks by time period. Cube will re-number zones (external and internal) in the subarea networks. Refer to “Extracting subarea highway network” in Cube help. Here are some steps to create sub-area network for a time-period (Note that sub-area networks for other time periods can be created using step “b” through step “f”. Step “a” is needed only once and that is for the first sub-area network creation.):
   a. Load “Create_SubAreaNetwork_AM_2015.VPR” in Cube.
   b. Go to “Layer Control” next to the layer drop down box and change the highway file by double clicking on the HWY file. Set it to statewide network file on your machine.
   c. Go to “Drawing Layer” and click on “Restore” under “Edit Polygon”. Choose “0=fresno_county”. This will make “Sub-Area Extraction” tool active under “Polygon Tools”.
   d. Click on “Sub-Area Extraction” tool and provide a name for output sub-area network file.
   e. Cube will open a “Sub-Area Extraction Node Renumbering” dialogue providing new ranges of zones (internal and external) and nodes in the output network. Don’t need to do anything here, just click “OK”.
   f. This will produce a sub-area network at the location provided in step “d”.
3. Run highway assignment with subarea analysis for four time periods. See “ASHWY00A_SubArea_[tod].S” to run assignments by time periods for base year. You would need to enter VOT=12 and HOVAutoOcc=3.36 in a run dialogue box. This will produce four OD matrices for Fresno sub-area. The OD matrices contain hourly demand for Fresno sub-area with zones that are different than the Fresno ABM network zones.

4. Disaggregate the OD matrices into Fresno zones and convert the hourly OD matrices into time-period matrices.
    a. First, create two zone equivalency text files (zone_eq_am_md.txt and zone_eq_pm_ev.txt) using spreadsheet Fresno_CSTDM_Zone.xlsx.
    b. Then run the following Cube scripts disaggregateToFresno\CSTDM_CreateFresnoMatrix_AMMD.S and CSTDM_CreateFresnoMatrix_PMNT.S. The scripts will create six OD demand matrices (four time periods and two peak hours). Before running, update file paths in the scripts.
5. Adjust the OD matrices to match external stations. See AdjustExtStations\CSTDM_FresnoMatrix_AdjExtSta.s for base year. Adjustment factors would be the same as base year as no future year count information at external stations are available. Before running, update file paths in the script.
6. In all, six output matrices will be created (four time periods and two peak hours). Copy the outputs to the inputs folder here: 1_Inputs\5_External.

APPENDIX C. PREPARE DAYSIM INPUTS

This appendix provides intersection to create some of the DaySim inputs files. The following inputs are included: input micro-zone file, transit stops file, and intersection file.

INPUT MICRO-ZONE FILE

The input MAZ file, "1_Inputs\8_DaySim\02_Parcel\maz_2014_parks.csv", provides landuse data input to DaySim. The input file is created from a shapefile provided by FresnoCOG.

Table 85 presents a list of fields that are needed in the MAZ shapefile to be exported into the MAZ input file. It also describes the action required for a field, whether you need to add it as a new field or it already exists, and the field data type. If a new field is required, the corresponding field in the shapefile is provided.

TABLE 85. MAZ SHAPEFILE FIELDS

DAYSIM FIELD	ACTION	FIELD TYPE	MAZ SHAPEFILE
Parcelid	Add field	Long	OBJECTID
xcoord_p	Add field	Double	Calculate geometry for X coordinate (in feet)
ycoord_p	Add field	Double	Calculate geometry for Y coordinate (in feet)
sqft_p	Add field	Double	Calculate geometry for polygon area (in sq. feet)
taz_p	Add field	Long	TAZ
hh_p	Add field	Long	HH
block_p	Add field	Double	Set to 0
stugrd_p	Add field	Long	ELEM
stuhgh_p	Add field	Long	HS
stuuni_p	Add field	Long	COLLEGE

empedu_p	Add field	Double	EMP_EDU
empfoo_p	Add field	Double	EMP_FOO
empgov_p	Add field	Double	EMP_GOV
empind_p	Add field	Double	EMP_IND
empmed_p	Add field	Double	EMP_MED
empofc_p	Add field	Double	EMP_OFC
empret_p	Add field	Double	EMP_RET
empsvc_p	Add field	Double	EMP_OTH
empoth_p	Add field	Double	EMP_AGR
emptot_p	Add field	Double	EMP
parkdy_p	Exist		
parkhr_p	Exist		
Ppricdyp	Exist		
Pprichrp	Exist		

Once these fields are added and populated in the shapefile:

- Go to shapefile properties and keep only these fields (see Table 85) and uncheck all other fields in the shapefile.
- Export the attribute table into a CSV file.
- Open the exported CSV file and make sure that all numbers are general numbers – look out for commas (,) in numbers. The CSV files should not have any formatted numbers so no commas should be present.
- Now need to add park sq feet into the maz file. The park square feet are stored in the field “block_p”. This field is populated using the base year MAZ file. This can be done either by doing a simple look up by “parcelid” from the base year data or sort the two

files by “parcelid” and copy column “block_p” from the base year file and paste into the new CSV file under the same column name.

- In the end, make sure that the fields are in the order listed in Table 85.

This will create the input MAZ file. Set the file name as “maz_2014_parks.csv” and place it here: “1_Inputs\8_DaySim\02_Parcel”.

Transit Stops File

The transit stops file is created using the input transit network: 1_Inputs\3_Highway\FresnoNetworks_Deliverable3.gdb\PTNetwork_14

Follow the sequential step below:

- Export the nodes (PTNetwork_14_PTNode) to a shapefile here:
- Open the shapefile in ArcMap and select the nodes that have STOPNODE=1. These are the transit stops.
- Export the selected nodes (stops) to a new stops shapefile
- Open the attribute table of the new stops file and add four fields: xcoord_p (float), ycoord_p (float), mode (int) and id (long).
- Calculate xcoord_p and ycoord_p using calculate geometry for the two fields.
- Join transit line by object id and calculate a new field “mode” equal to the MODE field in the line file (1-localbus, 2-exp bus, 4-brt)
- calculate stopid as “FID+1”.
- Keep only four fields in the attribute table: id, mode, x_coord, y_coord.
- Export all records in the attribute table to a csv file: transit_stops_2014.csv

The stops file is in the following format:

id	mode	xcoord_p	ycoord_p
1	1	6309142	2167872
2	1	6309117	2166434
3	1	6309112	2165240
4	1	6307677	2165242
5	1	6306517	2165245
6	1	6306529	2166560
7	1	6306503	2167889
8	1	6309149	2169329
9	1	6309154	2170526
10	1	6309171	2171830
11	1	6307831	2173200
12	1	6306543	2173215
13	1	6305216	2173227
14	1	6303896	2173225
15	1	6303887	2174830

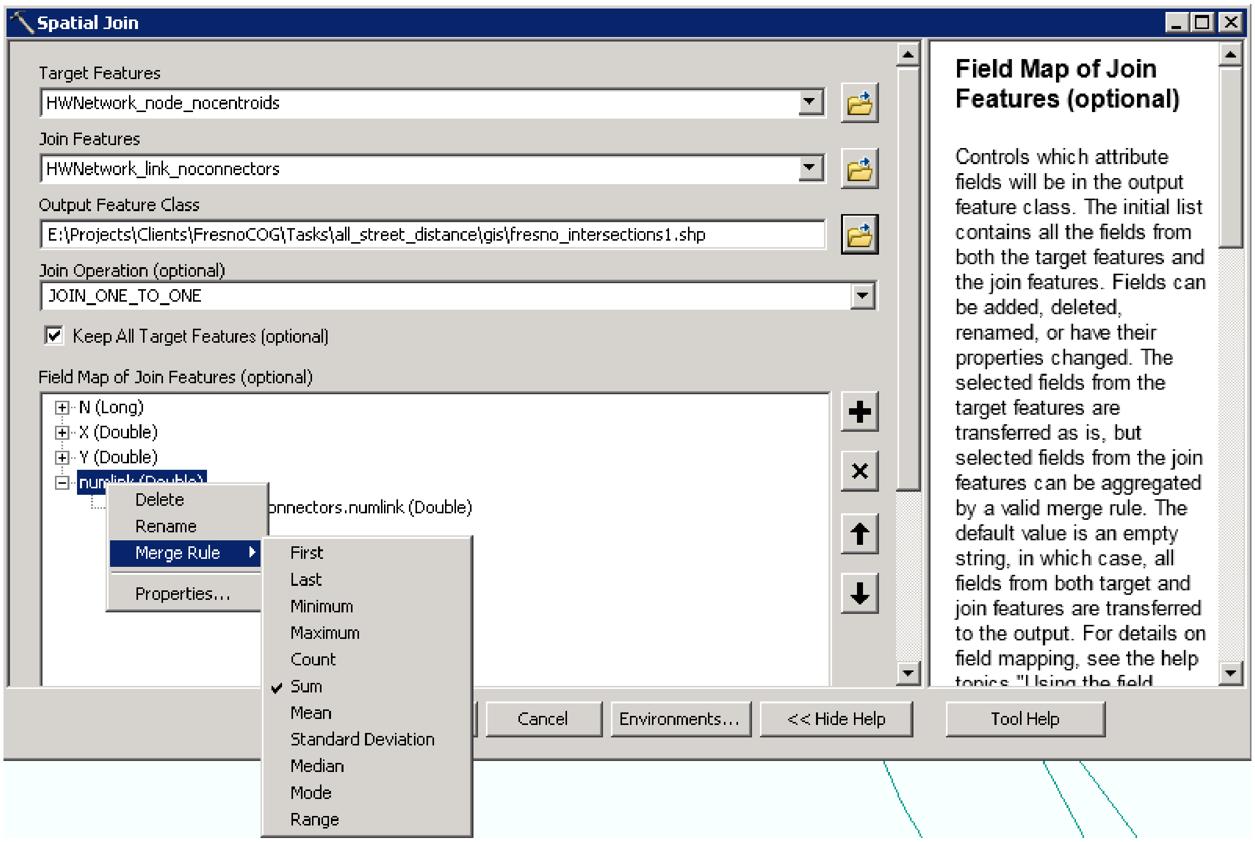
INTERSECTIONS FILE

The intersection file provides number of roads (links) at an intersection (node). The all-street network in the model is a dual road network where a two-way street is represented as one link for each direction. A two-way road is represented by two links representing to and from direction. Therefore, this network cannot simply be used to calculate number of links at an intersection. So, first we need to simplify the network and then process to create the intersection file.

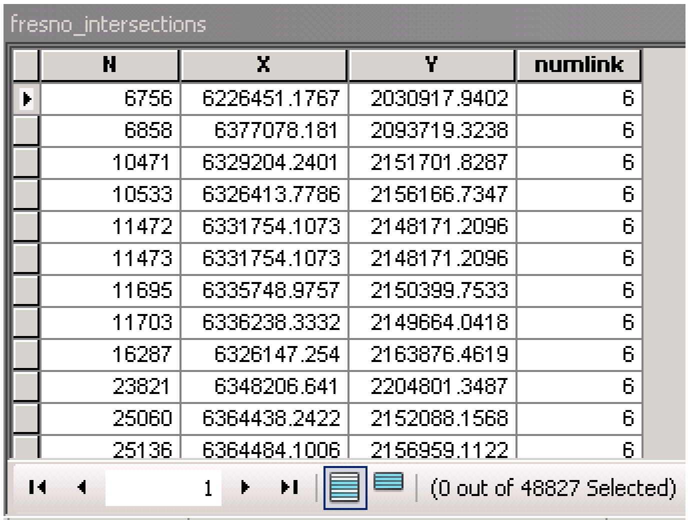
Follow the sequential steps below:

- first use “intersect” tool in ArcGIS to identify the links that have overlapped links. The output would be link shapefile containing only the spatially duplicate links or the two-way roads represented by two links.
- Join these links with the links in the original highway link file (no connectors, no highways, and no ramps) using field OBJECTID and select these links using OBJECTID>0.
- Create a new field (double) “numlinks” and assign a default value of 0.5 to the selected links.
- Switch selection and assign numlinks=1 to rest of the links. This way we have identified which links are spatially duplicate.

- Perform spatial join on highway nodes (target: node layer, join: link layer). In the spatial join tool, delete all fields except N, X, Y, and numlinks. Also, right click on the field "numlinks" as choose 'Merge Rule' as Sum.



- Now, in the output file only keep following fields: N, X, Y, numlink. Sort the field numlinks ascendingly and see if there are any node with numlinks less than 1. If there are some with 0.5, set them to 1.



fresno_intersections

	N	X	Y	numlink
▶	6756	6226451.1767	2030917.9402	6
	6858	6377078.181	2093719.3238	6
	10471	6329204.2401	2151701.8287	6
	10533	6326413.7786	2156166.7347	6
	11472	6331754.1073	2148171.2096	6
	11473	6331754.1073	2148171.2096	6
	11695	6335748.9757	2150399.7533	6
	11703	6336238.3332	2149664.0418	6
	16287	6326147.254	2163876.4619	6
	23821	6348206.641	2204801.3487	6
	25060	6364438.2422	2152088.1568	6
	25136	6364484.1006	2156959.1122	6

1 (0 out of 48827 Selected)

- Do attribute selection as “numlink>0” and export the selected records to a csv file: intersections_2014.csv
- Rename and arrange the columns in the order as required: id, links, xcoord_p, ycoord_p.

OPENSPACE AND PARKS FILE

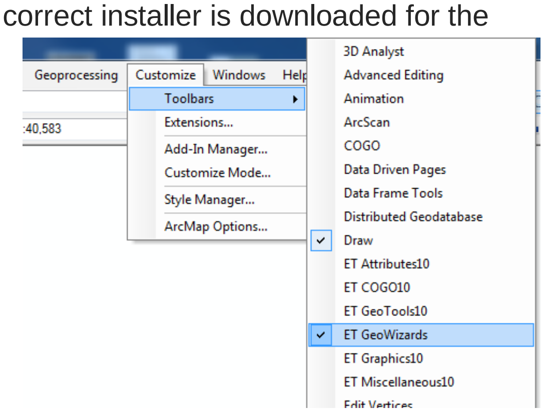
An ArcTool (ArcMap based) is created to process data in shapefile format and generate information in DaySim format. The tool is available on the Fresno ABM GitHub repository under tools.

Downloading and Installing ET Geowizards:

Continue to the next section if the current version of “ET Geowizards” is installed on the work computer.

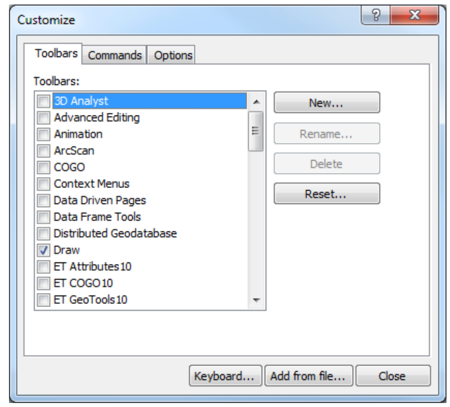
Following the steps below:

1. Locate the “ET Geowizards” installer (which can be found at the following Web address: http://www.ian-ko.com/ . *Note: Makes sure that correct installer is downloaded for the specific computer and the specific version of ArcGIS being run on that computer.
2. Install the software by following the instructions that accompany the installer file.
3. Open a new project in ArcMap



4. Click the “Customize” dropdown menu at the top of the ArcMap window→ hover over “Toolbars”→ Select “ET Geowizards” in the adjacent dropdown menu (see graphic to the right). Feel free to doc the “ET Geowizards toolbar”.

5. If “ET Geowizards” does not appear in the aforementioned dropdown menu, please recheck that the installer ran correctly. If it did install correctly, navigate to the bottom of the list of “Toolbars” and select “Customize” → a new window will appear which should look like the graphic to the right. Click on the “Add from file….” Button, and navigate to the *.dll file created by the Geowizards installer.



Finding Open source Representations of Protected Areas

Continue to the next section if adequate spatial representations (in polygon format) of the region’s open space areas have already been obtained.

Protected areas are locations which receive protection because of their recognized natural, *ecological* and/or cultural benefits. The USGS, as well as some state government agencies, actively monitor and inventory these areas. Spatial representations of California’s “Protected Areas” can be found here: *http://www.calands.org/* . For spatial representations of “Protected Areas” at a national level, follow this link: *http://gapanalysis.usgs.gov/data/padus-data/* *Note, the following steps use the California data, additional notes will supply directions for using the national data.

6. Navigate to the “CPAD: California Protected Areas Database” website using the link above. Download the most recently updated geospatial data. *Suggestion: download the data in *gdb format.

7. *Note: if using the national database, navigate to the second web address provided and download the equivalent data.

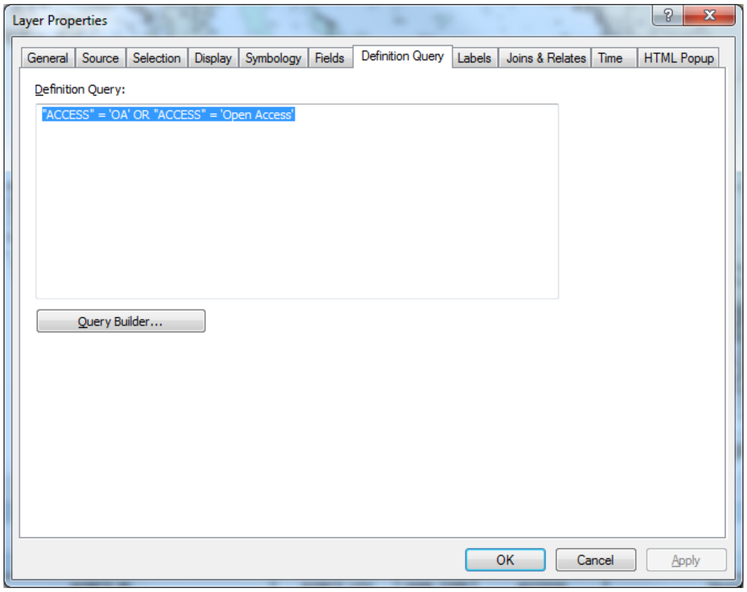
8. Once the data has completed downloading, open a new project in ArcMap and use the “Add Data” button  to add the shapefile (or feature class) named in the following convention “*_Holdings” (this data will be referred to as “_Holdings”).

9. Right click on the “_Holdings” file in the “Table of Contents” and select properties.

10. Navigate to the “Definition Query” tab.

11. In the box under “Definition Query:” write in the following expression:

"ACCESS" = 'OA' OR "ACCESS" = 'Open Access'



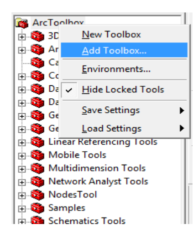
*Note: Users of the "national" database will have to determine their own qualifications for the equivalent of California's "Open Access" space

12. Add the shapefile/feature class representing the region's extent (forward referred to in the documentation as Extent").

Adding the "OpenspaceTool" to ArcToolbox

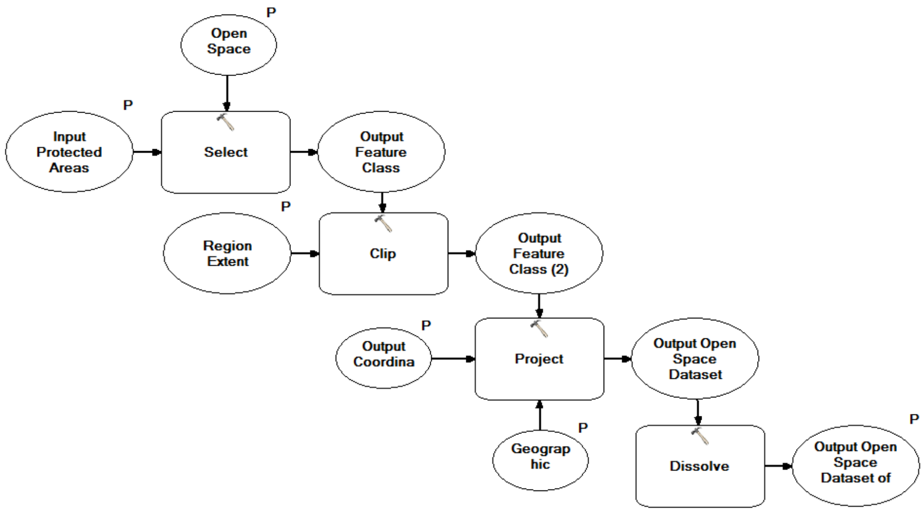
An ArcTool was created to process the data created in the previous steps.

13. Open the ArcToolbox 



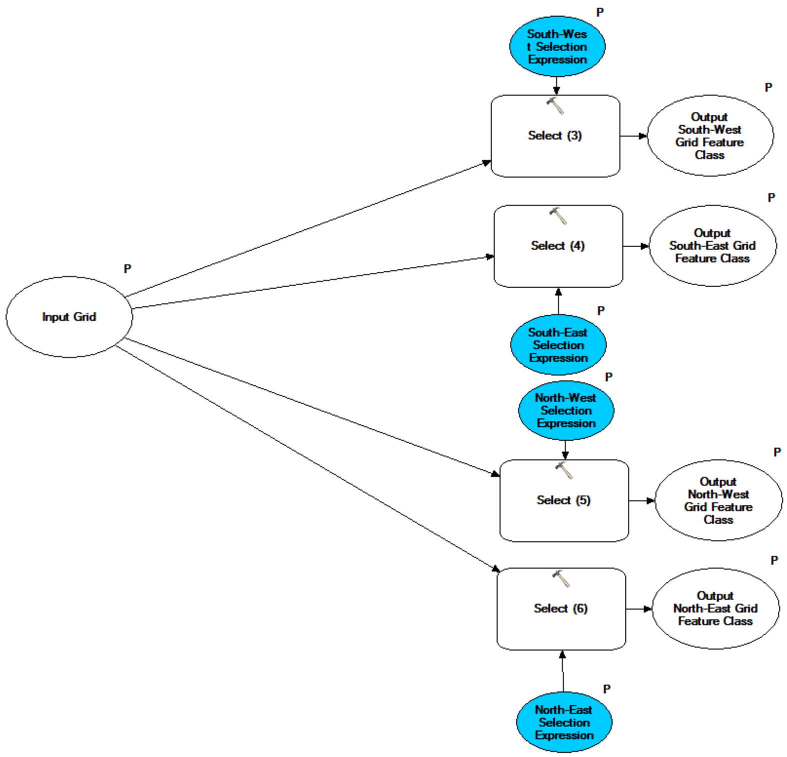
14. Right Click the "ArcToolbox" folder at the top of the directory and select "Add Toolbox"
15. Locate the "ParksTool" Toolbox on the computer and select "Ok"

Note: If problems arise with the "OpenspaceTool" Model consult the below image to understand the processes that make up the tool's operation.



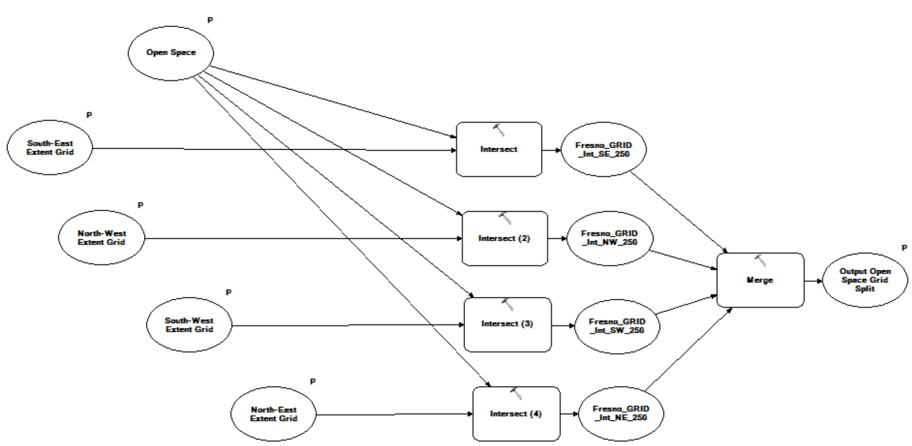
The tool first Selects the data by attribute to subset the "Open Space" data from the larger umbrella of "Protected Areas" (it is up to the user to decide what "attribute" in the database designates a feature as "Open Space", for the California Data we suggest the "ACCESS" field where attributes are either 'OA' or 'Open Access'); The tool then clips the subset to the extent of the region in question; Then, the tool projects the data into the proper projection (reminder: all data should be prepared in a common projection); Last, the output is dissolved into one feature, this output constitutes all the "Open Space" in the region in question.

*Note: If problems arise with the "Grid1" Model consult the below image to understand the processes that make up the tool's operation.



The tool merely selects data from the input grid and creates 4 new shapefiles/feature classes.

*Note: If problems arise with the “Grid2” Model consult the below image to understand the processes that make up the tool’s operation.



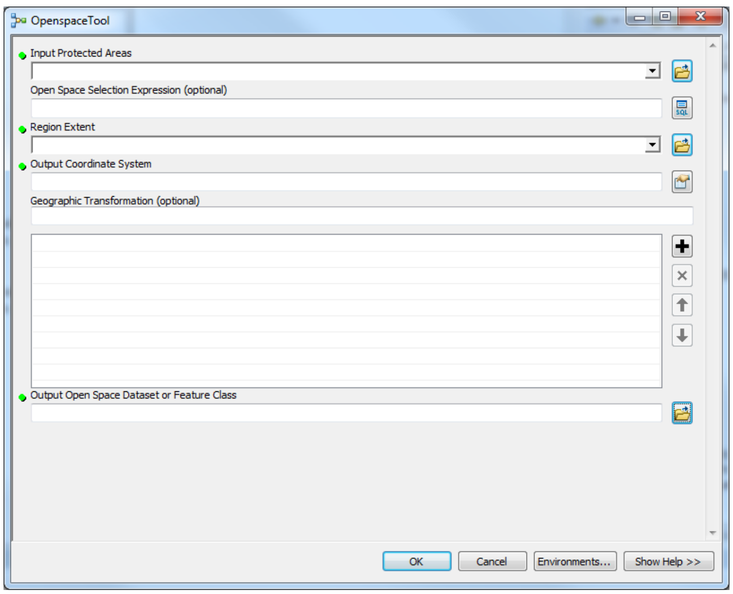
Based on user inputs the tool runs intersect the “Open Space” data with each of the 4 distinct sub-regions; once the data has been intersected, the 4 sub-regions are merged together.

Refining the "Protected Areas" data into "Open Space" Representations

NOTE: Steps 16-19 are not needed as FresnoCOG provided open space that are only in Fresno County.

The following steps will use the protected areas data, "OpenspaceTool", and Geowizards Toolbar to refine and separate the data into sections.

16. Open the "OpenspaceTool" Model in the "ParksTool", a window like the one below should appear.



17. Submit the form with the following inputs:

    -Input Protected Areas: -"_HOLDINGS" feature class

    -Open Space Selection Expression (optional): -*"ACCESS" = 'OA' OR "ACCESS" = 'Open Access'*

    -Region Extent -"Extent" feature class

    -Output Coordinate System -Default Projection for Project

    -Geographic Transformation -Leave blank unless asked to fill

    -Output Open Space Dataset or Feature Class - Name/Location of Open space feature class (forward referred to in the documentation as " Extent Open Space")

18. Click "OK".

**Note: if an error occurs run the "Select Tool" → "Clip Tool" → "Project Tool" and "Dissolve Tool" to get the same results*

19. Make sure the new feature class, "Extent OpenSpace", is added to the ArcMap display. Remove all other data in the table of contents. This new feature class represents all the "Open Space" in the region in question.

Refining "Open Space" data to create a DAYSIM Input File

The following steps: 1.) Divide the region into a grid with 250' by 250' grid cells; 2.) Intersect the open space data, that has already been created, with the new grid; 3.) Converts the intersected "Open Space" data into centroid points that contain area data.

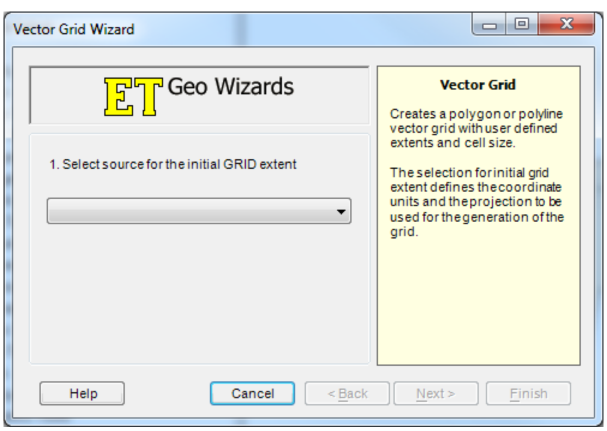
While ET Geowizards is a useful tool, there are many limitations to the tasks that the software can run. Many of the following steps concern getting around these limitations to create data that is as fine as possible. Feel free to attempt condensing the following steps into a single step (this may be possible for smaller, and/or less irregularly shaped, regions).

1. Open the "ParksExtent" excel workbook.

   *Note: the excel document will be used to calculate some values to guide the production of the grid.

2. Return to ArcMap and Click the "ET Geowizards" Toolbar 

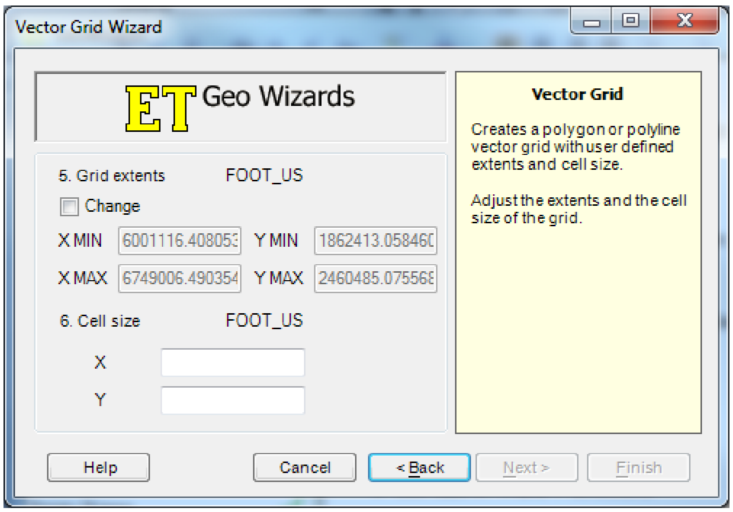
3. Navigate to the "Sampling" Tab → Double Click on the "Vector Grid" tool. A window similar to the one below will appear.



4. For object #1 "Select source for the initial GRID extent" use the "Extent OpenSpace" feature class.

5. For object #2 "Specify output feature class or shapefile" use a name and location of users choice (here forward referred to as "Extent_Grid") → Click on the "Next>" button.

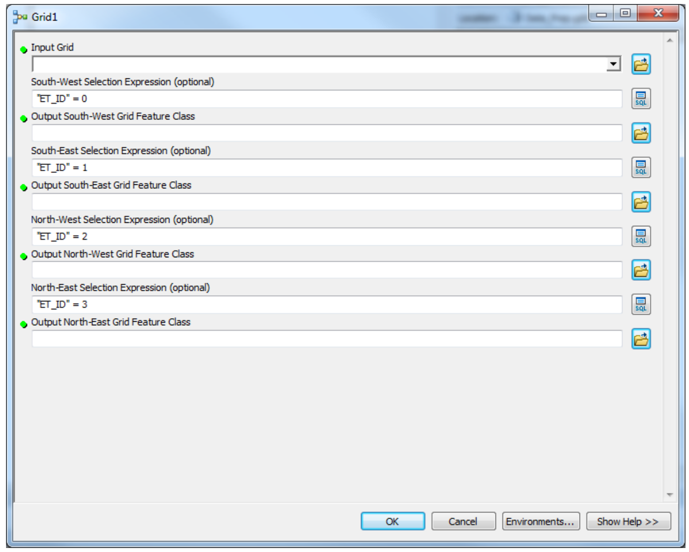
6. For #3 “Select output coordinate system”, click the “Select Output coordinate system” button and ensure that the coordinate system matches the default coordinate system for the project. If it does not, change the projection to match the default projection for the project.
7. For object #4 “Select GRID type” select the option for “Polygon”. Click on the “Next>” button. The window should now look like this:



8. Object #5 “Grid extents” will need to be updated. The excel workbook generates the required values based on the region specific data. Copy the “X Min”, “X Max”, “Y Min”, and “Y Max” values into the appropriate cells in the excel workbook. The work book generates new “X Max” and “Y Max” values, copy and paste the new values into their appropriate places in Object #5.

   The values can be copied by checking “Change“. This will turn boxes to edit box and you can copy value from each cell and paste into the excel workbook.
9. For Object #6 “Cell Size”, fill in the “X” and “Y” dimensions based on the values in the excel workbook. Click the “Finish” button.
10. Once the Geowizards processes have finished running, make sure that the new feature class, “Extent Grid,” is added to the display. The region should have been split into 4 different regions: two north and two south.

11. *Note: Given the limitation of the Geowizards software, the entire region can rarely ever be converted to a 250' by 250' Grid in one step. One can make an attempt at creating the 250' by 250' Grid in one step, but should not be surprised if/when the software fails.



12. Double click on the "Grid1" Model within the "ParksTool" Toolbox. A window similar to the one to the right will open.
13. *Note, the "Grid1" Tool separates the "Extent Grid" into 4 separate files. The defaults identify Geowizards conventions to separate the region into SE, SW, NE, and NW regions. If this tool fails to run, the step can be replicated by manually selecting each region, and exporting the selected data to shapefile.

    Submit the form with the following inputs:

    -Input Grid - "Extent_Grid" feature class

    - South-West Selection Expression (optional) - "ET_ID" = 0

    -Output South-West Grid Feature class - Name/Location of choosing (forward referred to as "Extent_Grid_SW")

    - South-East Selection Expression (optional) - "ET_ID" = 1

    -Output South-East Grid Feature class - Name/Location of choosing (forward referred to as "Extent_Grid_SE")

    - North-West Selection Expression (optional) - "ET_ID" = 2

    -Output North-West Grid Feature class - Name/Location of choosing (forward referred to as "Extent_Grid_NW")

    - North-East Selection Expression (optional) - "ET_ID" = 3

    -Output North-East Grid Feature class - Name/Location of choosing (forward referred to as "Extent_Grid_NE")

14. The following steps will subdivide each of the 4 regions created by step #19 into 250 foot grid cells. Re-open the "ET Geowizards" form and navigate to the "Vector Grid" tool.

Submit the form once for each region using the following inputs (* represents each of the four subdivisions):

#1 "Select source for the initial GRID extent" - once each for all "Extent_Grid_*"

#2 "Specify output feature class or shapefile" - Name/Location of users choosing

(forward referred to as "Extent_Grid_*_250")

#3 "Select output coordinate system"- Check that it matches the default projection for which all data is being prepared

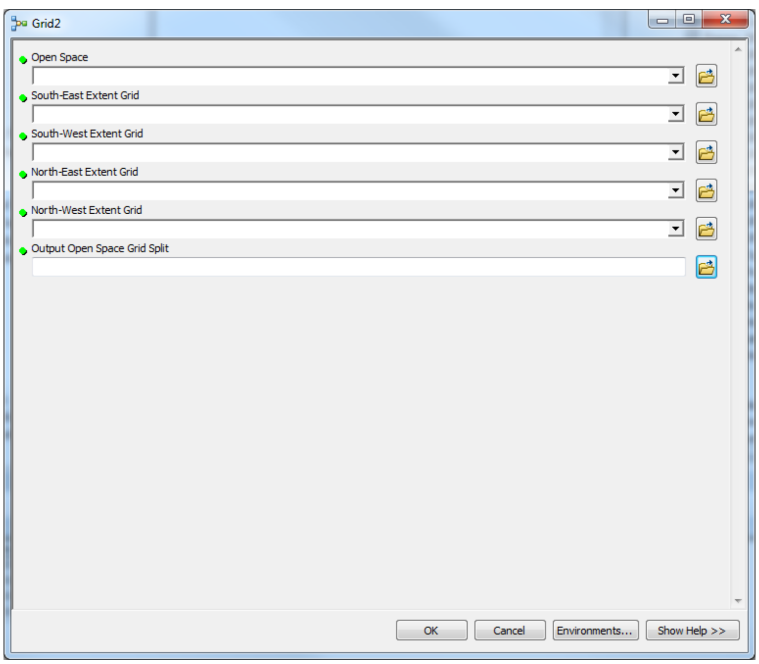
#4 "Select GRID type"- Polygon

#5 "Grid extents" - Use default

#6 "Cell Size" - both X and Y should = 250

Once generated, make sure all the "Extent_Grid_*_250" shapefiles are present in the ArcMAP display.

15. Double click the "Grid2" model within the "ParksTool" Toolbar. A window should appear matching the image below.



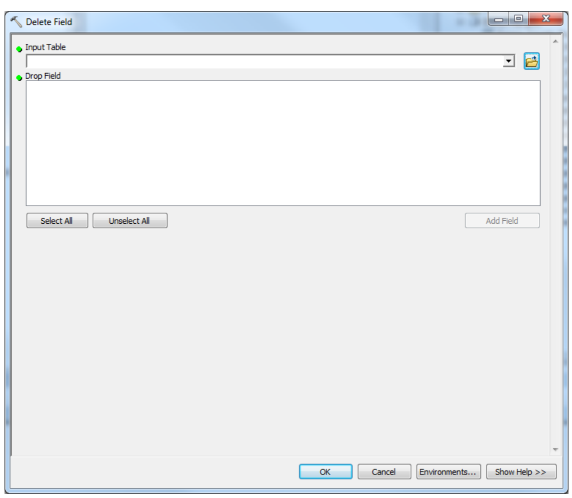
Fill in the appropriate fields with the following data:

“Open Space”	- “Extent Open Space” feature class
“South-East Extent Grid”	-“Extent_Grid_SE_250” feature class
“South-West Extent Grid”	-“Extent_Grid_SW_250” feature class
“North-East Extent Grid”	-“Extent_Grid_NE_250” feature class
“North-West Extent Grid”	-“Extent_Grid_NW_250” feature class
“Output Open Space Grid Split”	- Name/Location of choosing

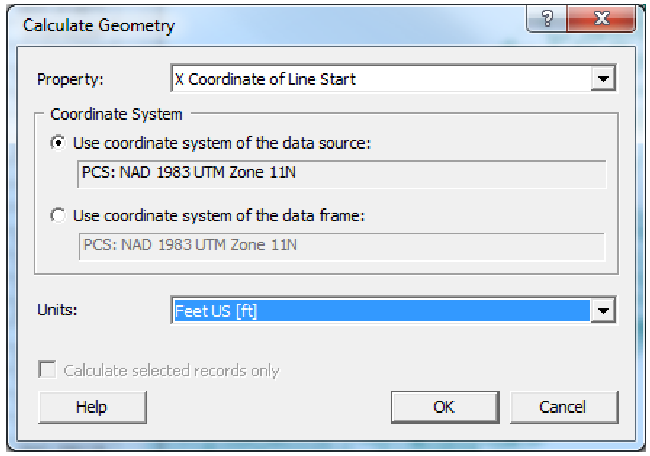
(forward referred to as “Extent Open Space 250”)

Click the “OK” button once all fields have been properly filled. Make sure the feature class generated through this tool, “Extent Open Space 250”, is added to the display.

16. Navigate to the “Delete Field” Tool in the ArcToolbox (“Data Management Tools” → “Fields” → “Delete Field”). A form like the image below should open.



17. The “Input Table is the “Extent Open Space 250” feature class. Check the box next to all the fields present in the “Drop Field” box. Click Ok.
18. Open the “Extent Open Space 250” feature class ‘s attribute table to check that all fields (“ID”, “FID”, “Name”, etc.) have been removed. **Note: some fields are structurally mandatory and cannot be removed.*
19. In the Attribute Table menu select  and click on the option to “Add Field”. The “Name” will be “sqft” and the “Type” will be Long Integer.
20. Repeat the above step and add fields “xcoord_p” (Long) and “ycoord_p” (Long)



21. Right click on the "sqft" column heading and click the option to "Calculate Geometry…" The "property to be calculated" is "Area"; The Coordinate System choice is "Use coordinate system of the data source" and the Units should be "Square Feet US [sq ft]". Click "OK"
22. Right click on the "xcoord_p" column heading and click the option to "Calculate Geometry…" The "property to be calculated" is "X Coordinate of Centroid"; The Coordinate System choice is "Use coordinate system of the data source" and the Units should be "Feet US [sq ft]". Click "OK"
23. Right click on the "ycoord_p" column heading and click the option to "Calculate Geometry…" The "property to be calculated" is "Y Coordinate of Centroid"; The Coordinate System choice is "Use coordinate system of the data source" and the Units should be "Feet US [sq ft]". Click "OK".

This will give you the openspace and parks data in required DaySim format.

FUTURE YEAR DAYSIM INPUTS

Table 86 provides instructions on what DaySim inputs to update for a future scenario (DaySim inputs are here: 1_Inputs\8_DaySim):

TABLE 86: INSTRUCTIONS FOR FUTURE YEAR DAYSIM INPUTS

INPUT LOCATION	INSTRUCTION
01_TAZ_Index	Don't need to change unless your scenario has updated zone system.
02_Parcel	Need to update three input files. Open space and park file would not need to be changed, unless you know open space and park information for the future scenario which is unlikely. This appendix provides instructions to create these files: • maz_2014_parks.csv • intersections_2014_nohwys.csv

INPUT LOCATION	INSTRUCTION
	• stops_transit_2014.csv
03_Household	Need to update. Output of PopSynIII converted to DaySim format.
04_Person	Need to update. Output of PopSynIII converted to DaySim format.
05_ixxi	Update if you have share of XI and IX workers by zone (XI-outside workers working in Fresno; IX-resident workers going out for work). Otherwise, use the same as base year
06_Roster	Update skim paths if scenario name is other than FC14_BASE. The paths are stored in field “name” (Column I).
07_Coefficients	Don’t need to change.
10_ParkAndRide	Change only if you have updated park and ride locations in the region.

APPENDIX D. GENERATE DAYSIM CALIBRATION SUMMARIES

A set of R scripts generates various calibration summaries using DaySim outputs. The setup, called “daysim_summaries”, is located under the root Fresno ABM folder.

SOFTWARE REQUIREMENT

The setup required the following software to be available on the machine:

- R (tested version - 3.3.1)
- R studio (tested version - 0.99.903)

FOLDER STRUCTURE

The folder “daysim_summaries” contains several R scripts and three sub directories (see Table 87). The outputs are produced under the folder “output”. Note that the folder “output” should always contain summary spreadsheet as the setup expects them to be there. The setup reads and writes new summaries to them during a new run.

TABLE 87. DAYSIM SUMMARIES FOLDER STRUCTURE

FOLDER/FILE	DESCRIPTION
data	Survey data and district crosswalks
template	Settings to populate spreadsheets in the output folder
output	DaySim summaries in spreadsheet format
*.R	R scripts

SETUP

The setup requires following updates before running it on a new location:

- daysim_output_config.R - DaySim output locations

```
dshhfile                              = "E:/Projects/Clients/FresnoCOG/Model/FresnoCOG_ABM_new/Scenarios/FC14_BASE/11_DaySim/_household.tsv"
dsperfile                             = "E:/Projects/Clients/FresnoCOG/Model/FresnoCOG_ABM_new/Scenarios/FC14_BASE/11_DaySim/_person.tsv"
dspdayfile                            = "E:/Projects/Clients/FresnoCOG/Model/FresnoCOG_ABM_new/Scenarios/FC14_BASE/11_DaySim/_person_day.tsv"
dstourfile                            = "E:/Projects/Clients/FresnoCOG/Model/FresnoCOG_ABM_new/Scenarios/FC14_BASE/11_DaySim/_tour.tsv"
dstripfile                            = "E:/Projects/Clients/FresnoCOG/Model/FresnoCOG_ABM_new/Scenarios/FC14_BASE/11_DaySim/_trip.tsv"
```

- main.R – working directory

```
setwd("E:/Projects/Clients/FresnoCOG/Tasks/calibration/daysim_summaries")
```

- Install R libraries - use the following commands on a R console to install R libraries required to run the R scripts in the setup:

```
install.packages("foreign")

install.packages("reshape")
install.packages("XLConnect")
install.packages("descr")
install.packages("Hmisc")
install.packages("data.table")
install.packages("plyr")
```

RUN SUMMARIES

The summaries are run using the following sequential steps:

- Open R Studio
- Open “main.R” within the R studio
- Select all text in the script “main.R” and click Run (Ctrl+Enter)

The outputs are produced under the folder “output”.

OUTPUT

The summaries produce following outputs under the folder “output”:

TABLE 88. DAYSIM SUMMARIES OUTPUTS

FILE	DESCRIPTION
VehAvailability.xlsm	Auto ownership summaries
WrkLocation.xlsm	Work location choice summaries
SchLocation.xlsm	School location choice summaries
DayPattern.xlsm	Person daily activity pattern summaries

TourDestination_[purpose].xlsm	Tour destination choice summaries by purpose (meal, shopping, personal business, social/recreation, escort, work-based)
TourDestination_maintenance.xlsm	Tour destination choice summaries for maintenance purpose – aggregation of shopping and personal business
TourDestination_discretionary.xlsm	Tour destination choice summaries for maintenance purpose – aggregation of meal and social/recreation
TripDestination.xlsm	Trip destination choice summaries
TourMode.xlsm	Tour mode choice summaries
TripMode.xlsm	Trip mode choice summaries
TourTOD.xlsm	Tour time of day summaries
TripTOD.xlsm	Trip time of day summaries

APPENDIX E. GENERATE ASSIGNMENT VALIDATION SUMMARIES

Assignment validation summaries are created in excel spreadsheets. The spreadsheets, one for each highway and transit, take input data and calculate validation summaries automatically. The required files for generating assignment validation summaries are here: 1_Inputs_Support\Validation.

SETUP

Install a python library called "simpledbf" (https://pypi.org/project/simpledbf/) using one of the two methods below:

- Install via pip as below (by simply entering the commands below on a DOS prompt):

```
$ pip install simpledbf
```

  If trouble installing using pip refer to the following help page: https://github.com/BurntSushi/nfldb/wiki/Python-&-pip-Windows-installation

- Alternatively, install using conda if you already have Anaconda installed:

```
$ conda install -c https://conda.binstar.org/rnelsonchem simpledbf
```

GENERATE INPUT DATA

A python script "assignment_validation_summaries.py" generates highway and transit validation inputs that go into validation spreadsheets. The script is run using the following steps:

- Open "assignment_validation_summaries.py
- Update working directory to match the folder with assignment results under the ABM run directory.

```
wd = r'E:\Projects\Clients\FresnoCOG\Model\FresnoCOG_ABM_new\Scenarios\FC14_BASE\09_Assignment'
```

- Open a DOS command window and run the python script (make sure that python is set to Ananconda installation) as below:

```
E:\Projects\Clients\FresnoCOG\Tasks\validation>python assignment_validation_summaries.py
```

- The run will generate two outputs under the working directory:
  - Fresno_hwy_validation.csv
  - trn_vol_abm.csv

UPDATE VALIDATION SPREADSHEETS

Highway

Highway validation summaries are in “hwy_validation.xlsx”. To update the spreadsheet for a new run, do the following:

- copy all data from “Fresno_hwy_validation.csv”
- paste the copied data in sheet “model” at cell C1.
- Now, go to sheets “CalTrans_Counts_Compare” and “Screenline” and refresh pivot tables.

The updated summaries are under sheets “Summary_All”, “Screenlines”, and “CalTrans_Counts_Compare”. The “CalTrans_Counts_Compare” include comparisons of CalTrans counts only. The “Summaries_All” include comparisons of all counts (CalTrans counts and Fresno counts). The “Screenline” contains comparisons by screenlines.

Transit

Transit validation summaries are in “trn_validation.xlsx”. To update the spreadsheet for a new run, do the following:

- copy all data from “trn_vol_abm.csv”
- paste the copied data in sheet “trn_vol_abm” at cell A1

The updated summaries are under sheet “compare” and include various comparisons of observed and estimated boarding by route.

Bike and Ped Validation

The bike and ped validation summaries are in “BikePed_Validation.xlsx”. To update the spreadsheet for a new run, do the following:

- Copy all data from “FC14_BASE_LOADEDNETWORK_BIKE.dbf” and paste it in sheet “FC14_BASE_LOADEDNETWORK_BIKE“ at cell B1.
- Copy all data from “FC14_BASE_LOADEDNETWORK_WALK.dbf” and paste it in sheet “FC14_BASE_LOADEDNETWORK_WALK“ at cell B1.

The updated summaries are under sheets “bike” and “walk” and include comparisons of observed and estimated bike/ped flows by count group location.

APPENDIX F. DAYSIM CONFIGURATION SETTINGS

This appendix describes the settings used in running DaySim. Note that a Cube script creates a configuration file before a DaySim run. So, any change to a setting, filename or file path needs to be done in two Cube scripts: DAMAT00F.S (Configuration_shadow_price.properties) and DAMAT00A.S (Configuration.properties). Both configuration files have the same settings, except that some models are turned off in shadow price runs. Table 89 presents the settings in the configuration files. The column "VALUE" contains the setting as in the Configuration.properties.

TABLE 89: DAYSIM CONFIGURATION FILE (CONFIGURATION.PROPERTIES)

SETTING	VALUE	DESCRIPTION
Sampling		
HouseholdSampling RateOneInX	1	The denominator of the fraction of households in the input sample to be simulated (e.g., 100 is for 1/100)
HouseholdSampling StartWithY	1	The household number to simulate first (e.g., 2, in combination with 100 above would simulate HH 2, 102, 202, etc.)
SamplingWeights SettingsType	SamplingWeights SettingsSimple	
General Path Settings		
BasePath	[CATALOG_DIR]\1_Inputs\8_DaySim	Base directory; all DaySim inputs will be stored

SETTING	VALUE	DESCRIPTION
OutputSubpath	[SCENARIO_DIR]\11_DaySim	Output folder path; DaySim outputs will be generated in this directory
WorkingDirectory	[SCENARIO_DIR]\11_DaySim \working	Directory for other DaySim outputs. DaySim generates other outputs in this directory.
WorkingSubpath	[SCENARIO_DIR]\11_DaySim\working	Directory path for other DaySim outputs
Threading Settings		
NProcessors	48	Number of processors (threads) to be used
NBatches	50	
Region Specific		
ShouldRunInputTester	true	Flag to run DaySim input checks
CustomizationDll	Fresno.dll	Region-specific DLL
Parcel Buffered Data		
ImportParcels	true	Flag to import parcel file
RawParcelPath	[SCENARIO_DIR]\Fresno_mzbuffer_allstreets_2014.dat	Buffered parcel file name
RawParcelDelimiter	32	Buffered parcel file delimiter (9=TAB, 32=space, 44=comma)

SETTING	VALUE	DESCRIPTION
UseParcelLandUseCodeAsSquareFeetOpenSpace	True	Switch for using open space square feet data from land-use type field (“lutype_p”) in the parcel file
Roster Impedance		
UseMicrozoneSkimsForBikeMode	TRUE	Switch for using bike skims by microzones
RosterPath	.\06_Roster\roster_mz.csv	Name of roster CSV file, including full directory path
RosterCombinations Path	.\06_Roster\roster_combinations.csv	Name of valid roster combinations CSV file, including full directory path
UseShortDistance NodeToNode Measures	true	TRUE to use node-to-node distance in accessibility measures calculations
RawParcelNode Path	[SCENARIO_DIR]\maz_node_2014.dat	File name providing the nearest node ID for a parcel
RawParcelNode Delimiter	32	Delimiter of the input parcel node file (9=TAB, 32=space, 44=comma)

SETTING	VALUE	DESCRIPTION
NodeIndexPath	[SCENARIO_DIR]\WALK_SKIM_MAZ_MAZ_SORTED_INDEX.TXT	File name for the file providing, for every node ID, starting and end record indices in node short-distance file (NodeDistancePath)
NodeIndexDelimiter	32	Delimiter of the node index file (9=TAB, 32=space, 44=comma)
NodeDistances Path	[SCENARIO_DIR]\WALK_SKIM_MAZ_MAZ_SORTED.TXT.dat	File name for the file providing short distances for node pairs
MaximumBlending Distance	3	The maximum (network) Distance for which short-distance blending should be used, in miles. For short trips, DaySim uses a linear combination of parcel-to-parcel distances from an all streets network and zone-zone distances from the skim matrices. When the zone-zone skim distance exceeeds this MaximumBlendingDistance, DaySim stop using the parcel-parcel distance and just use the zone-zone from the skims.
AllowNodeDistanceAsymmetry	True	

SETTING	VALUE	DESCRIPTION
UseShortDistance CircuityMeasures	false	true to use circuity distance in accessibility measures calculations
Value of Time		
VotVeryLowLow	0	Boundary between VeryLow and Low VOT groups, in monetary units per hour
VotLowMedium	6	Boundary between Low and Medium VOT groups, in monetary units per hour
VotMediumHigh	12	Boundary between Medium and High VOT groups, in monetary units per hour
VotHighVeryHigh	5001	Boundary between High and VeryHigh VOT groups, in monetary units per hour
Global Settings		
DataType	Default	Identifies the presence of client-specific household input data (currently only used for Actum)
ChoiceModelRunner	Default	Type of choice model runner

SETTING	VALUE	DESCRIPTION
Settings	DefaultSettings	
Debug Settings		
TraceSimulatedChoice Outcomes	false	true to trace simulated choice outcomes
TraceModelResult Validity	false	true to trace model result
InvalidAttempts BeforeTrace	100	
InvalidAttempts BeforeContinue	4	
ReportInvalidPerson Days	false	true to report invalid person days during a run
Shadow Price Settings for work and school locations		
ShouldUse ShadowPricing	true	true to apply shadow pricing for the Work Location and School Location models
UsualWorkParcel Threshold	5	Parcel-specific threshold used in the shadow price calculations of usual work location
UsualSchoolParcel Threshold	5	Parcel-specific threshold used in the shadow price calculations of usual school location

SETTING	VALUE	DESCRIPTION
UsualUniversity ParcelThreshold	5	Parcel-specific threshold used in the shadow price calculations of usual university location
NumberOfParcels InReportDiffs	10	Control for printing out reporting on shadow price calculations
UsualWork PercentTolerance	0	Percentage tolerance to trigger work parcel shadow price adjustment
UsualWork AbsoluteTolerance	0	Absolute tolerance to trigger work parcel shadow price adjustment
UsualSchool PercentTolerance	0	Percentage tolerance to trigger school parcel shadow price adjustment
UsualSchool AbsoluteTolerance	0	Absolute tolerance to trigger school parcel shadow price adjustment
UsualUniversity PercentTolerance	0	Percentage tolerance to trigger university parcel shadow price adjustment

SETTING	VALUE	DESCRIPTION
UsualUniversity AbsoluteTolerance	0	Absolute tolerance to trigger university parcel shadow price adjustment
Shadow Price Settings		
ShadowPriceDelimiter	9	Delimiter of the shadow price files (9=TAB, 32=space, 44=comma)
ShouldUse ParkAndRide ShadowPricing	TRUE	true to use park-and-ride shadow pricing in the model
ParkAndRide ShadowPrice Delimiter	9	Delimiter of the park-and-ride shadow pricing file (9=TAB, 32=space, 44=comma)
ParkAndRide ShadowPrice MaximumPenalty	-3	
ParkAndRide ShadowPriceTime Spread	20	
ParkAndRide ShadowPriceStep Size	0.15	
Model Run Flags		
ShouldRunChoiceModels	True	A toggle switch to run all choice models (true can be overridden by switches below and by individual model switches)

SETTING	VALUE	DESCRIPTION
ShouldRunHouseholdModels	true	A toggle switch to run household-level models (used to perform partial runs, TRUE can be overridden by individual model switches)
ShouldRun PersonModels	True	A toggle switch to run person-level models (used to perform partial runs, true can be overridden by individual model switches)
ShouldRun PersonDayModels	true	A toggle switch to run person-day-level models (used to perform partial runs, TRUE can be overridden by individual model switches)

SETTING	VALUE	DESCRIPTION
ShouldRunTour Models	True	A toggle switch to run tour-level models (used to perform partial runs, true can be overridden by individual model switches)
ShouldRunTourTripModels	true	A toggle switch to run trip-level models (used to perform partial runs, true can be overridden by individual model switches)
ShouldRunSubtourModels	True	A toggle switch to run subtour level models (used to perform partial runs, true can be overridden by individual model switches)
ShouldRunSubtourTripModels	true	A toggle switch to run trip level models for subtours (used to perform partial runs, true can be overridden by individual model switches)

SETTING	VALUE	DESCRIPTION
DestinationScale	1	For a model that uses parcels, this should be set at 0. If the model uses blocks (microzones) instead, then set this at 1, and it allows intra-microzone trips. But, with it set at 0, it does not allow intra-microzone trips.
ShowRunChoiceModelsStatus	True	true to show percentage of households simulated on the screen during simulation
ShouldRunRawConversion	True	If true, DaySim will convert and input all of the raw data files
Random Seed Settings		
RandomSeed	1234	Initial seed value for the random number generator
ShouldSynchronize RandomSeed	True	If true, DaySim will use a seed for each person/tour/trip/model combination that depends only on the initial seed

SETTING	VALUE	DESCRIPTION
Inernal-External Workers Settings		
IxxiPath	.\05_ixxi\Fresno_worker_ixxifractions.dat	Input worker IXXI fractions file name
IxxiDelimiter	32	Delimiter for the input file (9=TAB, 32=space, 44=comma)
IxxiFirstLineIsHeader	False	If true, DaySim expects a header record for this file (all other raw data files have headers)
Park-and-Ride Nodes		
RawParkAndRide NodePath	.\10_ParkAndRide\p_r_Nodes_2014.dat	Input park-and-ride node file; if none given, the park-and-ride mode will not be available)
RawParkAndRideNodeDelimiter	9	The delimiter for the input file (9=TAB, 32=space, 44=comma)
ImportParkAndRideNodes	true	If true, the raw file should be imported (always true if ShouldRunRawConversion=true)
ShouldReadParkAndRideNodeSkim	false	
Zones		

SETTING	VALUE	DESCRIPTION
ImportZones	true	If true, the zone file should be imported (always true if ShouldRunRawConversion=true)
RawZonePath	.\01_TAZ_Index\Fresno_taz_indexes.dat	Input zone index file name
RawZoneDelimiter	9	The delimiter for the input zone index file (9=TAB, 32=space, 44=comma)
RawHouseholdPath	.\03_Household\Fresno_household_sampled.dat	Input household file
RawHouseholdDelimiter	32	The delimiter for the input household file (9=TAB, 32=space, 44=comma)
RawPersonPath	.\04_Person\Fresno_person_sampled.dat	Input person file
RawPersonDelimiter	32	The delimiter for the input household file (9=TAB, 32=space, 44=comma)
DaySim Output Files		
ImportHouseholds	True	

SETTING	VALUE	DESCRIPTION
OutputHousehold Path	_household.tsv	The full path name for the household output file
OutputHousehold Delimiter	9	Delimiter for the household output file (9=TAB, 32=space, 44=comma)
ImportPersons	True	
OutputPersonPath	_person.tsv	Person output filename
OutputPerson Delimiter	9	Delimiter for the person output file (9=TAB, 32=space, 44=comma)
OutputHousehold DayPath	_household_day.tsv	
OutputHousehold DayDelimiter	9	Household day output filename
OutputPersonDay Path	_person_day.tsv	Person day output filename
OutputPersonDay Delimiter	9	Delimiter for the person day output file (9=TAB, 32=space, 44=comma)
OutputTourPath	_tour.tsv	Tour output filename
OutputTourDelimiter	9	Delimiter for the tour output file (9=TAB, 32=space, 44=comma)
OutputTripPath	_trip.tsv	Trip output filename

SETTING	VALUE	DESCRIPTION
OutputTripDelimiter	9	Delimiter for the trip output file (9=TAB, 32=space, 44=comma)
OutputJointTour Delimiter	9	Delimiter for the join tour output file (9=TAB, 32=space, 44=comma)
OutputJointTour Path	_joint_tour.tsv	Joint tour output filename
OutputFullHalfTour Path	_full_half_tour.tsv	Full half-tour output filename
OutputFullHalfTour Delimiter	9	Delimiter for the full half-tour output file (9=TAB, 32=space, 44=comma)
OutputPartialHalf TourPath	_partial_half_tour.tsv	Partial half-tour filename
OutputPartialHalf TourDelimiter	9	Delimiter for the partial half-tour output file (9=TAB, 32=space, 44=comma)
ShouldOutputTDM TripList	false	
Logsums		
ShouldLoadAggregate LogsumsFromFile	False	true to read the aggregate logsums from a file generated by a previous run (otherwise recalculated)

SETTING	VALUE	DESCRIPTION
ShouldOutput AggregateLogsums	True	true to write the aggregate logsums to a file for a subsequent run
OutputAggregate LogsumsPath	aggregate_logsums.dat	File name to write out aggregate logsums
ShouldLoadSampling WeightsFromFile	False	true to read the precalculated sampling weights from a file generated by a previous run (otherwise recalculated)
ShouldOutputSampling Weights	False	true to write the precalculated sampling weights to a file for a subsequent run
OutputSampling WeightsPath	sampling_weights.dat	File name write out sampling weights
Model Coefficients		
WorkLocationModelSampleSize	100	The maximum number of destinations to be sampled for this model
WorkLocationModelCoefficients	.\07_Coefficients\WorkLocationCoefficients.F12	Path of the coefficient file for the work location model

SETTING	VALUE	DESCRIPTION
ShouldRunWorkLocationModel	True	A toggle switch to run the work location model; can be used for partial runs, TRUE can be overridden by more general switches above
IncludeWorkLocationModel	True	false to always exclude this model from the set of models to be run. Set both - this and ShouldRUnWorkLocatinModel - to the same (true or false).
SchoolLocationModelSampleSize	100	The maximum number of destinations to be sampled for this model
SchoolLocationModelCoefficients	.\07_Coefficients\SchoolLocationCoefficients.F12	Path of the coefficient file for the school location model
ShouldRunSchool LocationModel	True	A toggle switch to run the school location model; can be used for partial runs, TRUE can be overridden by more general switches above

SETTING	VALUE	DESCRIPTION
IncludeSchool LocationModel	True	false to always exclude this model from the set of models to be run. Set both - this and ShouldRunSchoolLocationModel - to the same (true or false).
PayToParkAtWorkplace ModelCoefficients	.\07_Coefficients\PayToParkAtWorkplaceCoefficients.F12	Path of the coefficient file for the pay to park and workplace model
ShouldRunPayTo ParkAtWorkplace Model	True	A toggle switch to run the pay-to-park and workplace model; can be used for partial runs, true can be overridden by more general switches above
IncludePayTo ParkAtWorkplace Model	True	false to always exclude this model from the set of models to be run. Set both - this and ShouldRunPayToParkAtWorkplace Model - to the same (true or false).
TransitPassOwnership ModelCoefficients	.\07_Coefficients\TransitPassOwnershipCoefficients.F12	Path of the coefficient file for the transit pass ownership model

SETTING	VALUE	DESCRIPTION
ShouldRunTransit PassOwnershipModel	True	A toggle switch to run the transit pass ownership model; can be used for partial runs, true can be overridden by more general switches above
IncludeTransit PassOwnershipModel	True	false to always exclude this model from the set of models to be run. Set both - this and ShouldRunTransitPassOwnershipModel - to the same (true or false).
AutoOwnership ModelCoefficients	.\07_Coefficients\AutoOwnershipCoefficients.F12	Path of the coefficient file for the auto ownership model
ShouldRun AutoOwnershipModel	True	A toggle switch to run the auto ownership model; can be used for partial runs, true can be overridden by more general switches above
IndividualPerson DayPatternModel Coefficients	.\07_Coefficients\IndividualPersonDayPatternCoefficients.F12	Path of the coefficient file for the individual person-day pattern model

SETTING	VALUE	DESCRIPTION
ShouldRunIndividual PersonDayPattern Model	True	A toggle switch to run the individual person day pattern model; can be used for partial runs, true can be overridden by more general switches above
PersonExactNumberOf ToursModel Coefficients	.\07_Coefficients\PersonExactNumberOfToursCoefficients.F12	Path of the coefficient file for the person exact number of tours model
ShouldRunPerson ExactNumberOf Tours Model	True	A toggle switch to run the person exact number of tours model; can be used for partial runs, true can be overridden by more general switches above
WorkTourDestination ModelSampleSize	100	Maximum number of destinations to be sampled for the work tour destination model
WorkTourDestination ModelCoefficients	.\07_Coefficients\WorkTourDestinationCoefficients.F12	Path of the coefficient file for the work tour destination model
ShouldRunWorkTour DestinationModel	True	

SETTING	VALUE	DESCRIPTION
OtherTourDestination ModelSampleSize	100	Maximum number of destinations to be sampled for the other tour destination model
OtherTourDestination ModelCoefficients	.\07_Coefficients\OtherTourDestinationCoefficients.F12	Path of the coefficient file for the tour destination model
ShouldRunOtherTourDestinationModel	True	A toggle switch to run the tour destination model; can be used for partial runs, true can be overridden by more general switches above
WorkBasedSubtourGenerationModelCoefficients	.\07_Coefficients\WorkbasedSubtourGenerationCoefficients.F12	Path of the coefficient file for the work-based subtour generation model
ShouldRunWorkBasedSubtourGenerationModel	True	A toggle switch to run the work-based subtour generation model; can be used for partial runs, true can be overridden by more general switches above
WorkTourMode ModelCoefficients	.\07_Coefficients\WorkTourModeCoefficients.F12	Path of the coefficient file for the work tour mode model

SETTING	VALUE	DESCRIPTION
ShouldRunWorkTour ModeModel	True	A toggle switch to run the work tour mode model; can be used for partial runs, true can be overridden by more general switches above
SchoolTourMode ModelCoefficients	.\07_Coefficients\SchoolTourModeCoefficients.F12	Path of the coefficient file for the school tour mode model
ShouldRunSchool TourModeModel	True	A toggle switch to run the school tour mode model; can be used for partial runs, true can be overridden by more general switches above
WorkBasedSubtour ModeModelCoefficients	.\07_Coefficients\WorkBasedSubtourModeCoefficients.F12	Path of the coefficient file for the work-based subtour mode model
ShouldRunWorkBased SubtourModeModel	True	A toggle switch to run the work-based subtour mode model; can be used for partial runs, true can be overridden by more general switches above
EscortTourMode ModelCoefficients	.\07_Coefficients\EscortTourModeCoefficients.F12	Path of the coefficient file for the escort tour model model

SETTING	VALUE	DESCRIPTION
ShouldRunEscort TourModeModel	True	A toggle switch to run the escort tour model model; can be used for partial runs, true can be overridden by more general switches above
OtherHomeBased TourModeModel Coefficients	.\07_Coefficients\OtherHomeBasedTourModeCoefficients.F12	Path of the coefficient file for the other home-based tour mode model
ShouldRunOther HomeBasedTour ModeModel	True	A toggle switch to run the other home-based tour mode model; can be used for partial runs, true can be overridden by more general switches above
WorkTourTime ModelCoefficients	.\07_Coefficients\WorkTourTimeCoefficients.F12	Path of the coefficient file for the work tour time model
ShouldRunWorkTour TimeModel	True	A toggle switch to run the work tour time model; can be used for partial runs, true can be overridden by more general switches above
SchoolTourTime ModelCoefficients	.\07_Coefficients\SchoolTourTimeCoefficients.F12	Path of the coefficient file for the school tour time model

SETTING	VALUE	DESCRIPTION
ShouldRunSchoolTourTimeModel	True	A toggle switch to run the school tour time model; can be used for partial runs, true can be overridden by more general switches above
OtherHomeBasedTour TimeModelCoefficients	.\07_Coefficients\OtherHomeBasedTourTimeCoefficie nts.F12	Path of the coefficient file for the other home-based tour time model
ShouldRunOtherHomeBasedTour TimeModel	True	A toggle switch to run the home-based tour time model; can be used for partial runs, true can be overridden by more general switches above
WorkBasedSubtourTimeModelCoeff icients	.\07_Coefficients\WorkbasedSubtourTimeCoefficients. F12	Path of the coefficient file for the work-based subtour time model
ShouldRunWorkBasedSubtourTime Model	True	A toggle switch to run the work-based subtour time model; can be used for partial runs, true can be overridden by more general switches above

SETTING	VALUE	DESCRIPTION
IntermediateStopGenerationModelCoefficients	.\07_Coefficients\IntermediateStopGenerationCoefficients.F12	Path of the coefficient file for the intermediate stop generation model
ShouldRunIntermediate StopGenerationModel	True	A toggle switch to run the intermediate stop generation model; can be used for partial runs, true can be overridden by more general switches above
IntermediateStopLocation ModelSampleSize	100	The maximum number of destinations to be sampled for the intermediate stop location model
IntermediateStopLocation ModelCoefficients	.\07_Coefficients\IntermediateStopLocationCoefficients.F12	Path of the coefficient file for the intermediate stop location model
ShouldRunIntermediateStopLocationModel	True	A toggle switch to run the intermediate stop location model; can be used for partial runs, true can be overridden by more general switches above

SETTING	VALUE	DESCRIPTION
TripModeModelCoefficients	.\07_Coefficients\TripModeCoefficients.F12	Path of the coefficient file for the trip mode model
ShouldRunTripModeModel	True	A toggle switch to run the trip mode model; can be used for partial runs, true can be overridden by more general switches above
TripTimeModelCoefficients	.\07_Coefficients\TripTimeCoefficients.F12	Path of the coefficient file for the trip time model
ShouldRunTripTimeModel	True	A toggle switch to run the trip time model; can be used for partial runs, true can be overridden by more general switches above
Path Impedance Parameters		
PathImpedance_ PathChoiceScaleFactor	1.5	A scale factor for the coefficients of the path-type models; the inverse of a logsum coefficient in upper-level models. Not really used in BKRCast as BKRCast has only local bus under transit so don't have any path type competing under any modes.
PathImpedance_ AutoOperating CostPerMile	0.22	Auto operating cost, in monetary units per distance unit

SETTING	VALUE	DESCRIPTION
PathImpedance_ TransitInVehicle TimeWeight	1.0	Relative weight on transit in-vehicle time in the transit and park-and-ride path type models. These are all multiples of the auto in-vehicle time coeffcient, which is set in the VOT parameters.
PathImpedance_ TransitFirst WaitTimeWeight	2.0	Relative weight on transit first wait time in the transit and park-and-ride path-type models
PathImpedance_ TransitTransfer WaitTimeWeight	2.0	Relative weight on transit transfer wait time in the transit and park-and-ride path-type models
PathImpedance_ TransitNumber BoardingsWeight	8.0	Relative weight on transit number of boardings in the transit and park-and-ride path-type models
PathImpedance_ TransitDriveAccess TimeWeight	2.0	Relative weight on transit drive access in-vehicle time in the park-and-ride path-type models

SETTING	VALUE	DESCRIPTION
PathImpedance_ TransitWalkAccess TimeWeight	2.0	Relative weight on transit walk access and egress times in the transit and park-and-ride path-type models
PathImpedance_ WalkTimeWeight	2.5	Relative weight on walk mode time in the walk path-type model
PathImpedance_ BikeTimeWeight	4.0	Relative weight on bike mode time in the bike path-type model
PathImpedance_ WalkMinutesPerMile	20.0	Factor to convert parcel-based transit walk access/egress distance into time (in minutes per distance unit)
PathImpedance_ TransitWalkAccess DistanceLimit	1.0	Maximum parcel-based transit walk access or egress distance allowed for available transit paths
PathImpedance_TransitWalkAccessDirectLimit	1.0	Maximum parcel-based transit walk access or egress distance allowed for direct transit paths to be chosen over mixed paths

SETTING	VALUE	DESCRIPTION
athImpedance_ TransitSingleBoarding Limit	1.1	Maximum number of boardings for a transit path to be considered a "direct path" (no transfers). When DaySim figures out the walk time from a parcel to the nearest transit stop for path types other than local bus, DaySim wants to figure out whether those paths could include local bus feeder or if they are a direct path with no transfers (the boardings skim value is >=TransitSingleBoardingLimit). The fraction is because some transit skims are an average across multiple paths, some direct and some not.
PathImpedance_ AutoTolledPath Constant	0.0	Path-type constant for an auto path that includes a nonzero toll cost (reflects extra resistance to paying tolls)
PathImpedance_ AvailablePath UpperTimeLimit	200.0	Maximum total (unweighted) path travel time (in minutes) for a path to be considered an available option
PathImpedance_ TransitLocalBus PathConstant	0.0	Path-type constant for transit local bus only paths

SETTING	VALUE	DESCRIPTION
PathImpedance_ TransitPremiumBus PathConstant	0.0	Path-type constant for transit premium bus (possibly plus feeder) paths
PathImpedance_ TransitLightRail PathConstant	0.0	Path-type constant for transit light rail (possibly plus feeder) paths
PathImpedance_ TransitCommuterRail PathConstant	0.0	Path-type constant for transit commuter rail (possibly plus feeder) paths
PathImpedance_ TransitFerry PathConstant	0.0	Path-type constant for transit passenger ferry (possibly plus feeder) paths
PathImpedance_ TransitUsePathType SpecificTime	True	A switch to use additional skims and weights to reflect transit submode-specific in-vehicle times (SACOG-specific)
PathImpedance_ TransitPremiumBus TimeAdditiveWeight	0.00	An additive weight on premium bus submode-specific in-vehicle time (adds to TransitInVehicleTimeWeight)

SETTING	VALUE	DESCRIPTION
PathImpedance_ TransitLightRail TimeAdditiveWeight	-0.15	An additive weight on light-rail submode-specific in-vehicle time (adds to TransitInVehicleTimeWeight)
PathImpedance_ TransitCommuterRail TimeAdditiveWeight	-0.25	An additive weight on commuter rail submode-specific in-vehicle time (adds to TransitInVehicleTimeWeight)
PathImpedance_ TransitFerryTime AdditiveWeight	-0.40	An additive weight on passenger ferry submode-specific in-vehicle time (adds to TransitInVehicleTimeWeight)
PathImpedance_ BikeUseTypeSpecific DistanceFractions	False	A switch to use additional skims and weights to reflect bicycle distances on specific facility types (SACOG-specific)

SETTING	VALUE	DESCRIPTION
PathImpedance_ BikeType1Distance FractionAdditiveWeight	0.0	An additive weight on bike distance on Class 1 bike paths (adds to BikeTimeWeight, distance is converted to time)
PathImpedance_ BikeType2Distance FractionAdditiveWeight	0.0	An additive weight on bike distance on Class 2 bike paths (adds to BikeTimeWeight, distance is converted to time)
PathImpedance_ BikeType3Distance FractionAdditiveWeight	0.0	An additive weight on bike distance on Class 3 bike paths (adds to BikeTimeWeight, distance is converted to time)
PathImpedance_ BikeType4Distance FractionAdditiveWeight	0.0	An additive weight on bike distance on Class 4 bike paths (adds to BikeTimeWeight, distance is converted to time)
PathImpedance_ TransitUseFareDiscount Fractions	True	A switch (true/false) to use transit fare discount fractions based on person type and age

SETTING	VALUE	DESCRIPTION
PathImpedance_ TransitFareDiscount FractionChildUnder5	0.8	Transit fare discount fraction for children under age 5
PathImpedance_ TransitFareDiscount FractionChild5To15	0.5	Transit fare discount fraction for children ages 5 to 15
PathImpedance_ TransitFareDiscount FractionHighSchool Student	0.5	Transit fare discount fraction for high school students (children age 16+)
PathImpedance_ TransitFareDiscount FractionUniverity Student	0.5	Transit fare discount fraction for college students
PathImpedance_ TransitFareDiscount FractionAge65Up	0.5	Transit fare discount fraction for adults age 65+
PathImpedance_ TransitPassCost PercentChangeVersus Base	0	Policy input variable to change the cost of transit passes with respect to base year
Path-type Impedance Coefficients		
Coefficients_ BaseCostCoefficient PerDollar	-0.15	Base cost coefficient (per monetary unit), when income=BaseCostCoefficientIncom eLevel

SETTING	VALUE	DESCRIPTION
Coefficients_ BaseCostCoefficient IncomeLevel	30000	Household income level (monetary units per year) where the cost coefficient is the BaseCostCoefficient
Coefficients_ CostCoefficient IncomePower_Work	0.6	Power function exponent to use for adjusting the cost coefficient for income, for work tours
Coefficients_ CostCoefficient IncomePower_Other	0.3	Power function exponent to use for adjusting the cost coefficient for income, for nonwork tours
Coefficients_ MeanTimeCoefficient _Work	-0.03	Mean time coefficient (/minute) for work tours
Coefficients_ MeanTimeCoefficient _Other	-0.015	Mean time coefficient (/minute) for nonwork tours
Coefficients_ StdDeviationTime Coefficient_Work	0.8	Standard deviation of the time coefficient (/minute) for work tours, when using random VOT distribution

SETTING	VALUE	DESCRIPTION
Coefficients_ StdDeviationTime Coefficient_Other	1.0	Standard deviation of the time coefficient (/minute) for nonwork tours, when using random VOT distribution
Coefficients_ HOV2CostDivisor _Work	1.741	Divisor for the cost coefficient for the HOV2 mode for work tours (to reflect cost-sharing)
Coefficients_ HOV2CostDivisor _Other	1.741	Divisor for the cost coefficient for the HOV2 mode for nonwork tours (to reflect cost-sharing)
Coefficients_ HOV3CostDivisor _Work	2.408	Divisor for the cost coefficient for the HOV3+ mode for work tours (to reflect cost-sharing)
Coefficients_ HOV3CostDivisor _Other	2.158	Divisor for the cost coefficient for the HOV3+ mode for nonwork tours (to reflect cost-sharing)

SETTING	VALUE	DESCRIPTION
UseRandomVotDistribution	True	TRUE to randomly simulate a time coefficient for each tour, using a log-normal distribution
UrbanThreshold	500	

the science of insight

600 B Street, Suite 2202
San Diego, CA 92101
619.501.0559
www.rsginc.com

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

RSG promotes sustainable business practices that minimize negative impacts on the environment. We print all proposals and reports on recycled paper that utilizes a minimum of 30% post-consumer waste. RSG also encourages recycling of printed materials (including this document) whenever practicable. **For more information on RSG's sustainability practices, please visit www.rsginc.com**.