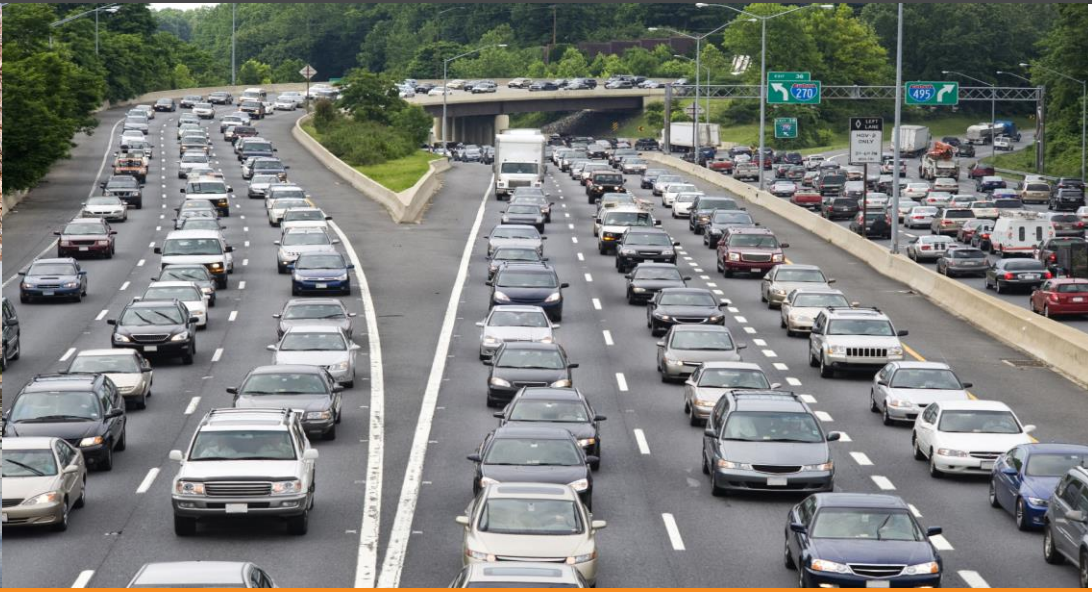
Metropolitan Washington Council of Governments

GEN3 MODEL DESIGN PLAN

Gen3 Model Development Project | July 2, 2020



PREPARED FOR:

METROPOLITAN WASHINGTON COUNCIL OF GOVERNMENTS

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

SUBMITTED BY:

RSG

WITH BASELINE MOBILITY GROUP

VERSION 1.0

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION

The Metropolitan Washington Council of Governments (MWCOG or COG) is an independent, nonprofit association that brings area leaders together to address major regional issues in the District of Columbia, suburban Maryland, and Northern Virginia. COG's membership comprises 300 elected officials from 24 local governments, the Maryland and Virginia state legislatures, and U.S. Congress. The Board of Directors is COG's governing body and is responsible for its overall policies. The National Capital Region Transportation Planning Board (NCRTPB or TPB) is the federally designated metropolitan planning organization (MPO) for metropolitan Washington. TPB is responsible for developing and carrying out a continuing, cooperative, and comprehensive transportation planning process in the metropolitan area. COG is the administrative agent for the TPB, and the TPB is staffed by COG's Department of Transportation Planning (DTP). The TPB staff, with some consultant assistance, develops, maintains, applies, and improves the TPB's family of regional travel demand forecasting models, which are used for regional, long-range transportation planning in the metropolitan Washington region. These regional travel demand models are developed under the guidance of the Travel Forecasting Subcommittee (TFS).

COG's current, production-use travel demand forecasting model is an aggregate, trip-based four-step model, known as the Generation-2 (Gen2), Ver. 2.3 Model. In 2018, following its strategic plan,[1] COG/TPB staff set out to develop a next-generation travel demand model, to be known as the Generation-3, or Gen3, Model. COG issued a request for information (RFI)[2] and a request for proposals (RFP)[3] to develop the Gen3 Model.

A team consisting of RSG and Baseline Mobility Group was selected for this project. This document is the main deliverable under Task Order 2 of that project: a Gen3 Model Design. **The project team recommends that COG transition from its current aggregate, trip-based travel demand model to a simplified activity-based model (ABM) implemented in the open-source ActivitySim software platform**. This report consists of an assessment of the strengths and weaknesses of the current MWCOG travel demand model, a proposed model form (ABM) and model application software (ActivitySim), a description of how that model design meets the project objectives, an outline of the data required for model development and

---

[1] "Strategic Plan for Model Development, Task Order 15.2, Report 3 of 3," Final Report (Washington, D.C.: Metropolitan Washington Council of Governments, National Capital Region Transportation Planning Board, October 15, 2015).

[2] "Request For Information No. 18-001, TPB Travel Demand Forecasting Model, Generation 3/NextGen" (Washington, D.C.: Metropolitan Washington Council of Governments, National Capital Region Transportation Planning Board, May 31, 2018).

[3] "Consultant Assistance to Develop the Next-Generation Travel Demand Forecasting Model, Known as the Gen3 Model, for the Metropolitan Washington Council of Governments and the National Capital Region Transportation Planning Board -- RFP #19-015," Request for Proposals (RFP) (Washington, D.C., May 29, 2019).

application, a description of the software system used to implement the model, quality assurance and quality control (QA/QC) procedures for model development and application, and other requirements outlined in the product requirements document released as part of the RFP and described further below.[4] The document also provides a plan for model development and associated costs with various activities and features.

It should be noted that a model design is a blueprint for model development. Its purpose is to provide enough information to the model development team and stakeholders to achieve a common understanding of the key features of the model system, how the model meets the product requirements, and the tasks to be undertaken to deliver a production-ready model by the end of the contract (end of calendar year 2022). The document does not describe every model formulation, coefficient, input, and output. **These details will be provided in subsequent model estimation, calibration and validation documentation and a model user's guide, to be delivered under separate task order(s)**.

---

[4] Mark S. Moran, "Product Requirements Document for the TPB Travel Demand Forecasting Model, Generation 3, the Next-Generation Model" (Washington, D.C.: National Capital Region Transportation Planning Board, Metropolitan Washington Council of Governments, May 25, 2018), https://www.mwcog.org/assets/1/25/Product_Requirements.pdf.

2.0 OVERVIEW OF GEN 3 PRODUCT REQUIREMENTS

Metropolitan Washington is a complex region and travelers are offered numerous transportation options. On the highway side, these include high-occupancy vehicle (HOV) lanes, high-occupancy/toll (HOT) lanes, and toll roads, both fixed price (e.g., Dulles Toll Road) and variably priced (e.g., the Intercounty Connector, or ICC). Some workers participate in casual carpooling, referred to in the DC area as “slugging.” On the transit side, these include local bus, express bus, bus rapid transit, streetcar, light-rail, heavy/urban rail, and commuter rail. As in many regions, non-motorized travel makes up a small but, nonetheless sizeable share of trips, with higher percentages of walking in the urban core. Mobility-as-a-service (MaaS) is increasingly being used as an alternative to private and public modes of transportation, including ride-hailing companies, also known as Transportation Network Companies (TNCs), car-on-demand, e-bike, bike-share, and e-scooter modes.

In addition to the resident and freight travel that is typically captured in a travel demand model, the region also has several unique travel markets, as shown in Table 1. There are three major commercial airports in the MWCOG region: Dulles International Airport (IAD), with approximately 7.2 million annual local enplanements, Reagan National Airport (DCA), with approximately 10.5 million local annual enplanements, and Baltimore/Washington International Airport (BWI), with approximately 8.9 million local annual enplanements.[5] There are at least 7 major universities in the MWCOG region, with combined attendance of approximately 145,000 students.[6] The District of Columbia attracts approximately 22.8 million annual visitors, who stay for an average of approximately 2.7 nights.[7] Internal-external and through travel are important components of demand. An analysis of 2015 American Community Survey indicates that approximately 6% of workers who reside inside the MWCOG modeled area have a regular workplace location outside the region, and 9% of the workers with a regular workplace inside the MWCOG modeling area reside outside the region.

TABLE 1: MWCOG SPECIAL TRAVEL MARKETS

Market	Site	Size	Type
Airports	Dulles International Airport (IAD)	7,245,000	2017 Local Annual Enplanements
	Reagan National Airport (DCA)	10,499,000	2017 Local Annual Enplanements

---

[5] Annual local enplanements from National Capital Region Transportation Planning Board. "Washington-Baltimore Regional Air Passenger Survey–2017 General Findings", June 2018.

[6] www.collegeraptor.com, accessed April 13, 2020.

[7] https://washington-org.s3.amazonaws.com/s3fs-public/2017_washington_dc_visitor_statistics_-_destination_dc.pdf, accessed April 13, 2020.

Market	Site	Size	Type
	Baltimore/Washington International Airport (BWI)	8,910,000	2017 Local Annual Enplanements
Major Universities	University of Maryland	37,610	2017 Enrollment
	George Mason University	37,316	Enrollment
	University of District of Columbia	4,688	2017 Enrollment
	American University	14,311	Enrollment
	Georgetown University	18,459	2018 Enrollment
	George Washington University	25,613	2017 Enrollment
	Howard University	6,276	2017-18 Enrollment
Overnight Visitors	District of Columbia	22,800,000	2017 annual visitors
Internal/ External Trips	IE Workers	225,622	2015 workers
	EI Workers	324,441	2015 workers

Below is a list of some of the policies that are important in the metropolitan Washington region, and, hence, should ideally be addressed in the Gen3 Model. Some of these were identified by TPB staff, others by modeling stakeholders. Many of these are discussed in more detail later in this report.

- Modeling of transit and transit sub-modes (e.g., bus versus light rail)

    - Mode choice and path-building: The trend has been to move some of this modeling of transit sub-modes out of mode choice and into path building
    - Transit assignment
        - All-or-nothing versus capacity restrained
        - Production/attraction format versus origin/destination format
        - Transit crowding. Even though there have been some declines in transit ridership in recent years, transit crowding/capacity, on both rail and some bus lines, remains an issue, especially in the long term.
- Modeling highway travel (private-use cars and trucks)
- Highway assignment
    - Very long run times to reach acceptable levels of convergence
    - Modeling HOV lanes, HOT lanes, and other managed-lane facilities
- Modeling non-motorized modes (walk and bike)
- Assessing the effect of land development patterns and job/housing balance on transportation system performance
- Estimating the impacts of infill development on mode share/choice, particularly with regards to walk and bike modes
- Modeling the effect of the employer-based transit subsidies that some workers, especially federal, currently receive
- Telework, which has risen substantially over the past decade[8]
- Increasing use of transportation network companies (TNCs) and other shared-mobility modes, including their effect on competing modes of travel
- Visitor/tourist travel: The Washington region receives many visitors; due in part to its role as the nation’s capital.
- Modeling peak spreading: Addressing the duration of the peak period, as opposed to focusing simply on the peak-hour condition
- Modeling the impact of travel time reliability (typically difficult to do with regional travel demand models)
- Representing/conveying the level of uncertainty in model inputs and outputs
- Impact of connected/autonomous vehicles (CAVs) in the coming years
- Modeling the impact of travel behavior of subsets of population, particularly those who are economically disadvantaged, such as for the purposes of environmental justice (EJ)/social equity
- Freight planning. Although the Washington, D.C. area is not considered a major freight city, freight and commercial vehicles are still an important segment of the travel market.
- Greenhouse gas analyses (identified by modeling stakeholders, tracked by COG/TPB staff, and the subject of the Transportation Climate Initiative)[9]

---

[8] During the coronavirus (COVID-19) pandemic, telework has risen dramatically. However, regional travel demand models are designed to make long-term predictions (5-25 years). Consequently, regional travel demand models are not generally designed to represent short-term (ca. one or two year) variations in travel.

[9] Transportation and Climate Initiative, “Homepage,” Transportation and Climate Initiative (TCI), 2019, https://transportationandclimate.org/.

- Effect of Internet on travel (identified by modeling stakeholders)
- Traffic microsimulation (identified by modeling stakeholders, though rarely modeled at the regional level, due to computing limitations)
- Modeling first/last-mile travel, transit access and transfers.
- Parking needs and impacts, e.g., related to environmental impacts, drive-access to transit (PNR lots), the internet of things (IoT). As an example, in Montgomery County, Maryland's General Plan Update, "Thrive Montgomery 2050," proposed policies are aimed at greatly reducing auto trips, reducing parking, including converting current parking lots and PNR lots in urban areas.

Additionally, modeling stakeholders noted several areas that they would like to see improved in the model:

- Improved ease of adapting the regional model for sub-regional travel analyses
- Improved ease of use
- Shorter model run times

Like all MPOs, the TPB must develop a Unified Planning Work Program (UPWP), which is updated on an annual basis and lists all the planning activities that the MPO staff will undertake for a given year. The TPB travel demand forecasting model (TDFM) is used for several of the tasks specified in the UPWP, including the following:

- Development of a Long-Range Transportation Plan (LRTP), which must extend at least 20 years into the future. The 2018 version of the TPB's LRTP is called Visualize 2045. The MPO must also designate a subset of the LRTP that is financially constrained to reasonably expected future revenues. In the past, TPB called this the constrained, long-range plan (CLRP). But, the new nomenclature is the constrained element of the LRTP (Visualize 2045).
- Assessment of the performance of the LRTP, both in general terms of interest to the MPO and in more specific terms dictated by performance-based planning and programming (PBPP).
- Air Quality Conformity Determination, since the metropolitan Washington area is a non-attainment area for one or more air pollutants.
- Regional scenario studies, where changes are made to one or more of the following: transportation networks, land use, or policy assumptions.
- Transportation-related corridor studies and project planning studies. Although these types of studies are often conducted by state and local governments (and their consultants), the TPB staff does perform these types of studies under technical assistance projects that are conducted by TPB staff for DC, Maryland, Virginia, and the Washington Metropolitan Area Transit Authority (WMATA or Metro).
- Analyses of the impacts of transportation projects and policies on environmental justice (EJ)/social equity. Much of this analysis is currently done using the aggregate, trip-based regional travel model.
- Parking needs and impacts, e.g., related to environmental impacts, drive-access to transit (PNR lots), the internet of things (IoT). As an example, in Montgomery County,

Maryland’s General Plan Update, “Thrive Montgomery 2050,” proposed policies are aimed at greatly reducing auto trips, reducing parking, including converting current parking lots and PNR lots in urban areas.

3.0 STRENGTHS AND WEAKNESSES OF THE CURRENT TRAVEL MODEL

3.1 | VERSION 2.3 MODEL OVERVIEW

MWCOG maintains at least two travel demand models: a production-use model, currently the Version 2.3 Model,[10] and one or more developmental models, such as the Version 2.4 Model. The Version 2.3 Travel Model became the adopted regional travel model on November 16, 2011. It operates with a 3,722 Transportation Analysis Zone (TAZ) system covering an area that encompasses 22 jurisdictions and extends over the District of Columbia and portions of three states: Maryland, Virginia, and West Virginia (see Figure 1). The model was calibrated to the 2007/08 Household Travel Survey, which included 11,400 households throughout the COG region. There were approximately 20 million expanded internal-internal person trips represented in the survey.

Inputs to the model include land-use data and transport networks. Land-use data is specified at the TAZ level, and includes households, household population, group quarters population, employment by category (i.e., retail, office, industrial, and other), zonal area, income index, jurisdiction code, airline distance to the nearest external station, and centroid XY coordinates. Transport networks consist of highway and transit networks.

The demographic forecasting model first splits households by TAZ into four household income groups and four household size groups using fitted distributions to Census data. A multinomial logit household vehicle availability model is used to further split the household distribution by four vehicle availability groups. This model also includes a peak-period transit accessibility variable, an area-type variable,[11] and a DC-specific constant term.

Trip generation cross-classification models are next applied to compute daily person trip productions (motorized and non-motorized) by five trip purposes: Home-Based Work (HBW), Home-Based Shop (HBS), Home-Based Other (HBO), Non-Home-Based Work (NHW), and Non-Home-Based Other (NHO). Cross classification models consider demographic variables household size, income, and vehicle availability. The HBO purpose incudes both K-12 and college/university school trips. Trip attractions are estimated using regression equations that are stratified by trip purpose and area type. The explanatory variables include total employment, retail employment, office employment, other employment, and total population. Home-based trip attractions are disaggregated by household income, based on percentages by area type.

A commercial vehicle purpose (consisting of both autos and light duty trucks), and two truck types, Medium and Heavy, are also modeled. Medium trucks are those with two axles and 6

---

[10] Ronald Milone et al., Calibration Report for the TPB Travel Forecasting Model, Version 2.3, on the 3,722-Zone Area System, Final Report (Washington, D.C.: Metropolitan Washington Council of Governments, National Capital Region Transportation Planning Board, January 20, 2012

[11] There are six area types, coded based on one-mile “floating” calculations of population and employment density.

tires. Heavy trucks represent all combination vehicles. Truck and commercial vehicle trip ends are estimated based on employment by type, total households, and area type.

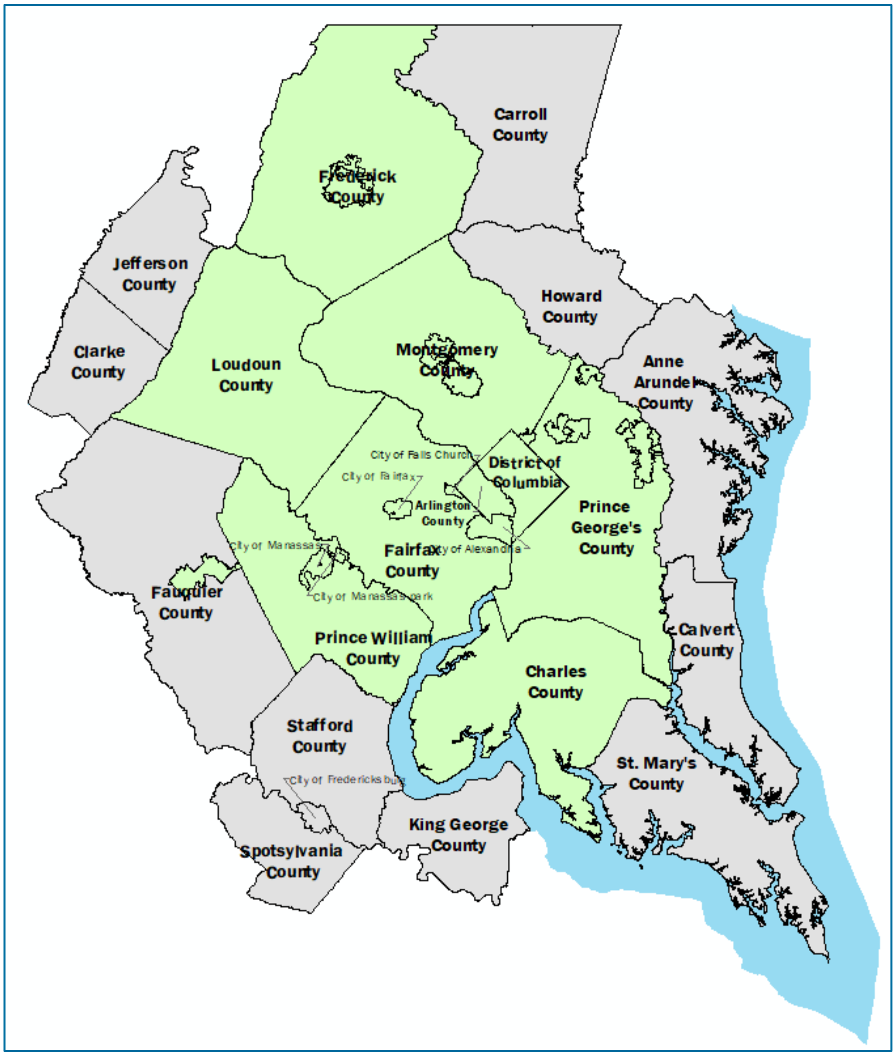


FIGURE 1: MWCOG VERSION 2.3 MODEL STUDY AREA (GREEN AND GRAY) AND TPB MEMBER JURISDICTIONS (GREEN)

Non-motorized trips are removed from the estimated productions and attractions prior to trip distribution, using linear regression (high density area types 1 and 2) or percentages (area types 3-6) that vary based on trip purpose and area type. Also, internal-external trips are removed from total productions, using an equation that takes into account distance from the nearest external station, and a variable indicating whether the TAZ is in the part of the region between Baltimore and Washington, D.C., which has different rates of IE trips than other parts of the MWCOG region.

Trip distribution is implemented as a gravity model that considers auto time, toll cost and Metrorail transit time in a harmonic mean formulation in which transit time is weighted by the base-year share of transit trips by trip purpose and household income. The HBW purposes use peak-period travel skims (associated with the A.M. period) and the other purposes use off-peak-period travel skims (associated with the midday period). Internal-internal home-based trip purposes are stratified by household income. Internal-internal non-home-based trips, commercial vehicles, and trucks are not stratified. External-internal person trip models are stratified by trips on interstates versus arterials, while commercial vehicles and trucks are not stratified.

Prior to the execution of the mode choice model, a model generates zonal files containing zonal parking costs and highway terminal times (the time to park and “un-park” a vehicle).[12] Daily and hourly parking costs are a function of area type and density. Terminal time is a function of area type.

Mode choice models are nested logit with three auto alternatives (drive-alone, shared 2, and shared 3+) and four transit alternatives (all bus, all Metrorail, bus/Metrorail, and commuter rail) differentiated by three access modes (walk, park-and-ride, and kiss-and-ride) for a total of 15 modes. The nesting structure is shown in Figure 2. The model separates paths that are bus-only from paths that are Metrorail only and paths that include both bus and Metrorail. Any transit path that includes commuter rail is treated as a commuter rail mode choice. Note that light-rail transit (LRT) is typically included in the Metrorail alternatives and both streetcar and bus-rapid transit (BRT) technologies are included in the bus alternatives. Like trip distribution, the HBW trip purpose uses A.M. peak period travel skims and the other purposes use midday period travel skims. Transit skims are built using TRNBUILD shortest paths, though combined headways are calculated within modes. Walk market segmentation is used to compensate for zone size aggregation bias. The model uses three walk markets: short (0.5 mile), long (1.0 mile), or none. The model was calibrated to 2007/8 Household Travel Survey data, 2005-8 transit on-board survey data and 2000 Census journey-to-work data. Alternative-specific constants are stratified by trip purpose, household income (home-based purposes only) and seven geographic “superdistrict” areas.

---

[12] Metropolitan Washington Council of Governments (COG) National Capital Region Transportation Planning Board (TPB). User’s Guide for the COG/TPB Travel Demand Forecasting Model, Version 2.3.75 Volume 1 of 2: Main Report and Appendix A (Flowcharts). December 5, 2018.

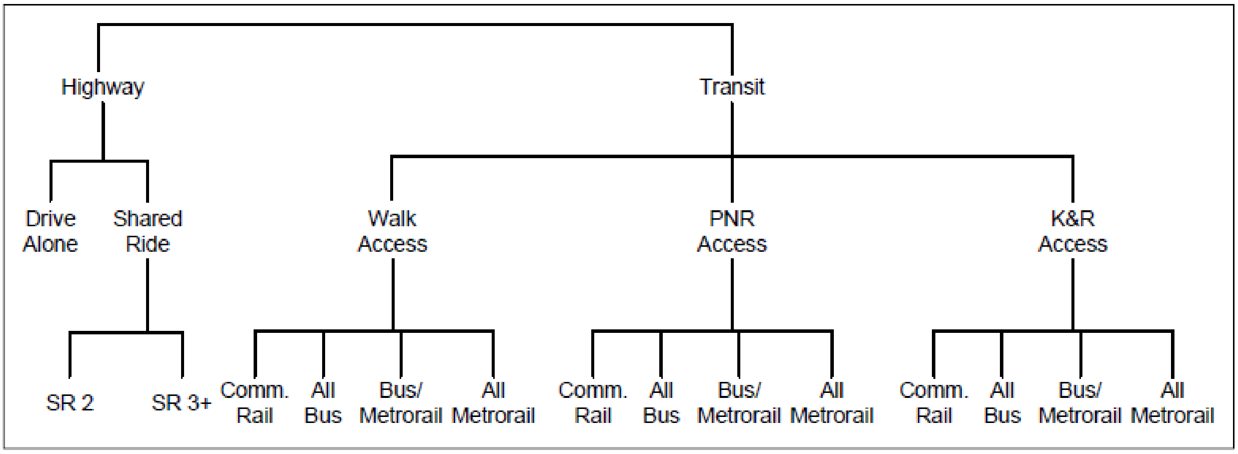


FIGURE 2: MWCOG VERSION 2.3 MODE CHOICE MODEL STRUCTURE

Between mode choice and assignment, trips are factored by time-of-day. The static factors are based on trip purpose and direction and are based on a trip-in-motion analysis of the 2007/8 Household Travel Survey data. The time-of-day model creates auto trip tables in origin-destination format for four time periods; A.M. peak period (6 – 9 A.M.), the midday period (9 A.M. – 3 P.M.), the PM peak period (3 P.M. – 7 P.M.) and the nighttime/early morning period (7 P.M. – 6 A.M.).

Auto assignment is performed for all four time periods listed above, using a bi-conjugate Frank-Wolfe algorithm run to a relative gap of 10^-4 (0.0001) or 1000 user equilibrium iterations (whichever is attained first).[13] The model assigns six user classes (drive-alone, shared 2, shared 3+, commercial vehicles, medium & heavy trucks, and airport driver trips) and uses a conical volume-delay function. Free-flow speeds and per-lane-hour capacities are calculated based on lookup tables by facility type and area type. Passenger car equivalents are not currently used in the assignment process. The A.M. and P.M. traffic assignments are performed sequentially where first, non-HOV3+ demand is assigned to general purpose lanes. This is followed by another assignment in which HOV3+ demand is assigned to all facilities. The model calibration report suggests that the sequential assignment results in a closer match to volumes on HOV/HOT traffic on the Capital Beltway in Virginia and the I-395 Shirley Highway.

---

[13] Ray Ngo, Feng Xie, and Mark S. Moran, “User’s Guide for the COG/TPB Travel Demand Forecasting Model, Version 2.3.78” (Washington, D.C.: Metropolitan Washington Council of Governments, National Capital Region Transportation Planning Board, April 14, 2020), 204, https://www.mwcog.org/transportation/data-and-tools/modeling/model-documentation/.

The model system is iterated five times, including an initial assignment and four subsequent speed-feedback iterations. The method of successive averages is used to average link volumes which calculate travel times for the next iteration of the model.

The Ver. 2.3 Model has been validated three times: to 2007, 2010, and 2014 conditions. For the purposes of this report, we summarize the 2007 validation results.[14] The model was validated to VMT reported in state Highway Performance Monitoring System (HPMS) summaries by jurisdiction, as well as traffic counts and transit boardings. The model matches VMT by state very well; within 1% of each state's estimated VMT. As one might expect, there is more variation in goodness-of-fit at the sub-state jurisdiction level, with estimates ranging from 66% to 125% of observed. The model estimated flows were also compared to counts at screenline locations, though the analysis is muddied by the fact that not all links crossing each screenline are counted. On average only 57% of links crossing screenlines have counted volume, though one might expect that the percent would increase with respect to volume. The overall percent root mean square error for the model is 43%.

On the transit side, estimated Metrorail productions and attractions were compared to observed productions and attractions by route and segment (aka station group). Overall total Metrorail productions and attractions are under-estimated by 7%, while they are under-estimated by 12% for the District of Columbia stations. The difference in productions and attractions by station group range from 65% to 148% of observed.

The Ver. 2.3 Travel Demand Model also includes an automated script that can be used to calculate tolls on toll facilities and managed lanes. The script starts with assumed lower-bound toll costs of 20 peak and 15 off-peak cents per mile and adjust those costs upwards until the tolled lanes operate at a volume/capacity (V/C) ratio of 0.95 or better.[15] This can be considered the boundary V/C ratio designating a level-of-service between D and E. The toll adjustment script is typically run by COG staff and then the “converged” toll price file is provided to model users as a standard input. COG staff found that the toll prices estimated by the script approach the actual toll costs, though they are somewhat lower than actual toll costs.

There are three special markets considered in the model besides commercial vehicles and trucks. These exogenous travel markets consist of taxis, school, and visitor/tourist auto driver trips (collectively referred to as “miscellaneous trips”) and airport-passenger auto driver trips. The miscellaneous trip totals, shown by year on Table 8, are based on surveyed travel patterns that have been growth factored through time. The airport-passenger forecasts are based on the

---

[14] Milone, Ronald, Hamid Humeida, Maria Martchouk, Mark Moran, and Meseret Seifu. *Calibration Report for the TPB Travel Forecasting Model, Version 2.3, on the 3,722-Zone Area System*. Final Report. Washington, D.C.: Metropolitan Washington Council of Governments, National Capital Region Transportation Planning Board, January 20, 2012.

[15] Jinchul Park to Files, “HOT Lane Modeling Process of MWCOG/TPB (Draft),” Memorandum, October 12, 2012.

2011 Washington-Baltimore Regional Air Passenger Survey.[16] The trip tables represent auto travel to each of the three major airports serving the Washington/Baltimore area.

3.2 | VERSION 2.5 MODEL OVERVIEW

The MWCOG Version 2.5 Travel Model was a developmental model, developed in 2017[17] and tested by MWCOG staff for two years, but which was never brought into production use. The Ver. 2.5 Model differs from Ver. 2.3 Model in the following ways:

1) The area-type specific factors used to estimate non-motorized productions and attractions in Ver. 2.3 were replaced with a disaggregate model that takes into account more explanatory variables.
2) The transit path-builder used in Ver. 2.3 (TRNBUILD) was replaced with Cube Public Transport (PT). PT differs from TRNBUILD in a number of ways, but the key difference is that PT is a multipath transit path-builder in which level-of-service between any given TAZ pair is represented as a weighted average of multiple transit "hyperpaths,"[18] and estimated transit trips between zone pairs are split across transit hyperpaths in proportion to their weights. It should be noted, though, that the Ver. 2.5 Model implemented the "BestPathOnly" mode in PT (to be comparable to the Ver. 2.3 Model), despite the fact that PT is mainly designed as a multipath transit path-builder.
3) Flattened mode choice. Instead of having a complex mode choice model, where transit sub-mode choice is determined in the mode choice model, the Ver. 2.5 Model used a "flattened mode choice" paradigm, which features a simplified mode choice model and divides transit sub-mode choice across both mode choice and transit path-building. Thus, the transit portion of the mode choice model was simplified: Transit technologies (also referred to as transit sub-modes or line-haul modes) such as commuter rail, Metrorail, and bus, were removed from the mode choice model. Together with the simplification of the mode choice model, the PT transit path-builder is relied on to model the choice of transit sub-mode/technologies. The transit mode of access choice was retained in the mode choice model. The revised mode choice model eliminated geographic, alternative-specific constants.

---

[16] Ngo, Xie, and Moran, "User's Guide for the COG/TPB Travel Demand Forecasting Model, Version 2.3.78," 214.

[17] Cambridge Systematics, Inc. and Gallop Corporation, "FY 17 Task Orders," Final Report (Metropolitan Washington Council of Governments, National Capital Region Transportation Planning Board, June 30, 2017), https://www.mwcog.org/file.aspx?&A=YiUe54YhmPVA0q1IahkVpmf4CjB%2fkVfhr3mZDJJ1ACM%3d.

[18] The definition of a hyperpath is "an acyclic subnetwork with at least one link connecting the origin to the destination, and where at each node, there are probabilities for choosing the alternative links". For more information, see Sang Nguyen, Stefano Pallottino, Michel Gendreau. Implicit Enumeration of Hyperpaths in a Logit Model for Transit Networks, Transportation Science, Vol. 32, No. 1. February 1998.

4) Value-of-time segmentation was added to highway assignment, increasing the number of trip tables assigned in each time period and the number of level-of-service matrices skimmed. This was done to increase the sensitivity of the model to priced transportation infrastructure.
5) The conical volume-delay functions were replaced with modified Bureau of Public Roads (BPR) functions on freeways in an attempt to increase the accuracy of travel times predicted by the model.

MWCOG staff spent approximately 2 years testing and debugging the Ver. 2.5 Travel Demand Model, making 13 updates to the model. In the end, however, COG staff decided not to adopt the model for use in production. Thus far COG staff feel that the benefits of the Ver. 2.5 Model do not outweigh the costs, particularly increased runtime (mostly due to the introduction of value-of-time segmentation in highway skimming and assignment) and complexity.[19] While the highway validation of the Ver. 2.5 Model is roughly comparable to that of the 2.3 Model, transit boardings are significantly under-estimated, especially on MARC, VRE, and local bus modes.[20] The Ver. 2.5 Model runtime is twice that of Ver. 2.3. Given the underwhelming performance of the Ver. 2.5 Model, COG staff chose not to perform dynamic validation of the model.

3.3 | STRENGTHS AND WEAKNESSES OF CURRENT MODELING SYSTEM

Since model version 2.3 is the current production version of the MWCOG travel demand model, we will focus on that version of the model for this review. Our assessment of the current model system is that the MWCOG Ver. 2.3 Model represents state of the practice in trip-based modeling, and many of the weaknesses of the model apply to most or all trip-based models. Our assessment is based on the following key points:

1) The auto ownership model is sensitive to area type and transit accessibility. However, auto ownership influences only trip generation; auto ownership does not affect mode choice. Many regional travel demand models use auto ownership as a segmentation variable in mode choice in order to reflect the influence of transit investments on the size of the transit-dependent market (e.g. trips in 0-auto households and households where the number of autos is less than workers or licensed drivers).
2) The demographic stratification (household size, household income, vehicle availability) used in trip generation is consistent with other regions. The model represents a commonly used range of trip purposes, with non-home-based trips broken out by work

---

[19] See, for example, Mark S. Moran, “Status Report on the TPB’s Developmental Travel Demand Forecasting Models” (November 30, 2018 meeting of the COG/TPB Travel Forecasting Subcommittee, held at the Metropolitan Washington Council of Governments, Washington, D.C., November 30, 2018), 7–9.

[20] Ron Milone. Ver 2.5 Travel Demand Model Development Status Report. TPB Travel Forecasting Subcommittee March 15, 2019.

versus other. However, school and university trips are not represented explicitly. This may reduce the usefulness of the model for evaluating certain types of transportation infrastructure and policies, particularly around major universities.

3) The use of a composite impedance term in trip distribution that considers auto time, toll cost, and Metrorail transit time is a reasonable alternative to use of a mode choice logsum[21] in a destination choice model. The disadvantage of this approach is that it is not fully consistent with mode choice model variables or parameters, and the shares used in the composite impedance term are based on regional mode share rather than market-specific probabilities. The model appears to be well-calibrated to trip length frequency distributions and district-level scatterplots demonstrate a good match to observed travel patterns.

4) The parking cost model is a convenient tool to estimate parking costs and reduce model maintenance burden. However, the model is relatively simple in that it is based on only total employment density and area type. We note that the maximum daily parking cost indicated in the model user guide may be low ($12/day) compared to current price of parking in downtown DC,[22] possibly due to incorporation of industrial data in total employment or because the parking cost model represents an average parking cost paid across all workers.

5) The mode choice model has reasonable parameters. The differentiation of transit by technology (bus-only, Metrorail-only, bus + Metrorail, and commuter rail) is a good compromise between more a simple local versus premium structure and a complicated structure with an alternative for every transit sub-mode. The use of walk-access transit segmentation is a useful way to mitigate spatial aggregation bias, particularly for larger zones. The model was well calibrated to a combination of random household travel survey data and choice-based, on-board survey data. The drawbacks of the model include the lack of non-motorized modes (which are extracted prior to trip distribution), the lack of vehicle availability as a segmentation variable, and the use of geographic constants which tend to modify the sensitivity of the model to changes in transit level of service. The mode choice documentation would be improved by inclusion of a few estimated versus observed summaries of trips by mode and trip length, transit trip distance by access mode, and number of transfers, compared to on-board survey data.

6) The use of static time-of-day factors to convert trip tables from production-attraction format to origin-destination format by time period is a common limitation in trip-based models. However, like most trip-based models, the Ver. 2.3 Model has limited sensitivities to policies and level-of-service that varies by time of day. Peak-period level-of-service affects only HBW trip distribution and mode choice, while off-peak-period

---

[21] The logsum is a measure of composite utility or expected maximum utility of a choice set. It is equal to the log of the denominator of a multinomial logit model.

[22] https://en.parkopedia.com/parking/washington_dc/?arriving=202004221930&leaving=202004222130, accessed April 20, 2020.

level-of-service affects only non-HBW purposes. The percent of trips in each time period is not affected by changes in congestion.

7) Four time-of-day periods for assignment is consistent with best practices and the estimation of peak-period delay. The Bi-conjugate Frank Wolfe assignment algorithm used is state-of-the-practice. Highway assignment appears to be run to a sufficient relative gap. The values of time used in the model are based on careful analysis of household income and trip share by purpose in each period. The sequential assignment approach used to assign non-HOV 3+ and then HOV3+ demand is unique and we are not prepared to comment on the value of the approach compared to a simultaneous multi-class assignment. It would be useful to compare the results and model runtime with and without pre-assignment.

8) The toll optimization algorithm represents best practice, though it may be useful to optimize the algorithm by adjusting assignment iterations between toll adjustments. It may also be useful to test changing the stopping criteria for the algorithm from an LOS E threshold (V/C ratio of 0.95 to 1.00) to lower thresholds, such as an LOS D/E threshold (V/C ratio of 0.90-0.95) and an LOS C/D threshold (V/C ratio of 0.8 to 0.85). This would depend on the policy of the toll road authority to manage to a specific target speed on the facility (e.g. 55 miles per hour) and the V/C ratio at which the speed estimated by the volume-delay function degrades below that threshold.

9) Path weights used in transit assignment are consistent with the mode choice model. According to information provided by MWCOG staff, the model system may over-estimate potential future Metrorail ridership because the skimming and assignment process does not take into account transit capacity restraint, which could result in over-estimated rail ridership in the future. The model under-estimates ridership on commuter rail lines, though this is a small overall portion of transit demand with only about 50k boardings per day (about 4% of transit ridership).

10) The use of fixed auto trip tables for representation of special markets is suited to accounting for traffic impacts of the markets but ignores non-motorized and transit modes and is incapable of analyzing market-specific policies that affect them such as parking pricing, provision of transit, etc. The special market trip tables are based on relatively old data, and there may be some "double-counting" of university-related travel due to the inclusion of university auto trips in both the exogenous trip tables and in the Home-based Other trip tables, though it is unlikely that the HBO trip tables would model the correct attraction end of these trips due to lack of enrollment data considered in the model.

Key Model Strengths

Based on the above assessment, we summarize the current model strengths as follows.

1. Model usability: The model has been successfully applied by MWCOG staff and TPB-member agencies for key planning activities, including:

   a. Assessment of investments in infrastructure including highway and transit alternatives, including:
      i. Highway and transit capacity changes, including toll roads, high occupancy vehicle (HOV) lanes, high occupancy toll (HOT) lanes, transit investments, etc.
      ii. Estimation of benefits of the above investments for the Long-Range Transportation Plan and Transportation Improvement Program.
      iii. Air Quality Conformity Determination.
   b. The model is well-documented, with a thorough user guide.
   c. Model runtime is reasonable given the size and complexity of the MWCOG region.
   d. Networks are maintained in customized software (COGTools) with an intuitive user interface and scenario management capabilities.
   e. Supplemental scripts provide additional capabilities (such as toll cost adjustment) and summaries.
   f. GitHub is used by model development staff for code versioning, though not all staff are using version control.
2. Model Sensitivity: The model is sensitive to the key inputs in the following ways:
   a. Changes in total households and employment by TAZ, and density of development, on the magnitude and cost of travel.
   b. The effects of key household socio-demographic variables (size, income and vehicle ownership) on trip generation, and a subset of these variables on trip distribution and mode choice.
   c. The effects of auto travel time, cost, and Metrorail transit time on trip distribution.
   d. The effects of travel time and cost on mode choice, including, for auto modes, travel time, terminal time, parking cost and toll charges, and for transit modes, changes in transit service frequency and time, provision and cost of parking at transit stations, differences in transit technology, and fare.
   e. The effects of congestion and cost on route choice, mode choice, and trip distribution.
3. Model Credibility: The model has been generally well calibrated and validated, as noted above and in various technical reports.

Key Model Weaknesses[23]

1. Model Usability
    a. The model is incapable of addressing some of the features of importance to MWCOG. For example:
        i. Transit crowding. Transit capacity is not considered by the model.
        ii. External transit travel. The model does not address transit trips entering or leaving the region, which means it has difficulty of providing good estimates of transit travel to Washington Union Station and to the area’s three commercial airports.
        iii. Modeling non-motorized modes (walk and bike). Although walk is one of the modes of access to transit, and non-motorized travel is included in trip generation, non-motorized trips are extracted after trip generation, so are not used in trip distribution or mode choice. The model does not consider the provision of bicycle infrastructure on travel demand.
        iv. Employer-based transit subsidies. The model does not consider transit subsidies for specific groups of workers.
        v. Modeling the effect of travel time unreliability in travel behavior. Unreliability is not modeled.
        vi. Telework is not modeled.
        vii. Transportation Network Companies (TNCs) and other shared mobility modes are not considered. This shortcoming is due to the fact that the Ver. 2.3 Model was calibrated using household and transit surveys from 2007 and 2008, when TNCs did not yet exist.
        viii. Visitor/tourist travel: Only auto trips are represented, the trip tables are based on old data, and the trip tables are insensitive to key inputs such as hotel rooms, changes in visitor attractions, etc.
        ix. Connected/autonomous vehicles (CAVs) are not modeled.
        x. Micro-mobility is not modeled (e.g., e-bike, e-scooter)
2. Model Sensitivity. The model system suffers from typical limitations of trip-based models that have been reported in the literature:
    a. Various types of aggregation biases, such as
        i. Temporal. Only HBW trips are exposed to peak level of service, non-work trip distribution and mode choice is based on off-peak level of service. The models do not represent the effect of congestion on the timing of

---

[23] See also Table 8 (p. 47) Moran, “Product Requirements Document for the TPB Travel Demand Forecasting Model, Generation 3, the Next-Generation Model.”

    trips, therefore peak spreading is not considered. In many trip-based models, accessibilities do not affect trip generation, therefore induced demand related to the frequency of travel is not explicitly represented. In the COG Ver. 2.3 Model, however, trip distribution is influenced by peak-period transit accessibility, so it has a better representation of induced travel than trip-based models that do not consider transit in trip distribution, but it does not include the accessibilities from other non-transit modes.

    ii. Spatial. The smallest unit of space in the model is the TAZ, which tend to be larger for zones further from the center of the region. Large TAZs make it more difficult to measure the effects of density on travel, and represent non-motorized (walk and bike) travel times in the model. Sub-zones used in mode choice partially compensate for this bias.

    iii. Markets. Three socio-demographic variables are used in trip generation, only household income affects trip distribution and mode choice. Trip-based aggregate models do not consider person variables such as age, and using additional variables adds computational burden which leads to unacceptable increases in runtime. This has implications for testing certain policies, such as parking pricing, in which every trip is exposed to an average parking cost, whereas, in reality, some travelers have parking provided by their employer and some travelers see the full cost of parking.

  b. Independence of trips. This limitation leads to potentially inconsistent changes between home-based and non-home-based trips, inability of the model to consider variables that affect activity duration, and lack of socio-economic variables in non-home-based trip models.

  c. Treatment of special markets. Trip tables representing overnight visitors, university students, and other special markets are based on old data and not responsive to key variables or changes in land-use and transport supply over time.

  d. Given the aggregate nature of trip-based models, difficulty to segment on specific population groups, which could be useful for social equity analyses.

3. Model credibility. Although the model is generally well-calibrated and validated, transit ridership on certain lines and modes is under-estimated and certain highway screenlines could be improved. Model validation has focused on cross-sectional validation where model outputs are compared to traffic counts, travel times, and transit volumes for a given set of inputs. Dynamic validation, in which key inputs are systematically varied to observe model elasticities and response surfaces, should be performed to ensure reasonable longitudinal sensitivities.

3.4 | RECOMMENDED MODEL FORM

We propose that MWCOG move from its aggregate, trip-based (four-step) travel demand model to a simplified activity-based model (ABM), implemented in the open-source ActivitySim modeling platform, in order to meet the product requirements described above and in the Product Requirements Document. ABMs are disaggregate models, meaning they operate at the person and household level, and are tour-based. A tour is sequence of connected trips. Tour-based models can guarantee a level of consistency across trips in a tour that is not possible in an aggregate trip-based model. Below, we provide a brief introduction to ActivitySim. We then describe how ActivitySim meets the product requirements and addresses shortcomings in the MWCOG Version 2.3 trip-based travel demand model. We provide an overview of ActivitySim and a description of the key dimensions of the system - treatment of time, treatment of space, and person types. We then provide a model development plan and resources requirements.

4.0 INTRODUCTION TO ACTIVITYSIM

ActivitySim is an open platform for activity-based modeling. It is a software effort that is supported both financially and managerially by a consortium of MPOs. The travel model that is currently implemented in the ActivitySim platform is based on a fully functional activity-based model that was originally designed for the San Francisco Metropolitan Transportation Commission (MTC) and Atlanta Regional Commission (ARC) starting around 2006. The model system was initially developed starting in the early to mid-2000s and has been in use by these agencies for practical transportation planning and policy analysis since approximately 2010.[24] The model currently implemented in the ActivitySim framework is a member of the Coordinated Travel - Regional Activity-based Modeling Platform family of models.[25] The system relies on logit choice models to represent travel decisions (how frequently to travel, where to travel to, by what mode, etc.) and was designed to achieve behavioral realism within a practical system of components. The existing model addresses many of the limitations noted above with respect to the MWCOG Ver. 2.3 trip-based model and other product requirements will be addressed via enhancements in Gen3 and Gen4 as noted below.

The original CT-RAMP model was developed jointly for both MTC and ARC, originally implemented in the Java programming language. In 2014, a consortium of Metropolitan Planning Organizations (MPOs) created the ActivitySim project to "create and maintain advanced, open-source, activity-based travel behavior modeling software based on best software development practices for distribution at no charge to the public".[26] The consortium decided to adopt the MTC Travel Model One activity-based model as the basis for the new software tool, and subsequently contracted for consultant services under the Association of MPOs (AMPO) to convert the model to Python, enhance, and maintain the software code. The new Python-based software is very flexible, configurable, and easy-to-use.

Current members of the consortium include ARC, MTC, San Diego Association of Governments (SANDAG), San Francisco County Transportation Authority (SFCTA), Puget Sound Regional Council (PSRC), Southeast Michigan Council of Governments (SEMCOG), Oregon Department of Transportation (ODOT), Metropolitan Council (Met Council), and AMPO. MTC Travel Model One has been fully implemented in ActivitySim. Model deployments are currently underway for ARC (there are minor differences between the MTC and ARC models that are being implemented) and SEMCOG. SANDAG has contracted for services to implement their Mexican Resident tour-based travel model in ActivitySim as an initial step. MTC is currently procuring consultant services to implement their second-iteration activity-based model (Travel Model Two)

---

[24] Metropolitan Transportation Commission. Plan/Bay Area: Technical Summary of Predicted Traveler Responses to First Round Scenarios Technical Report. March 22, 2011, available https://mtcdrive.app.box.com/s/3qj8egg1esg01ac68qtnlq8e0c4l4h6s.

[25] Davidson, Vovsha, Freedman, and Donnelly. CT-RAMP Family of Activity-Based Models. Australasian Transport Research Forum 2010 Proceedings. 29 September – 1 October 2010, Canberra, Australia. Publication website: http://www.patrec.org/atrf.aspx

[26] “ActivitySim: An Open Platform for Activity-Based Travel Modeling,” 2020. https://activitysim.GitHub.io/, accessed April 23, 2020.

in ActivitySim. As can be gleaned from the above information, ActivitySim has a robust and active user community.

The ActivitySim model framework has the following characteristics:

- Utilizes tours (sequences of trips beginning and ending at an anchor location such as home or work) as an organizing principle for the generation of travel and to ensure consistency across trips within a tour.
- Utilizes micro-simulation for modeling travel choices, in which a synthetic population is generated, and explicit mobility and travel choices are made for each decision-maker in the population according to contextual probability distributions.
- Addresses both household-level and person-level travel choices including limited intra-household interactions between household members.
- Schedules tours into time-windows to ensure there are no overlapping travel episodes.
- Reflects and responds to detailed demographic information including household structure, aging, changes in wealth, and other key attributes.

4.1 | HOW DOES ACTIVITYSIM MEET THE GEN3 PROJECT OBJECTIVES?

Objective 1: To ensure that the COG/TPB travel demand forecasting methods are either state of the practice or state of the art with respect to the modeling practices of peer MPOs.

According to a survey of 23 peer MPOs conducted in 2015,[27] 70% have or were developing a production-use activity-based travel demand model. Since the survey was conducted, one of the MPOs that did not have an AB model in development now has an ActivitySim model under development (SEMCOG). It is therefore the conclusion of the RSG team that the state of the practice for peer MPOs is an activity-based model, and state of the art of peer MPOs is an activity-based model that is “advanced” compared to peer MPO models in one or more ways - treatment of space, time, behavior, special travel markets, integration with dynamic traffic assignment, etc. We believe that ActivitySim represents state of the practice in terms of activity-based model form and function. We believe that an initial deployment of the current ActivitySim model (**Gen3, Phase I model development**) followed by selected enhancements to address policies and markets of specific interest to MWCOG (**Gen3, Phase II model development)** is the most prudent approach for Gen3 model development. Phase I would occur in the first half of the three-year contract and Phase II would occur in the second half of the contract. These two phases are further discussed later in this report. We suggest further development of state-of-the-art AB model features and functionality in Gen4.

---

[27] Cambridge Systematics, Inc., *Status of Activity-Based Models and Dynamic Traffic Assignment at Peer MPOs, Task Order 15.2, Report 2 of 3*, 10–11.

Of the peer MPOs in the survey, only two reported having a production use DTA model, with seven other models in development. We therefore consider AB/DTA model integration “state of the art” and recommend that it be considered for future model development (Gen4). We discuss this further below.

Objective 2: To address current shortcomings with the TPB’s adopted, production-use travel demand model (currently the Ver. 2.3 Model).

One of the key shortcomings of the Ver. 2.3 Model is the aggregation bias due to the trip-based model structure. ActivitySim is a disaggregate activity-based model. The structure of the model allows the use of any number of explanatory variables without affecting computational burden. Of course, the ability of MWCOG staff to forecast and maintain such variables is an important consideration, which is addressed further below. However, unlike aggregate models, an activity-based model structure does not preclude the incorporation of desired variables.

The ActivitySim model effectively responds to some of the key shortcomings noted above with respect to the Ver. 2.3 Model. For example, time-of-day choice is explicitly represented at the tour level. ABMs, such as ActivitySim, take into account accessibility and therefore respond to changes in congestion. As peak-period travel gets more congested relative to off-peak periods, the utility and probability of travel in peak periods decreases, all else being equal.

The following list is taken from the Ver. 2.3 Model key shortcomings summary above. We note how the Gen3 Model will or will not address each shortcoming.

Key features of importance to MWCOG (order reflects importance to Gen3 Model, not necessarily importance of a policy issue, since some important policy issues may be difficult to model):

1. Modeling non-motorized modes (walk and bike) through mode choice. Walk and bike modes are explicitly considered in ActivitySim mode choice models. We recommend that MWCOG consider a more refined spatial system (micro-zones) and development of an all-streets network for Gen4 model development. See data development, below.
2. Transportation Network Companies (TNCs) and other shared mobility modes. TNC modes have been added to ActivitySim and will be included in the Phase I models.
3. Teleworking. We propose to add a telework frequency model in Phase II model deployment. We note that RSG has estimated a telework frequency model for SANDAG and is developing a telework frequency model as part of the SEMCOG model development project so we will have several examples to start with.
4. Transit crowding. We plan to test link-level transit crowding in Cube PT for the current year using observed transit trip tables from on-board survey data and for future-year scenarios using the Phase I future-year model trip tables and networks. If the functionality results in reasonable results and the increase in runtime is acceptable to MWCOG staff, transit crowding will be implemented in the Phase II model deployment. See transit assignment, below.

5. Special travel markets. This section on special travel markets is expanded from the ones described above. Also see special market section below.
    a. External transit travel. External transit trip tables will be created from on-board survey data and assigned along with internal transit trips. A growth factoring methodology will be developed to grow transit trips into the future.
    b. Airport travel. Phase I models will use existing airport trip tables. We suggest replacing these trip tables with an explicit airport ground access simulation model in Phase II. The existing airport passenger survey is sufficient to develop such a model, but the development would be dependent on resource constraints.
    c. Overnight visitors. Gen3 Phase I models will use existing visitor trip tables. We suggest that MWCOG begin collecting visitor data, including total visitor counts from public sources, hotel/motel room inventory, and trip attraction data. We also suggest that MWCOG consider collecting either disaggregate and/or aggregate visitor travel data. An explicit overnight visitor travel model can be developed in either Gen3 Phase II or Gen4 depending on data availability and project resources.
    d. University students. University student travel is an explicit tour/trip purpose in ActivitySim. We suggest collecting and using university enrollment as a size term variable in school destination choice in Phase I models and building an explicit group quarters synthetic population that considers university student dwelling type. We suggest potential future enhancements in Gen3 Phase II or Gen4 models.
6. Employer-based transit subsidies. We propose to estimate and implement a transit pass ownership model in which partially and fully subsidized transit passes will be considered explicitly. We note that RSG is planning to develop a transit pass model as part of the SEMCOG model development project. See model enhancements, below.
7. Connected/autonomous vehicles (CAVs). We plan to enhance ActivitySim Phase II models to explicitly consider CAVs. RSG is developing CAV functionality in ActivitySim for SEMCOG, which can be transferred to MWCOG. Modeling the effect of travel time unreliability in travel behavior. We propose to address travel time unreliability in the Gen4 Model. See traffic assignment, below.
8. Model Sensitivity. We describe the temporal, spatial, and market segmentation of the ActivitySim model below.

Objective 3: To ensure that the new model has the capability to address the most pressing regional transportation planning issues in the Washington, D.C. region.

As noted above, the CT-RAMP model system upon which ActivitySim (MTC Travel Model One and ARC Activity-based model) has been used successfully for long-range regional transportation planning, transportation improvement program and air quality analysis for approximately 10 years. The model system has been used to analyze the impacts of land-use on transportation demand, the effects of highway capacity increases, planning for priced infrastructure including toll roads and managed lanes, demand for at-grade and grade-separated transit investments, and many other relevant projects and policies.

We provide the following references for further reading:

- Metropolitan Transportation Commission. *Plan/Bay Area: Technical Summary of Predicted Traveler Responses to First Round Scenarios Technical Report.* March 22, 2011
- Metropolitan Transportation Commission. *Plan/Bay Area: Technical Summary of Predicted Traveler Responses to Second Round Scenarios Technical Report.* January 5, 2012
- Metropolitan Transportation Commission with Parsons Brinckerhoff. *Plan/Bay Area: Project Performance Assessment Travel Modeling Methodological Approach Technical Paper.* February 2, 2012.
- Metropolitan Transportation Commission. *Plan Bay Area 2040: Technical Summary of Predicted Traveler Responses to Planning Scenarios Technical Paper.* May 2016.
- Metropolitan Transportation Commission and Association of Bay Area Governments. *Travel Modeling Report Plan Bay Area 2040 Final Supplemental Report.* July 2017.
- Metropolitan Transportation Commission and Association of Bay Area Governments. *Futures Final Report: Resilient And Equitable Strategies For The Bay Area's Future.* January 2020
- Atlanta Regional Commission. *Conformity Determination Report: Atlanta Non-attainment and Maintenance Areas in Support of the Atlanta Regional Plan (2020), Gainesville-Hall Regional Transportation Plan (2020), and Bartow On the Move (2020).* February 2020.

4.2 | OVERVIEW OF THE ACTIVITYSIM MODEL SYSTEM

Figure 3 shows the design of the MTC Travel Model One CT-RAMP model implemented in ActivitySim.. In order to understand the flow chart, some definitions are required. These are described in more detail below and in the appendix.

- Tour: A sequence of trips that start and end at an anchor location. In ActivitySim, anchors are home or work.
- Primary destination: The “main” activity of the tour; this activity determines the tour purpose. It also divides the tour into two "legs"; the sequence of trips from the anchor location to the primary destination is the outbound leg, and the sequence of trips from the primary destination back to the anchor location is the inbound or return leg.
- Mandatory activity: Work or school

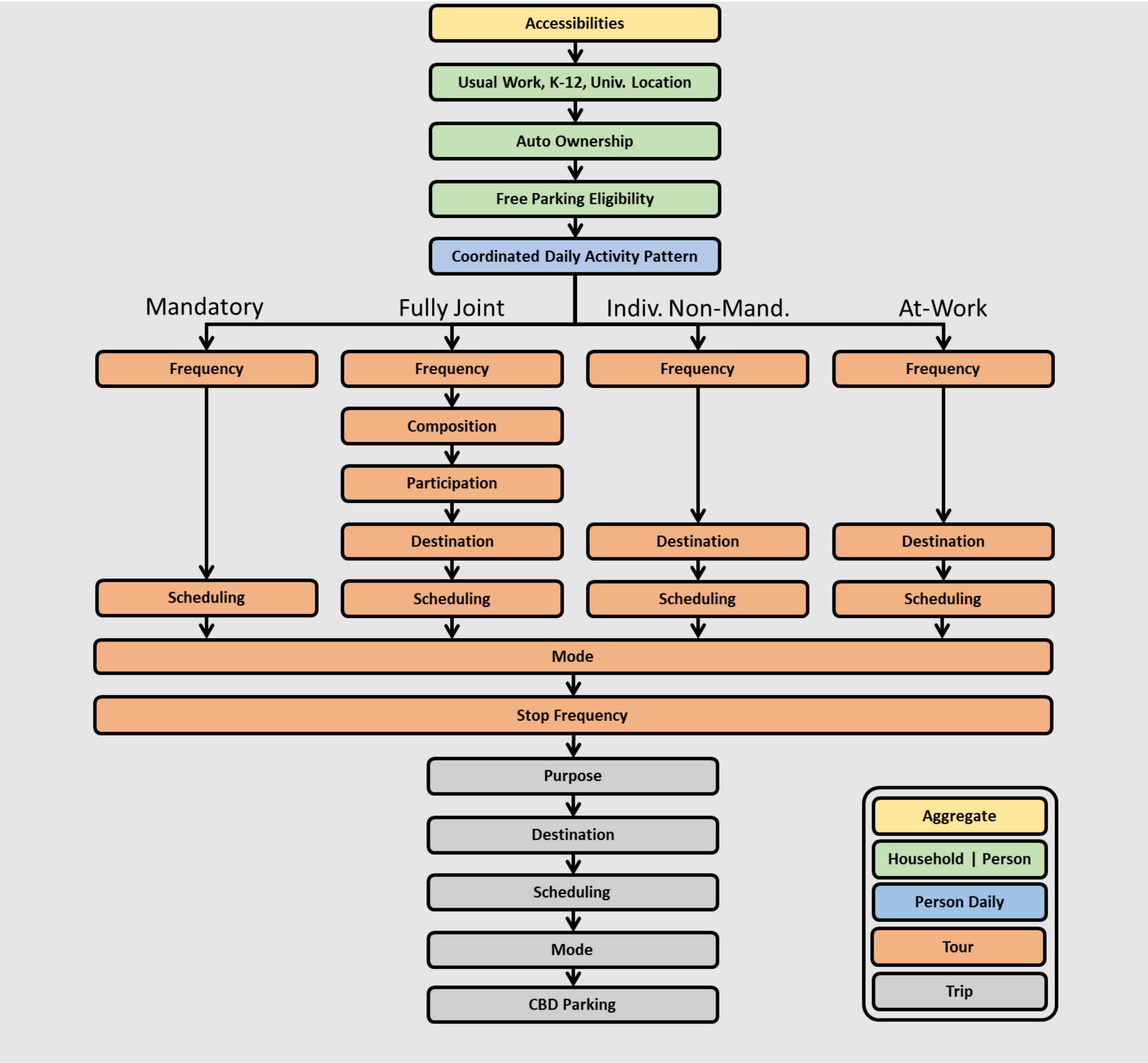
- Non-mandatory activity: Any out of home activity that is not work or school, including maintenance activities such as shopping as well as discretionary activities such as out-of-home recreation and eating out.
- Fully joint tour: A tour in which two or more household members travel together to all out-of-home activity locations and return home together. In other words, no household member is picked-up or dropped-off en route.
- Intermediate stop: An out-of-home activity location on the tour other than the anchor location or the primary destination. Intermediate stops are made on the way from the anchor location to the primary destination (outbound) or on the way from the primary destination back to the anchor location (inbound).
- Tour mode: The “main mode” or “preferred mode” of the tour. This is an abstract concept used categorize the tour with respect to accessibility and constrain the availability of modes for trips on the tour to ensure some consistency of modes used for each trip.

The first model in the sequence is mandatory location choice; this model is run for all workers and students regardless of whether they attend work or school on the simulated day. Next, one or more mobility models are run. In the Phase I models this includes household auto ownership and worker free parking eligibility. Next, the daily activity pattern model is run, which predicts the general activity pattern type for every household member. Then Mandatory tours are generated for workers and students, the tours are scheduled (their location is already predicted by the work/school location choice model), and the tour mode is chosen. Fully joint tours are generated at a household level, their composition is predicted (adults, children or both), the participants are determined, and a tour mode is chosen. The primary destination of fully joint tours is predicted, the tours are scheduled, and a tour mode is chosen. Next, non-mandatory tours are generated, their primary destination is chosen, they are scheduled, and a tour mode is chosen for each. At-work subtours are tours that start and end at the workplace. These are generated, scheduled (with constraints that the start and end times must nest within the start and end time of the parent work tour), a primary destination is selected, and a tour mode is chosen.

At this point, all tours are generated, scheduled, have a primary destination, and a selected tour mode. The next set of models fills in details about the tours - number of intermediate stops, location of each stop, the departure time of each stop, and the mode of each trip on the tour. Finally, the parking location of each auto trip to the central business district (CBD) is determined.

After the model is run, the output files listed above are created. The trip lists are then summarized into origin-destination matrices by time period and vehicle class or transit mode and assigned to the transport network. Skims are created based on congested times, and the model system is iterated multiple times until either some convergence threshold is attained, or a predetermined number of iterations is reached.

FIGURE 3: EXISTING ACTIVITYSIM MODEL DESIGN



ActivitySim would be used to represent all internal travel made by residents of the MWCOG region (modeled area). The decision-makers in the model system include both persons and households. These decision-makers are created (synthesized) for each simulation year and land-use scenario, based on Census data and forecasted distributions of households and persons by key socio-economic categories. The decision-makers are used in the subsequent discrete-choice models in a microsimulation framework where a single alternative is selected from a list of available alternatives according to a probability distribution. The probability distribution is generated from a logit model which considers the attributes of the decision-maker and the attributes of the various alternatives. The application paradigm is referred to as Monte

Carlo simulation, since a random number draw is used to select an alternative from the probability distribution. The decision-making unit is an important element of model estimation and implementation and is explicitly identified for each model specified in the following sections.

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

Person Types

A total of eight segments of person-types, shown in Table 2, are used for the model system. The person-types are mutually exclusive and collectively exhaustive with respect to age, work status, and school status. These person types are coded in the synthetic population according to person-level attributes. The same methodology is used to code person type in travel behavior survey data, so that apples-to-apples comparisons can be made between model results and observed data. Note that the 2017-18 MWCOG Regional Travel Survey[28] does not specifically ask number of hours worked per week. However, the survey does ask number of jobs, number of trips to work per week, and usual work start and end time, so it should be possible to infer full or part-time status from response to those and other questions. All other person-type categories can be calculated from response to age, work, and school questions in the household survey, and all segments can be identified in the American Community Survey (ACS) Public Use Microdata Sample (PUMS) person file.

---

[28] RSG. *Metropolitan Washington Council of Governments Regional Travel Survey Final Report*. July 31, 2019.

TABLE 2: PERSON TYPES

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

Activity Types

The 2017-18 MWCOG Regional Travel Survey included 23 different activity codes plus a write-in option for “other.” Modeling all activity types would add significant complexity to estimating and implementing the model system, so these detailed activity types are grouped into more aggregate activity types, based on the similarity of the activities. The activity types are used in most model system components, from developing daily activity patterns and to predicting tour and trip destinations and modes by purpose.

The activity types explicitly modeled in ActivitySim are shown in Table 3. The activity types are also grouped according to whether the activity is mandatory, maintenance, or discretionary, and eligibility requirements are assigned determining which person-types can be used for generating each activity type. The classification scheme of each activity type reflects the relative importance or natural hierarchy of the activity, where work and school activities are typically the most inflexible in terms of generation, scheduling and location, whereas discretionary activities are typically the most flexible on each of these dimensions. Each out-of-home location that a person travels to in the simulation is assigned one of these activity types, as shown in the table. Any activity with missing purpose is recoded as discretionary activity. Note that “change mode” trips are linked according to a set of trip-linking rules, and in-home activities 1 (At home activity) and 2 (Work at home or telework) are not differentiated. This is one of the characteristics of a “simpler” activity-based model; in-home activities are not modeled explicitly in ActivitySim.

In order to avoid double-counting commercial vehicle trips, the MWCOG ActivitySim model will exclude travel made by delivery drivers for work-related purposes. The specific variables that can be used to identify the work-related travel by those who typically drive for work are workplace equal to 4 (Drives for a living) and activity purpose equal to 3 or 4 (work at regular workplace, other location or work-related activity).

Tours made by delivery drivers with stops for these purposes will be dropped from model estimation and calibration summaries to avoid double-counting. Trips made for the purposes of delivering persons (taxi, school bus trips) will be retained in the data but these are expected to be a very small overall share of travel.

Also note that the current ActivitySim implementation does not have a separate purpose for work-related activities, such as meetings and service calls. Work activities are specifically work at fixed locations. Work at non-fixed locations and other work-related activities will be grouped with “Other Maintenance.” A more explicit treatment for these activities, such as adding a work-related purpose and extending mandatory tour frequency models, is not currently in scope.

TABLE 3: OUT-OF-HOME ACTIVITY TYPES

TYPE	PURPOSE	DESCRIPTION	CLASSIFICATION	RELEVANT PURPOSES FROM SURVEY
1	Work	Working at regular workplace.	Mandatory	3 - Work at regular workplace or other work location
2	University	College +	Mandatory	7 - Attend school/class as a student 8 - Attend other school-related activity (if student and at regular school location)
3	Pre-School/Grade School/High School	Grades K-12	Mandatory	7 - Attend school/class as a student 9 - Receive childcare or preschool services (if not home)
4	Escorting	Pick-up/drop-off passengers (auto trips only).	Maintenance	6 - Drop off/pick up someone

5	Shopping	Shopping away from home.	Maintenance	11 - Shop in store 13 - Quick stop to pick up food or coffee
6	Other Maintenance	Personal business/services, and medical appointments.	Maintenance	4 - Work-related activity 10 - Receive adult care services (if not home) 14 - Fuel vehicle/get gas 15 - Receive healthcare services 16 - Receive personal services 22 - Mail package/letter or other postal activity
7	Social/Recreational	Recreation, visiting friends/family.	Discretionary	8 - Attend other school-related activity (if not student or not at regular school location) 17 - Entertainment 18 - Socialize 19 - Recreation 20 - Exercise
8	Eat Out	Eating outside of home.	Discretionary	12 - Eat a meal/have coffee or drink
9	Other Discretionary	Volunteer work, religious activities.	Discretionary	5 - Volunteer activity 21 - Governmental, civic, or religious activity

Treatment of time

The MTC Travel Model One version of ActivitySim utilizes a one-hour temporal resolution, with aggregation for the first and last periods of the day. The ARC and SEMCOG Phase II versions of ActivitySim use a half-hourly time period resolution. We propose to implement MWCOG phase I model system with an hourly temporal resolution, beginning with 3 A.M. and ending with 3 A.M. the next day. We propose to switch to a half-hourly temporal resolution in MWCOG's Gen3, Phase II model, to provide greater consistency with potential DTA integration in Gen4.

Temporal integrity is ensured so that no activities are scheduled with conflicting time windows, except for short activities/tours that are completed within an hour increment. For example, a person may have a short tour that begins and ends within the 8am-9am period, as well as a second longer tour that begins within this time period but ends later in the day.

A critical aspect of the model system is the relationship between the temporal resolution used for scheduling activities, and the temporal resolution of the network simulation (skimming and assignment) periods. Although each activity generated by the model system is identified with a start time and end time in one-hour increments, level-of-service matrices will be created for four aggregate time periods consistent with the current Ver. 2.3 trip-based model: A.M. Peak, Midday, P.M. Peak, and Evening/Early A.M. The trips occurring in each time period reference the appropriate transport network depending on their trip mode and the mid-point trip time. The definition of time periods for level-of-service matrices is given in Table 4, below. Note that in Phase II, the model system temporal resolution will be at 30 minutes, so half-hour breakpoints can be used for mapping model time periods to skims and assignment if desired.

TABLE 4: TIME PERIODS FOR LEVEL-OF-SERVICE SKIMS AND ASSIGNMENT

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

Treatment of Space

Currently ActivitySim uses Transportation Analysis Zones (TAZs) as the unit of space in the model. We recommend utilizing the current Ver. 2.3 Model system TAZs for the Gen3 Model. We suggest that MWCOG consider development of a Micro-Analysis Zone (MAZ) system for the Gen4 Model, since a key desired enhancement of the MWCOG travel demand model is enhanced representation of non-motorized accessibility.

The construction of MAZs can be done in several ways. One approach is to use Census blocks as the basic geographic unit. Another approach is to use an all-streets network (excluding freeways and ramps) to generate an MAZ layer. This approach has the advantage that MAZs will be more consistent with the underlying street network. In either case, one must intersect the resulting coverage with TAZs to ensure that MAZs nest within TAZs and that every internal TAZ has at least one MAZ within it. This may result in "slivers" that must be cleaned up by further processing the MAZ layer. For a region the size of MWCOG, we expect the MAZ layer to have approximately 50-60k polygons. This may require collapsing some MAZs; this can be performed semi-manually by searching for MAZs smaller than a certain size and dissolving their boundaries until the desired number of MAZs is reached. We also recommend that COG add MAZ detail to “greenfield” areas that are likely to have significant future land development in the future. These MAZs are likely to be very large using either method described above due to lack of local streets. However, it is likely that local streets would be built as the land develops. Leaving MAZs large may bias future year model results against non-motorized travel as greenfields fill in with future development.

Once the MAZ layer is developed, it must be populated with land-use data described in section 7.1 below. There are multiple ways to do this. Typically census data at the block level is aggregated by MAZ for base-year estimates of households, and disaggregate employment data is geocoded to the MAZ layer for base-year employment estimates. **Future-year data can be more challenging to forecast at the MAZ level, particularly if relying on local jurisdictions to provide such forecasts**. For agencies that struggle with forecasting land-use at the MAZ level, we recommend continuing to request allocations at the TAZ level, and perform automated allocation procedures to MAZs. Ideally such procedures would take into account buildable land in each MAZ.

ActivitySim is currently being enhanced to consider MAZs, but the runtime implications of implementing this functionality are unknown. Note that ActivitySim does not currently use walk market segmentation for transit walk access. **In order to enable this functionality, the MAZ functionality must be implemented**. If available for Phase II deployment, we can enable walk market segmentation similar to what is currently being used in the Ver. 2.3 Model by breaking each TAZ into short walk, long walk, and no walk segments using the same walk buffering process currently used by the Ver. 2.3 Model. This would essentially replicate the walk market segmentation in a simulation context. The runtime implications of doing so will be monitored to ensure that the model runtime is acceptable to MWCOG.

Trip Modes

The current MWCOG model structure differentiates three auto modes, four transit technology options (all bus, all Metrorail, bus and Metrorail, and commuter rail), and three access modes (walk, park-and-ride (PNR) and kiss-and-ride (KNR)). We propose to utilize the same auto and transit alternatives in ActivitySim. We believe that this structure adequately represents the competition between key transit options in the region and the unique unobserved characteristics of each option. We will continue to build transit skims using the same mode groupings as are currently defined in the Ver. 2.3 Model, unless assignment of transit on-board survey trip tables indicates that the structure should be revised in some way.

We propose to extend the structure in three ways in Phase I models. First, we will add a non-motorized nest with walk and bike alternatives. The travel time for walking and biking will be based on a "non-motorized" shortest path skim that minimizes distance, excluding freeways and ramps. Second, we will add a "ride hail" nest, with two TNC modes (single-payer and pool TNC) as well as traditional taxi. This enhancement has recently been implemented in the ActivitySim framework. The 2017-18 Regional Travel Survey identifies TNC trips as a separate mode; however, the survey does not differentiate between single-payer and pooled trips. We will have to assert a reasonable split between these modes from other data sources such as the recent California TNC survey. Finally, we will add a school bus mode for K-12 school trips.

Table 5 lists the trip modes that will be defined in MWCOG ActivitySim model. We do not anticipate there being any difference in mode definition between Phase I and Phase II models. The table also describes the mapping between these modes and the 2017-18 Regional Travel Survey, and the relevant network skims that will be utilized for each mode. Note that the actual mode coding rules will be more complicated than what is identified in this table, since trip linking logic (where change mode activities are linked out of trips) will be employed. It is possible that COG survey staff have already implemented these rules. We will document these procedures after we have processed the survey data. For Phase II models, we will test value-of-time segmentation, and if MWCOG staff determine that model sensitivities outweigh runtime, we will skim auto modes by value-of-time bin.

Current skimming procedures do not build PNR or KNR skims with auto as an egress mode; this is typical in trip-based models in which mode choice is estimated in a production-attraction (PA) format. Since PA format is used in trip-based mode choice models, drive is only represented as an access mode as the vast majority of drive-transit trips use auto at the production, or home, end of the trip. In activity-based models, origin-destination (OD) format is used for mode choice. Therefore, it is necessary to represent auto as both an access option (i.e. for the outbound leg of the tour) and an egress option (for the inbound leg of the tour). We will modify skimming procedures accordingly.

The current ActivitySim mode choice model does not have a “bike” access or egress to transit option. We suggest carefully analyzing on-board transit survey data to better understand the size and characteristics of this market. In many models without an explicit bike-transit mode, these trips are often combined with kiss-and-ride, since the length of their bike leg is typically consistent with the average length of the drive leg of kiss-and-ride trips. If the market is large enough to warrant doing so (in other words, the same size or larger than one of the other access markets) or if there are special policy needs (for example, the need to size bike carrying capacities of buses), some model systems break out bike-transit as a separate mode. This provides the advantage of representing bike mode as both access and egress; according to analysis of data from Valley Metro's (Phoenix Arizona) on-board survey, 90% of bike-transit trips use the bike on both ends of the transit trip.[29]

---

[29] 2014-2015 Valley Metro Onboard Survey Final Report, Prepared for Maricopa Association of Governments (MAG) and Valley Metro Transit System by ETC Institute (no date)

TABLE 5: TRIP MODES FOR ASSIGNMENT

NUMBER	MODE	RELEVANT MODES FROM HOUSEHOLD SURVEY	RELEVANT SKIMS FROM MWCOG MODEL
1	Auto SOV	3 Household vehicle 4 Other car, truck, or van (e.g., someone else's car, rental vehicle, etc.) Where occupancy = 1	SOV skims where HOV2 and HOV3+ lanes are unavailable
2	Auto 2 Person	3 Household vehicle 4 Other car, truck, or van (e.g., someone else's car, rental vehicle, etc.) Where occupancy = 2	HOV2 skims where HOV3+ lanes are unavailable
3	Auto 3+ Person	3 Household vehicle 4 Other car, truck, or van (e.g., someone else's car, rental vehicle, etc.) Where occupancy = 3+	HOV3+ skims, all facilities are available.
4	Walk	1 Walk (or jog/wheelchair)	Non-motorized distance
5	Bike	2 Bicycle	Non-motorized distance
6	Walk to All Bus	Linked trip with 1 Walk and 23 Local Bus and/or 55 Commuter or express bus	Walk-bus-walk skim
7	Walk to All Metrorail	Linked trip with 1 Walk and 42 Light rail or streetcar and/or 14 Subway (e.g., Metrorail, Baltimore Metro)	Walk-Metrorail-walk skim
8	Walk to Bus-Metrorail	Linked trip with 1 Walk and 23 Local Bus and/or 55 Commuter or express bus	Walk-Bus/Metrorail-walk skim

		and 42 Light rail or streetcar and/or 14 Subway (e.g., Metrorail, Baltimore Metro)	
9	Walk to Commuter Rail	Linked trip with 1 Walk and 13 Commuter rail (e.g., MARC, VRE), plus any other transit mode	Walk-Commuter Rail-walk skim
10	PNR to All Bus	Linked trip with 3 Household vehicle 23 Local Bus and/or 55 Commuter or express bus	PNR-bus-walk skim (outbound) walk-bus-PNR skim (inbound)
11	PNR to All Metrorail	Linked trip with 1 Walk 42 Light rail or streetcar and/or 14 Subway (e.g., Metrorail, Baltimore Metro)	PNR-Metrorail-walk skim (outbound) walk-Metrorail-PNR skim (inbound)
12	PNR to Bus-Metrorail	Linked trip with 3 Household vehicle 23 Local Bus and/or 55 Commuter or express bus and 42 Light rail or streetcar and/or 14 Subway (e.g., Metrorail, Baltimore Metro)	PNR-Bus/Metrorail-walk skim (outbound) walk-Bus/Metrorail-PNR skim (inbound)
13	PNR to Commuter Rail	Linked trip with 3 Household vehicle and 13 Commuter rail (e.g., MARC, VRE), plus any other transit mode	PNR-Commuter Rail-walk skim (outbound) walk-Commuter Rail-PNR skim (inbound)

14	KNR to All Bus	Linked trip with 3 Household vehicle 23 Local Bus and/or 55 Commuter or express bus	KNR-bus-walk skim (outbound) walk-bus-KNR skim (inbound)
15	KNR to All Metrorail	Linked trip with 1 Walk 42 Light rail or streetcar and/or 14 Subway (e.g., Metrorail, Baltimore Metro)	KNR-Metrorail-walk skim (outbound) walk-Metrorail-KNR skim (inbound)
16	KNR to Bus-Metrorail	Linked trip with 3 Household vehicle 23 Local Bus and/or 55 Commuter or express bus and 42 Light rail or streetcar and/or 14 Subway (e.g., Metrorail, Baltimore Metro)	KNR-Bus/Metrorail-walk skim (outbound) walk-Bus/Metrorail-KNR skim (inbound)
17	KNR to Commuter Rail	Linked trip with 3 Household vehicle and 13 Commuter rail (e.g., MARC, VRE), plus any other transit mode	KNR-Commuter Rail-walk skim (outbound) walk-Commuter Rail-KNR skim (inbound)
18	Taxi	5 Taxi or private limo service	HOV 2 skim
19	TNC-single	6 Ride-hailing service (e.g., Uber, Lyft)	HOV 2 skim
20	TNC-pool	6 Ride-hailing service (e.g., Uber, Lyft)	HOV 3+ skim
21	School Bus	7 School Bus	HOV3+ skims

4.3 | DESCRIPTION OF STANDARD ACTIVITYSIM MODEL

Accessibilities

The accessibilities model computes aggregate (zonal) measures of accessibility used by the downstream models. The accessibility measure is the equivalent of a destination choice logsum where the level-of-service variable is restricted to a certain mode (or set of modes) and the size term is restricted to a certain employment variable. These origin-based logsums are used as explanatory variables in models where the tour destination is unknown, to reflect the accessibility of the household to non-mandatory activities. For example, they are used as explanatory variables in tour frequency; as accessibility to retail employment increases, one would expect the frequency of shopping tours to increase. The equation is shown below.

$$A_i = \ln\left[\sum_{j=1}^{I} S_j \times \exp\left(-\gamma c_{ij}\right)\right]$$

where:

A_i is the accessibility for origin TAZ i

S_j is the size term for destination TAZ j

C_{ij} is the cost of travel from origin TAZ i to destination TAZ j

γ is a parameter indicating the sensitivity to the cost of travel

Accessibilities are used in daily activity pattern models and tour frequency models. The output accessibility measures by zone are:

- autoPeakRetail – the accessibility by auto during peak conditions to retail employment
- autoPeakTotal – the accessibility by auto during peak conditions to all employment
- autoOffPeakRetail – the accessibility by auto during off-peak conditions to retail employment
- autoOffPeakTotal – the accessibility by auto during off-peak conditions to all employment
- transitPeakRetail – the accessibility by transit during peak conditions to retail employment
- transitPeakTotal – the accessibility by transit during peak conditions to all employment
- transitOffPeakRetail – the accessibility by transit during off-peak conditions to retail employment
- transitOffPeakTotal – the accessibility by transit during off-peak conditions to all employment
- nonMotorizedRetail – the accessibility by walking during all time periods to retail employment
- nonMotorizedTotal – the accessibility by walking during all time periods to all employment

Mandatory (workplace/university/school) Activity Location Choice

Number of Models: 4 (Work, K-8, 9-12, University)

Decision-Making Unit: Workers for Work Location Choice; Persons attending K-8 for K-8 students model, Persons attending High School for High School model; University Students for University Model

Model Form: Multinomial Logit

Alternatives: Zones

A workplace location choice model assigns a workplace TAZ for every employed person in the synthetic population. Every worker is assigned a regular work location zone (TAZ) according to a multinomial logit destination choice model. Size terms vary according to worker household income, to reflect the different types of jobs that are likely to attract different (white collar versus blue-collar) workers. Accessibility is measured by a “representative” mode choice logsum based on peak period travel (8-9 A.M. departure and 5-6 P.M. return), as well as distance to the workplace. The mode choice logsum represents the total ease of travel between two zones across all available modes.

Since mode choice logsums are required for each destination, a two-stage procedure is used for all destination choice models in ActivitySim in order to reduce computational time (it would be computationally prohibitive to compute a mode choice logsum for each of approximately 3,000 zones and every worker in the synthetic population). In the first stage, a simplified destination choice model is applied in which all zones are alternatives. The only variables in this model are the size term and distance. This model creates a probability distribution for all possible alternatives (zones with no employment are not sampled). A set of alternatives are sampled from the probability distribution and these alternatives constitute the choice set in the full destination choice model. Mode choice logsums are computed for these alternatives and the destination choice model is applied. A discrete choice of TAZs is made for each worker from this more limited set of alternatives. In the case of the work location choice model, a set of 40 alternatives is sampled.

The application procedure utilizes an iterative shadow pricing mechanism in order to match workers to input employment totals. Shadow prices are alternative-specific constants that are calculated by the ratio of total employment to total workers in each destination zone (ln(input employment/estimated workers)). The use of shadow prices in destination choice is mathematically equivalent to a doubly constrained gravity model in which the resulting trip table matches both input productions and input attractions. The shadow prices are written to a file and can be used in subsequent model runs to cut down computational time.

A grade school location choice model assigns a school location for every person enrolled in grade school in the synthetic population. The size term in this model will be K-12 enrollment, unless K-8 grade school enrollment can be broken out separately, in which case separate models will be applied for grade school students versus high school students. Note that enrollment should include both public and private grade schools. Grade school parameters include person/household characteristics, representative school mode choice logsums (outbound 8-9 A.M. and return 2-3 P.M.), distance, and size terms. A high school location choice model assigns a school location for every person enrolled in high school, like the grade school location choice model.

A university location choice model assigns a university location for every university student in the synthetic population. The size term in this model is university enrollment. The University location choice model parameters include person/household characteristics, representative university mode choice logsums (same periods as school logsum), distance, and size terms.

Car Ownership Model

Number of Models: 1

Decision-Making Unit: Households

Model Form: Multinomial Logit

Alternatives: Four (0, 1, 2, 3+ autos)

The car ownership models predict the number of vehicles owned by each household. It is formulated as a choice model with four alternatives, including “no cars”, “one car”, “two cars”, and “three or more cars”. The model includes the following explanatory variables:

- Household size and composition
- Income
- Accessibilities and urban form

Free Parking Eligibility Model

Number of Models:	1
Decision-Making Unit:	Person
Model Form:	Multinomial Logit
Alternatives:	2 (has free parking or not)

The free parking eligibility model determines whether a worker has free parking at their workplace. The explanatory variables include workplace location (currently based on county), household size, household income, and auto ownership.

Coordinated Daily Activity Pattern (DAP) Model

Number of Models:	1
Decision-Making Unit:	Households
Model Form:	Multinomial Logit

Alternatives:	3 alternatives for one person households
	9 alternatives for two person households
	27 alternatives for three person households
	81 alternatives for four person households
	243 alternatives for five-or-more person households
	363 total alternatives across all household sizes

The DAP is classified by three main pattern types:

- Mandatory pattern (M) that includes at least one of the three mandatory activities – work, university or school. This constitutes either a workday or a university/school day, and may include additional non-mandatory activities such as separate home-based tours or intermediate stops on the mandatory tours.
- Non-mandatory pattern (N) that includes only individual and/or joint maintenance and discretionary tours. By virtue of the tour primary purpose definition, maintenance and discretionary tours cannot include travel for mandatory activities.
- At-home pattern (H) that includes only in-home activities. At-home patterns are not distinguished by any specific activity (e.g., work at home, take care of child, being sick, etc.). Cases with complete absence from town (e.g., business travel) are also combined with this category.

Statistical analysis implemented with the Columbus, Atlanta, and San Francisco Bay Area data has shown that there is an extremely strong correlation between DAP types of different household members, especially for joint N and H types. For this reason, the DAP for different household members are not modeled independently but rather simultaneously across multiple household members. The total number of possible DAP type combinations is significant for large households. However, there are several important considerations that significantly reduce the dimensionality of the simultaneous model. First, mandatory DAP types are only available for appropriate person types (workers and students). Intra-household coordination of DAP types is relevant only for the N- and H-type patterns. Thus, simultaneous modeling of DAP types for all household members is essential only for the trinary choice (M, N, H), while the sub-choice of the mandatory pattern can be modeled for each person separately.

The Coordinated DAP model features simultaneous modeling of trinary pattern alternatives for all household members with the subsequent modeling of individual alternatives, as shown in Figure 4 **Error! Reference source not found.Error! Reference source not found.Error! Reference source not found.Error! Reference source not found.**for a 3-person household (resulting in 27 alternatives). Tour frequency choice is a separate choice model, conditional upon the choice of alternatives in the trinary choice.

Simultaneous modeling of potentially joint alternatives for all household members assumes that for each person only a trinary choice (M, N, H) is considered. For a household of five persons the simultaneous combination of trinary models results in a total of 243 alternatives. For households with size greater than five, the model is applied for the first five household members by priority while the rest of the household members are processed sequentially, conditional upon the choices made by the first five members.

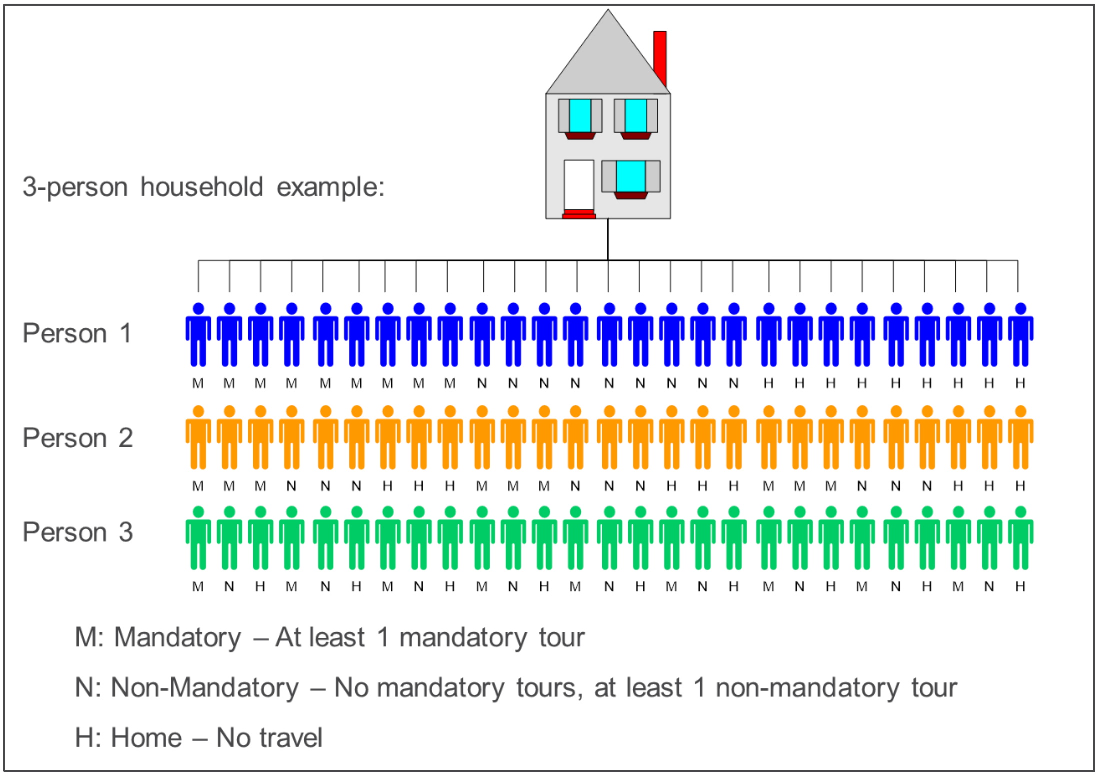


FIGURE 4: DAILY ACTIVITY PATTERN CHOICES

The Coordinated DAP model contains explanatory variables that include person and household attributes, accessibility measures, and density/urban form variables. Since the model features intra-household interactions, a subset of the parameters in the model are specified as interaction terms. These terms are based on the contribution to the total utility of an alternative from either a two-person interaction, a three-person interaction, or an entire-household interaction. For example, the contribution of a two-worker interaction to the utility for each worker to stay home on the simulation day is positive, indicating that it is more likely that both workers will attempt to coordinate their days off to engage in recreational opportunities together. Similarly, the contribution of a pre-school child to a worker’s mandatory pattern is negative, indicating the likelihood that if a pre-school child stays at home, a worker also is more likely to stay at home with the child. .

Individual Mandatory Tour Frequency

Number of Models: 1

Decision-Making Unit:	Persons
Model Form:	Multinomial Logit
Alternatives:	5 (1 Work Tour, 2 Work Tours, 1 School Tour, 2 School Tours, 1 Work/1 School Tour)

Based on the DAP chosen for each person, individual mandatory tours, such as work, school and university tours are generated at person level. The model predicts the exact number and purpose of mandatory tours (e.g., work and school/ university) for each person who chose the mandatory DAP type at the previous decision-making stage. Since the DAP type model at the household level determines which household members engage in mandatory tours, all persons subjected to the individual mandatory tour model implement at least one mandatory tour. The model has the following five alternatives:

- One work tour,
- One school tour,
- Two or more work tours,
- Two or more school tours,
- One work tour plus one school tour.

DAPs and subsequent behavioral models of travel generation include various explanatory variables that relate to household composition, income, car ownership, location of work and school activities, land-use development, residential and employment density, and accessibility factors.

Individual Mandatory Tour Time of Day Choice

Number of Models:	1
Decision-Making Unit:	Persons
Model Form:	Multinomial Logit
Alternatives:	190

After individual mandatory tours have been generated, the tour departure time from home and arrival time back at home is chosen simultaneously. Note that it is not necessary to select the destination of the tour, as this has already been determined. The model is a discrete-choice construct that operates with tour departure-from-home and arrival-back-home time combinations as alternatives. The proposed utility structure is based on "continuous shift" variables, and represents an analytical hybrid that combines the advantages of a discrete-choice structure (flexible in specification and easy to estimate and apply) with the advantages of a duration model (a simple structure with few parameters, and which supports continuous time). The model has a temporal resolution of an hour, with the period between 3 AM and 5 AM collapsed and the period between 10 PM and 3 AM the next day collapsed. The model utilizes direct availability rules for each subsequently scheduled tour, to be placed in the residual time window left after scheduling tours of higher priority. This conditionality ensures a full consistency for the individual's entire-day activity and travel schedule as an outcome of the model.

The model utilizes household, person, and zonal characteristics, most of which are generic across time alternatives. However, network level-of-service (LOS) variables vary by time of day and are specified as being alternative-specific based on each alternative's departure and arrival time. By using generic coefficients and variables associated with the departure period, arrival period, or duration, a compact structure of the choice model is created, where the number of alternatives can be arbitrarily large depending on the chosen time unit scale, but the number of coefficients to be estimated is limited to a reasonable number. Duration variables can be interpreted as "continuous shift" factors that parameterize the duration in such a way that if the coefficient multiplied by the variable is positive, this means the whole distribution is shifted to the longer durations. Negative values work in the opposite direction, collapsing the distribution toward shorter durations.

In the CT-RAMP model structure, the tour-scheduling model is placed after destination choice and before mode choice. Thus, the destination of the tour and all related destination and origin-destination attributes are known and can be used as variables in the model estimation.

The choice alternatives are formulated as tour departure from home/arrival at home in hour combinations (g,h), and the mode choice logsums and bias constants are related to departure/arrival periods (s,t). Tour duration is calculated as the difference between the arrival and departure hours ($h-g$) and incorporates both the activity duration and travel time to and from the main tour activity, including intermediate stops.

The tour time-of-day (TOD) choice utility has the following general form:

$$V_{gh} = V_g + V_h + D_{h-g} + \mu \ln\left(\sum_m V_{stm}\right)$$ ***EQUATION 1***

where:

V_g, V_h = departure and arrival time-specific components

D_{h-g} = duration-specific components

m = entire-tour modes (SOV, HOV, walk to transit, drive to transit, non-motorized)

V_{stm} = mode utility for the tour by mode *m*, leaving home in period *s* (containing hour *h*) and returning home in period *t* (containing *g*)

μ = mode choice logsum coefficient

The network simulations to obtain travel time and cost skims are implemented for four broad periods:

- A.M. Peak

- Midday
- P.M. Peak
- Evening/ Early AM

Mode-choice logsums are used for all relevant combinations of the four time periods above.

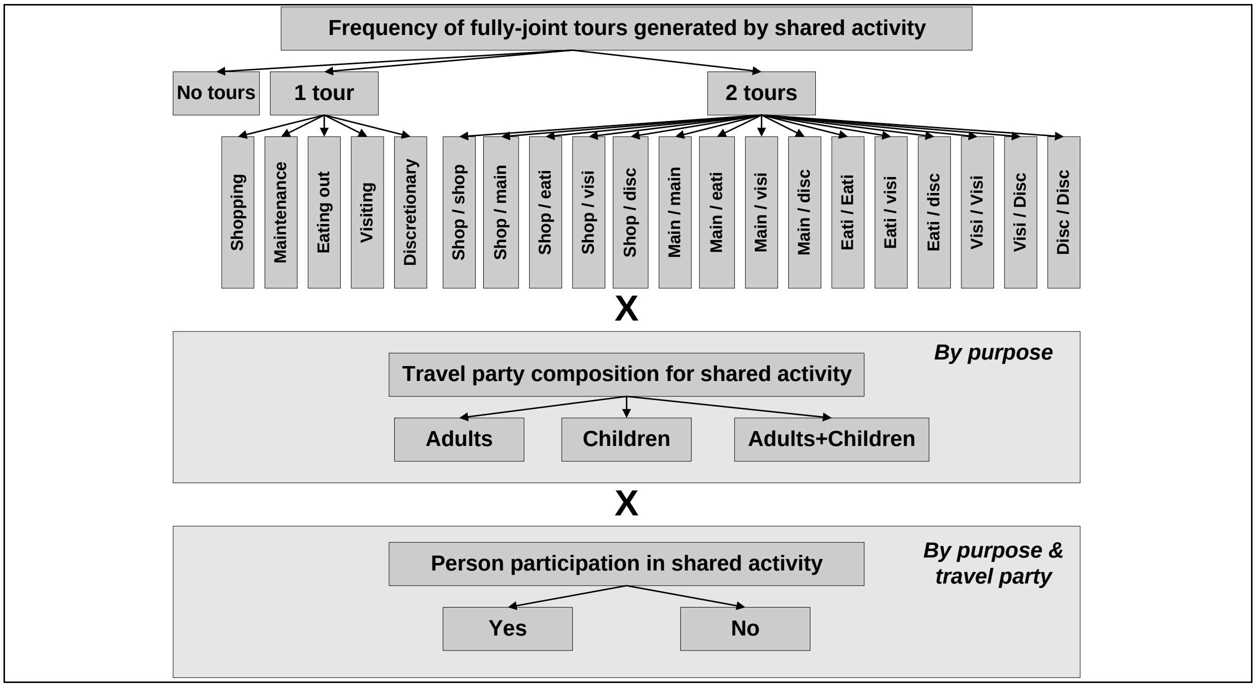
Generation of Joint Household Tours

In the current model structure, joint travel for non-mandatory activities is modeled explicitly in the form of fully joint tours, where all members of the travel party travel together from the very beginning to the end and participate in the same activities along the way. Other types of joint travel like carpooling of workers and escorting children are not explicitly considered currently, though they are handled implicitly through shared-ride alternatives in tour and trip mode choice.

Each fully joint tour is considered a unit of modeling with a group-wise decision-making for the primary destination, mode, frequency and location of stops, etc. Formally, modeling joint activities involves three linked stages (Figure 5):

- A joint tour frequency model that generates the number of joint tours by purpose/activity type made by the entire household.
- A travel party composition model that predicts whether the tour is composed of adults-only, children-only, or adults and children. The purpose of the joint party composition model is to narrow down the set of possible person participation choices modeled by the subsequent sub-model.
- And a person participation model that predicts which household members participate in the tour. Participation choice is modeled for each person sequentially. In this approach, a binary choice model is calibrated for each activity, party composition and person type. The model iterates through household members and applies a binary choice to each to determine if the member participates. The model is constrained to consider only those members with available time-windows overlapping with the generated joint tour.

FIGURE 5: JOINT TOUR MODELS



The joint tour frequency, composition, and participation models are described below.

Joint Tour Frequency

Number of Models:	1
Decision-Making Unit:	Households
Model Form:	Multinomial Logit
Alternatives:	21 (No Tours, 1 Tour segmented by purpose, 2 tours segmented by purpose combination)

Joint tour frequencies are generated by households and include the number and purposes of the joint tours. The explanatory variables in the joint tour frequency model include household variables, accessibilities, and other urban form type variables. One of the most significant variables in the joint tour frequency model is the presence and size of overlapping time-windows, which represent the availability of household members to travel together after mandatory tours have been generated and scheduled.   This formulation provides “induced demand” effects on the generation and scheduling of joint tours; the frequency and duration of mandatory tours affects whether joint tours are generated.

Joint Tour Composition

Number of Models:	1

Decision-Making Unit:	Joint Tour
Model Form:	Multinomial Logit
Alternatives:	3 (Adults-only, Children-only, Adults + Children)

Joint tour party composition is modeled for each tour and determines the person types that participate in the tour. The model is multinomial logit, and explanatory variables include the maximum time window that overlaps across adults, children and adults or children after mandatory tours have been scheduled. Other variables include household structure, area type, and the purpose of the joint tour.

Joint Tour Participation

Number of Models:	1
Decision-Making Unit:	Persons
Model Form:	Multinomial Logit
Alternatives:	2 (Yes or No)

Joint tour participation is modeled for each person and each joint tour. If the person does not correspond to the composition of the tour determined in the joint tour composition model, they are ineligible to participate in the tour. Similarly, persons whose daily activity pattern type is home are excluded from participating. The model relies on heuristic process shown in Figure 6 to assure that the appropriate persons participate in the tour as per the composition model. Explanatory variables include the person type of the decision-maker, the maximum pair-wise overlaps between the decision-maker and other household members of the same person type (adults or children), household and person variables, and urban form variables.

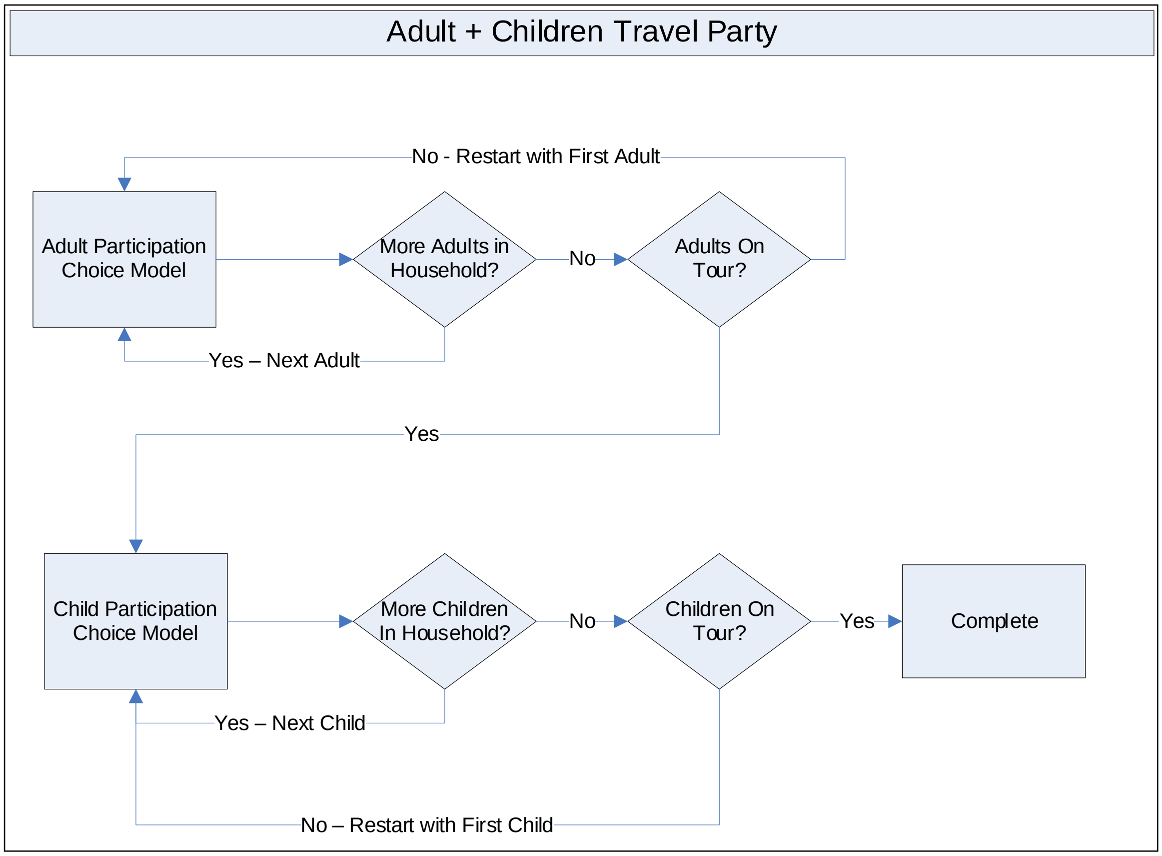


FIGURE 6: JOINT TOUR PARTICIPATION MODEL

Joint Tour Primary Destination Choice

Number of Models:	1
Decision-Making Unit:	Tour
Model Form:	Multinomial Logit
Alternatives:	Zones

The joint tour primary destination choice model determines the location of the tour primary destination. The destination is chosen for the tour and assigned to all tour participants. The model works at a zone level, and sampling of destination alternatives is implemented in order to reduce computation time, similar to mandatory location choice models. Explanatory variables include household and person characteristics, the tour purpose, the natural log of size (i.e. attraction) variables, round-trip mode choice logsum, distance, and other variables. Note that the mode choice logsum used is based on a “representative” time period for joint tours (outbound and return period is 2-3 P.M.), since the actual time period is not chosen until a later model.

Joint Tour Time of Day Choice

Number of Models: 1
Decision-Making Unit: Person
Model Form: Multinomial Logit
Alternatives: 190

After joint tours have been generated and assigned a primary location, the tour departure time from home and arrival time back at home is chosen simultaneously. The model is fully described above. However, a unique condition applies when applying the time-of-day choice model to joint tours. That is, the tour departure and arrival period combinations are restricted to only those available for each participant on the tour, after scheduling mandatory activities. Once the tour departure/arrival time combination is chosen, it is applied to all participants on the tour.

Individual Non-Mandatory Tour Frequency

Number of Models: 8 (segmented by 8 person types)
Decision-Making Unit: Person
Model Form: Multinomial Logit
Alternatives: 89 (Corresponding to most frequently observed combinations of number of individual maintenance and discretionary tours by purpose)

The third tour frequency model generates all non-mandatory individual tours at the person level. This model determines the number of both maintenance and discretionary tours simultaneously, at the person level, by purpose. There are six different kinds of maintenance and discretionary activities (escort, shop, other maintenance, eat out, visit, other discretionary), and a large number of possible combinations of each (assuming a maximum of 4 individual maintenance/discretionary tours per day, the number of possible combinations is 6^4 = 1,296 alternatives, many of which are not observed in the data). Therefore, a simplification of the alternatives is used in which only the most frequently observed combinations of tours by purpose and number are available in the logit model. There are 89 alternatives, as shown in Table 6. Certain alternatives are defined as “one or more tours” of a certain purpose. If such alternatives are chosen, a subsequent frequency model determines the exact number of tours for those cases (either 1 or 2), based on the person type and the number of mandatory and fully joint tours already generated for the decision-maker, as shown in Table 7. Only rows with probabilities for at least one additional tour are shown in the table.

Individual Non-Mandatory Tour Primary Destination Choice

Number of Models:	1
Decision-Making Unit:	Person
Model Form:	Multinomial Logit
Alternatives:	Zones

The individual non-mandatory tour primary destination choice model determines the location of the tour primary destination. The model works at a zone level, and sampling of destination alternatives is implemented in order to reduce computation time. Explanatory variables include household and person characteristics, the tour purpose, logged size (i.e. attraction) variables, round-trip mode choice logsum, distance, and other variables. Note that the mode choice logsum used is based on a "representative" time period for individual non-mandatory tours (outbound and return periods 2-3 P.M.), since the actual time period is not chosen until later in the system.

Individual Non-Mandatory Tour Time of Day Choice

Number of Models:	1
Decision-Making Unit:	Person
Model Form:	Multinomial Logit
Alternatives:	190 (combinations of tour departure hour and arrival hour back at home)

After individual non-mandatory tours have been generated and assigned a primary location, the tour departure time from home and arrival time back at home is chosen simultaneously. The model is fully described above. The tour departure and arrival period combinations are restricted to only those available for each participant on the tour, after scheduling individual mandatory tours and joint tours.

TABLE 6: INDIVIDUAL NON-MANDATORY TOUR FREQUENCY ALTERNATIVES

Alternative	Number of tours by purpose					
	Escorting	Shopping	Maintenance	Eating	Visiting	Discretionary
1	0	0	0	0	0	0
2	0	0	0	0	0	1+
3	0	0	0	0	1+	0
4	0	0	0	0	1+	1+
5	0	0	0	1+	0	0

Alternative	Number of tours by purpose					
	Escorting	Shopping	Maintenance	Eating	Visiting	Discretionary
6	0	0	0	1+	0	1+
7	0	0	0	1+	1+	0
8	0	0	0	1+	1+	1+
9	0	0	1+	0	0	0
10	0	0	1+	0	0	1+
11	0	0	1+	0	1+	0
12	0	0	1+	0	1+	1+
13	0	0	1+	1+	0	0
14	0	0	1+	1+	0	1+
15	0	0	1+	1+	1+	0
16	0	0	1+	1+	1+	1+
17	0	1+	0	0	0	0
18	0	1+	0	0	0	1+
19	0	1+	0	0	1+	0
20	0	1+	0	0	1+	1+
21	0	1+	0	1+	0	0
22	0	1+	0	1+	0	1+
23	0	1+	0	1+	1+	0
24	0	1+	0	1+	1+	1+
25	0	1+	1+	0	0	0
26	0	1+	1+	0	0	1+
27	0	1+	1+	0	1+	0
28	0	1+	1+	0	1+	1+
29	0	1+	1+	1+	0	0
30	0	1+	1+	1+	0	1+
31	0	1+	1+	1+	1+	0
32	0	1+	1+	1+	1+	1+
33	1	0	0	0	0	0
34	1	0	0	0	0	1+
35	1	0	0	0	1+	0
36	1	0	0	0	1+	1+
37	1	0	0	1+	0	0
38	1	0	0	1+	0	1+
39	1	0	0	1+	1+	0

Alternative	Number of tours by purpose					
	Escorting	Shopping	Maintenance	Eating	Visiting	Discretionary
40	1	0	0	1+	1+	1+
41	1	0	1+	0	0	0
42	1	0	1+	0	0	1+
43	1	0	1+	0	1+	0
44	1	0	1+	0	1+	1+
45	1	0	1+	1+	0	0
46	1	0	1+	1+	0	1+
47	1	0	1+	1+	1+	0
48	1	0	1+	1+	1+	1+
49	1	1+	0	0	0	0
50	1	1+	0	0	0	1+
51	1	1+	0	0	1+	0
52	1	1+	0	0	1+	1+
53	1	1+	0	1+	0	0
54	1	1+	0	1+	0	1+
55	1	1+	0	1+	1+	0
56	1	1+	0	1+	1+	1+
57	1	1+	1+	0	0	0
58	1	1+	1+	0	0	1+
59	1	1+	1+	0	1+	0
60	1	1+	1+	0	1+	1+
61	1	1+	1+	1+	0	0
62	1	1+	1+	1+	0	1+
63	1	1+	1+	1+	1+	0
64	0	0	0	0	0	0
65	0	0	0	0	0	1+
66	0	0	0	0	1+	0
67	0	0	0	0	1+	1+
68	0	0	0	1+	0	0
69	0	0	0	1+	0	1+
70	0	0	0	1+	1+	0
71	0	0	0	1+	1+	1+
72	0	0	1+	0	0	0
73	0	0	1+	0	0	1+

Alternative	Number of tours by purpose					
	Escorting	Shopping	Maintenance	Eating	Visiting	Discretionary
74	0	0	1+	0	1+	0
75	0	0	1+	0	1+	1+
76	0	0	1+	1+	0	0
77	0	0	1+	1+	0	1+
78	0	0	1+	1+	1+	0
79	0	1+	0	0	0	0
80	0	1+	0	0	0	1+
81	0	1+	0	0	1+	0
82	0	1+	0	0	1+	1+
83	0	1+	0	1+	0	0
84	0	1+	0	1+	0	1+
85	0	1+	0	1+	1+	0
86	0	1+	1+	0	0	0
87	0	1+	1+	0	0	1+
88	0	1+	1+	0	1+	0
89	0	1+	1+	1+	0	0

TABLE 7: INDIVIDUAL NON-MANDATORY TOUR EXTENSION CUMULATIVE PROBABILITIES

Person Type	Number of Mandatory Tours	Number of Joint Tours	Individual Discretionary Tour Purpose	Additional Tours		
				0	1	2
1	0	0	1	83.0%	100.0%	100.0%
2	0	0	1	76.9%	100.0%	100.0%
3	0	0	1	89.4%	100.0%	100.0%
4	0	0	1	75.0%	100.0%	100.0%
5	0	0	1	84.2%	100.0%	100.0%
6	0	0	1	71.4%	100.0%	100.0%
7	0	0	1	81.5%	100.0%	100.0%
8	0	0	1	75.0%	100.0%	100.0%
1	1	0	1	78.9%	100.0%	100.0%
2	1	0	1	60.0%	100.0%	100.0%
5	1	0	1	82.6%	100.0%	100.0%
6	1	0	1	83.7%	100.0%	100.0%
7	1	0	1	60.0%	100.0%	100.0%
1	0	1	1	84.2%	100.0%	100.0%
5	1	1	1	77.8%	100.0%	100.0%
1	0	0	2	89.3%	99.1%	100.0%
2	0	0	2	84.1%	99.3%	100.0%
3	0	0	2	97.1%	100.0%	100.0%
4	0	0	2	97.0%	100.0%	100.0%
5	0	0	2	87.0%	99.4%	100.0%
6	0	0	2	86.7%	100.0%	100.0%
7	0	0	2	97.1%	100.0%	100.0%
8	0	0	2	93.1%	100.0%	100.0%
1	1	0	2	88.5%	100.0%	100.0%
2	1	0	2	72.7%	100.0%	100.0%
3	1	0	2	97.1%	100.0%	100.0%
5	1	0	2	89.6%	99.3%	100.0%
6	1	0	2	88.5%	100.0%	100.0%
1	0	1	2	91.0%	99.3%	100.0%
2	0	1	2	88.0%	100.0%	100.0%
3	0	1	2	80.0%	100.0%	100.0%
6	1	1	2	96.5%	100.0%	100.0%
8	1	1	2	88.9%	100.0%	100.0%

Person Type	Number of Mandatory Tours	Number of Joint Tours	Individual Discretionary Tour Purpose	Additional Tours		
				0	1	2
1	0	0	3	93.6%	99.8%	100.0%
2	0	0	3	90.6%	100.0%	100.0%
3	0	0	3	97.9%	100.0%	100.0%
4	0	0	3	92.9%	100.0%	100.0%
5	0	0	3	90.2%	99.2%	100.0%
6	0	0	3	86.4%	100.0%	100.0%
7	0	0	3	94.7%	100.0%	100.0%
8	0	0	3	91.3%	100.0%	100.0%
1	1	0	3	89.3%	98.7%	100.0%
4	1	0	3	85.7%	100.0%	100.0%
5	1	0	3	91.6%	99.6%	100.0%
6	1	0	3	85.6%	98.4%	100.0%
7	1	0	3	100.0%	100.0%	100.0%
8	1	0	3	100.0%	100.0%	100.0%
1	0	1	3	91.6%	99.2%	100.0%
2	0	1	3	91.2%	98.2%	100.0%
6	0	1	3	83.3%	100.0%	100.0%
7	0	1	3	96.2%	100.0%	100.0%
1	1	1	3	97.8%	98.9%	100.0%
2	1	1	3	97.3%	100.0%	100.0%
5	1	1	3	99.6%	100.0%	100.0%
6	1	1	3	92.2%	98.0%	100.0%
1	0	0	4	92.2%	99.6%	100.0%
2	0	0	4	90.1%	100.0%	100.0%
3	0	0	4	99.7%	100.0%	100.0%
4	0	0	4	99.1%	100.0%	100.0%
5	0	0	4	92.2%	98.0%	100.0%
6	0	0	4	95.5%	100.0%	100.0%
8	0	0	4	95.5%	100.0%	100.0%
1	1	0	4	94.1%	97.1%	100.0%
2	1	0	4	92.6%	100.0%	100.0%
4	1	0	4	87.5%	100.0%	100.0%
5	1	0	4	91.5%	100.0%	100.0%
6	1	0	4	94.8%	99.4%	100.0%

Person Type	Number of Mandatory Tours	Number of Joint Tours	Individual Discretionary Tour Purpose	Additional Tours		
				0	1	2
7	1	0	4	66.7%	100.0%	100.0%
1	0	1	4	92.6%	98.8%	100.0%
2	0	1	4	90.4%	100.0%	100.0%
2	1	1	4	91.1%	100.0%	100.0%
6	1	1	4	96.3%	100.0%	100.0%
1	0	0	5	97.7%	100.0%	100.0%
2	0	0	5	98.2%	100.0%	100.0%
3	0	0	5	98.6%	100.0%	100.0%
8	0	0	5	87.5%	100.0%	100.0%
3	1	0	5	96.4%	100.0%	100.0%
5	1	0	5	98.6%	100.0%	100.0%
6	1	0	5	95.2%	100.0%	100.0%
1	0	1	5	92.7%	100.0%	100.0%
2	0	1	5	94.1%	100.0%	100.0%
3	1	1	5	97.3%	100.0%	100.0%
6	1	1	5	93.3%	100.0%	100.0%
1	0	0	6	93.8%	98.9%	100.0%
2	0	0	6	88.9%	100.0%	100.0%
3	0	0	6	96.7%	99.8%	100.0%
4	0	0	6	94.2%	100.0%	100.0%
5	0	0	6	88.0%	100.0%	100.0%
6	0	0	6	92.6%	100.0%	100.0%
7	0	0	6	96.8%	100.0%	100.0%
8	0	0	6	90.6%	100.0%	100.0%
1	1	0	6	85.9%	100.0%	100.0%
2	1	0	6	81.8%	97.0%	100.0%
4	1	0	6	95.2%	100.0%	100.0%
5	1	0	6	87.9%	99.8%	100.0%
6	1	0	6	86.3%	98.5%	100.0%
7	1	0	6	90.0%	100.0%	100.0%
1	0	1	6	92.8%	99.7%	100.0%
2	0	1	6	85.9%	99.2%	100.0%
5	0	1	6	92.0%	100.0%	100.0%
7	0	1	6	90.5%	100.0%	100.0%

Person Type	Number of Mandatory Tours	Number of Joint Tours	Individual Discretionary Tour Purpose	Additional Tours		
				0	1	2
1	1	1	6	98.3%	100.0%	100.0%
2	1	1	6	92.8%	98.8%	100.0%
3	1	1	6	98.3%	100.0%	100.0%
4	1	1	6	93.9%	100.0%	100.0%
6	1	1	6	93.8%	100.0%	100.0%

At-Work Sub-Tour Frequency

Number of Models:	1
Decision-Making Unit:	Persons
Model Form:	Multinomial Logit
Alternatives:	6 (None, 1 eating out tour, 1 business tour, 1 maintenance tour, 2 business tours, 1 eating out tour + 1 business tour)

Work-based sub-tours are modeled last and are relevant only for those persons who implement at least one work tour. These underlying activities are mostly individual (e.g., business-related and dining-out purposes), but may include some household maintenance functions as well as person and household maintenance tasks. There are six alternatives in the model, corresponding to the most frequently observed patterns of at-work sub-tours. The alternatives define both the number of at-work sub-tours and their purpose. Explanatory variables include household and person attributes, duration of the parent work tour, the number of joint and individual non-mandatory tours already generated in the day, and accessibility and urban form variables.

At-Work Sub-Tour Primary Destination Choice

Number of Models:	1
Decision-Making Unit:	Person
Model Form:	Multinomial Logit
Alternatives:	Zones

The at-work sub-tour primary destination choice model determines the location of the tour primary destination. The model works at a zone level, and sampling of destination alternatives is implemented in order to reduce computation time. Explanatory variables include household and person characteristics, the tour purpose, logged size (i.e. attraction) variables, round-trip mode choice logsum, distance, and other variables. Note that the mode choice logsum used is based on a “representative” time period for individual non-mandatory tours (outbound and return periods 2-3 P.M.). The model is constrained such that only destinations within a reasonable time horizon from the workplace are chosen, such that the tour can be completed within the total available time window for the sub-tour.

At-Work Sub-Tour Time of Day Choice

Number of Models: 1

Decision-Making Unit: Person

Model Form: Multinomial Logit

Alternatives: 190 (combinations of tour departure hour and arrival hour back at work)

After at-work sub-tours have been generated and assigned a primary location, the tour departure time from workplace and arrival time back at the workplace is chosen simultaneously. The tour departure and arrival period combinations are restricted to only those available based on the time window of the parent work tour.

Tour Mode Choice Model

Number of Models: Coefficients segmented by purpose (10 with at-work subtours)

Decision-Making Unit: Person

Model Form: Nested Logit

Alternatives: 10 main alternatives, 21 elemental alternatives

This model determines the “main tour mode” used to get from the origin to the primary destination and back. The tour-based modeling approach requires a certain reconsideration of the conventional mode choice structure. Instead of a single mode choice model pertinent to a four-step structure, there are two different levels where the mode choice decision is modeled:

- The tour mode level (upper-level choice),
- The trip mode level (lower-level choice conditional upon the upper-level choice).

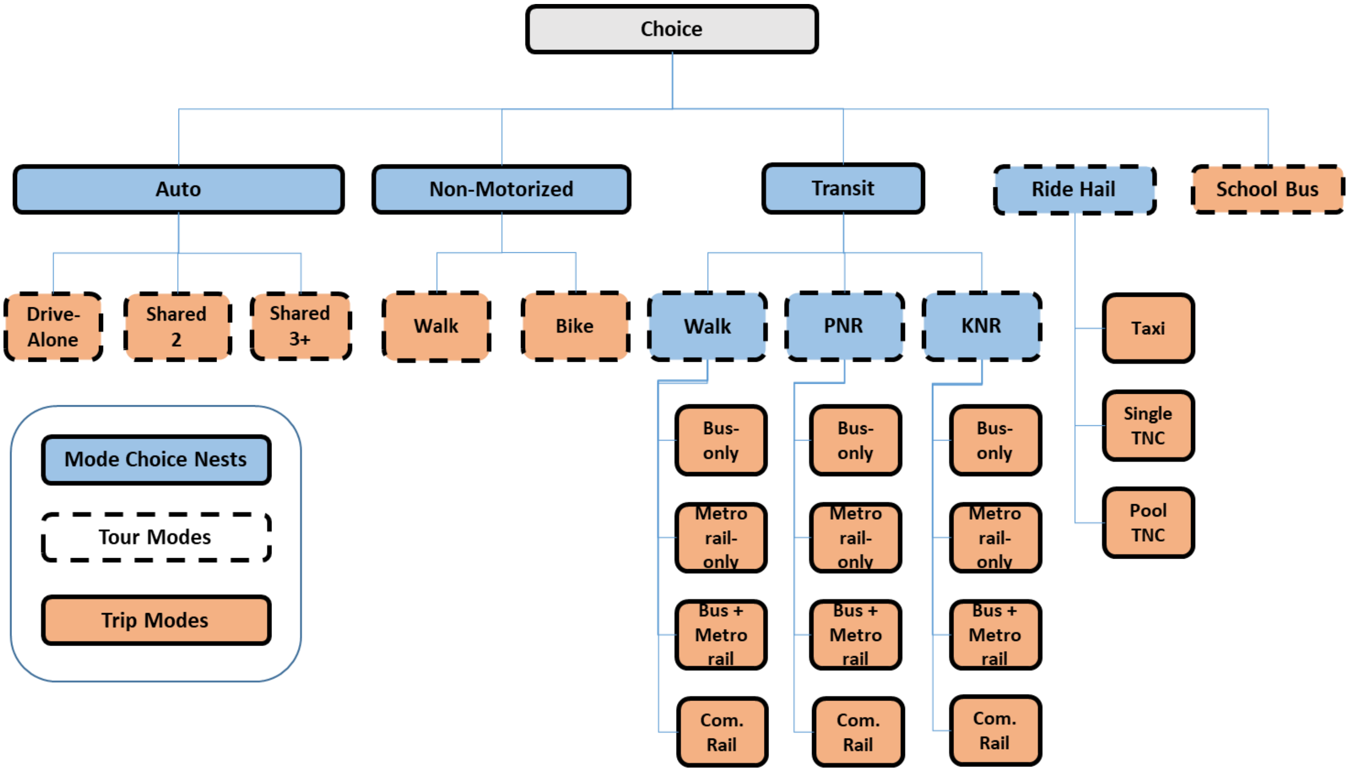


FIGURE 7: MODE CHOICE STRUCTURE

The tour mode level can be thought of as a mode preference model, while the trip mode choice model can be thought of as a mode switching model. Tour mode choice is used to constrain stop location choice as well as trip mode choice. The modes, or elemental alternatives, for both models are the same, but the higher level of the nesting structure constrains lower level decisions. This can be visualized in Figure 7, which shows the entire nesting structure for both tour and trip mode choice. However, for the purposes of downstream models, only tour modes (indicated by the 10 alternatives with dashed lines in the figure) are retained from the tour mode choice model. Lower level choices, such as transit path type (local only, premium only, or mixed) are used to calculate the upper level nest logsums but are not used to constrain trip mode choice.

The tour mode choice model is based on the round-trip level-of-service (LOS) between the tour anchor location (home for home-based tours and work for at-work sub-tours) and the tour primary destination. The tour mode is chosen based on LOS variables for both directions according to the time periods for the tour departure from the anchor and the arrival back at the anchor. This is one of the fundamental advantages of the tour-based approach. For example, a commuter can have very attractive transit service in the a.m. peak period in the outbound direction, but if the return home time is in the midday or later at night, the commuter may prefer private auto due to lower off-peak transit service.

The appropriate skim values for the tour mode choice are a function of the TAZ of the tour origin and TAZ of the tour primary destination. The tour mode choice model contains many household and person attributes, including income, auto sufficiency (typically defined as a comparison of number of autos to number of drivers in the household), age, etc. Urban form variables are also important, particularly related to the choice of non-motorized modes.

Intermediate Stop Frequency Model

Number of Models: 10 (By purpose plus one model for at-work subtours)

Decision-Making Unit: Person

Model Form: Multinomial Logit

Alternatives: Maximum 6 total, 3 per tour direction

The stop frequency choice model determines the number of intermediate stops on the way to/from the primary destination up to a maximum of 3 per direction, for a total of 8 trips per tour (four on each tour leg). However, for many tour purposes, the number of intermediate stops observed in the data is significantly less than 3 per direction. An additional constraint placed on intermediate stop models is that no stops are allowed on drive-transit tours. This is enforced to ensure that drivers who drive to transit pick up their cars at the end of the tour. If one of the drive-transit alternatives is chosen at the tour level, the choice applies to both outbound and return tour legs so that there is mode consistency. In most cases this would also ensure that the same parking location is chosen in both directions. Note that this is a simplification compared to other models which might allow stops on drive-transit tours.

Stop frequency is based on a number of explanatory variables, including household and person attributes, the duration of the tour (with longer durations indicating the potential for more stop-making) the distance from the tour anchor to the primary destination (with intermediate stop-making positively correlated to tour distance), and accessibility and urban form variables.

Once the number of intermediate stops is determined, each intermediate stop is assigned a purpose based on a frequency distribution created from observed data. The distribution is segmented by tour purpose, tour direction (outbound versus return) and person type and is based on survey data summaries. Work tours are also segmented by departure or arrival time period.

Intermediate Stop Location Choice Model

Number of Models: 10 (By purpose plus one model for at-work subtours)

Decision-Making Unit: Person

Model Form: Multinomial Logit

Alternatives: Zones

The stop location choice model predicts the location of stops along the tour other than the primary destination. The stop location model is structured as a multinomial logit model using a zone attraction size variable and route deviation measure as impedance. The alternatives are sampled from the full set of zones, subject to availability of a zonal attraction size term. The sampling mechanism is also based on accessibility between tour origin and primary destination and is subject to certain rules based on tour mode. All destinations are available for auto tour modes, so long as there is a positive size term for the zone. Intermediate stops on walk tours must be within 4 miles of both the tour origin and primary destination zones. Intermediate stops on bike tours must be within 8 miles of both the tour origin and primary destination zones. Intermediate stops on walk-transit tours must either be within 4 miles walking distance of both the tour origin and primary destination or have transit access to both the tour origin and primary destination. Additionally, only short and long walk zones are available destinations on walk-transit tours.

The intermediate stop location choice model works by cycling through stops on tours. The level-of-service variables (including mode choice logsums) are calculated as the additional utility between the last location and the next known location on the tour. For example, the LOS variable for the first stop on the outbound direction of the tour is based on additional impedance between the tour origin and the tour primary destination. The LOS variable for the next outbound stop is based on the additional impedance between the previous stop and the tour primary destination. Stops on return tour legs work similarly, except that the location of the first stop is a function of the additional impedance between the tour primary destination and the tour origin. The next stop location is based on the additional impedance between the first stop on the return leg and the tour origin, and so on.

Trip Departure Time Model

Number of Models: 1

Decision-Making Unit: Person

Model Form: Multinomial-Logit

Alternatives: Lookup from probabilities

The trip departure time model simulates the departure time for each trip on a tour, based on a lookup of probabilities by tour purpose, inbound versus outbound indicator, tour departure hour, and stop index. These probabilities are created from survey data.

Trip Mode Choice Model

Number of Models: 10 (By purpose plus one model for at-work subtours)

Decision-Making Unit: Person

Model Form:	Nested logit with constraints by tour mode
Alternatives:	15

The trip mode choice model determines the mode for each trip along the tour. Trip modes are constrained by the main tour mode. The linkage between tour and trip levels is implemented through correspondence rules (which trip modes are allowed for which tour modes). The model can incorporate asymmetric mode combinations, but in reality, there is a great deal of symmetry between outbound and inbound modes used for the same tour. In particular, symmetry is enforced for drive-transit tours, by excluding intermediate stops from drive-transit tours.

The tour and trip mode correspondence rules are shown in Table 8. Note that in the trip mode choice model, the trip modes are exactly the same as the modes in the tour mode choice model. However, every trip mode is not necessarily available for every tour mode, and tour mode constraints are generalized. Also note that "X" in the table means that a trip mode is NOT allowed in a tour mode. For example, the drive-alone tour mode allows only one type of trip mode (drive alone). The correspondence rules depend on a kind of hierarchy, which is similar to that used for the definition of transit modes. The hierarchy is based on the following principles:

1) The auto occupancy of the tour mode is determined by the maximum occupancy across all auto trips that make up the tour. Therefore, the auto occupancy for the tour mode is the maximum auto occupancy for any trip on the tour.

2) Transit tours can include auto shared-ride trips for particular legs. Therefore, "casual carpool", wherein travelers share a ride to work and take transit back to the tour origin, is explicitly allowed in the tour/trip mode choice model structure.

3) The walk mode is allowed for any trip on a tour except for drive-alone, wherein the driver must use the vehicle for all trips on the tour.

4) The transit mode of the tour is determined by the highest transit mode used for any trip in the tour according to the transit mode hierarchy as described in **TABLE 8**.

5) As previously mentioned, free shared-ride modes are also available in transit tours, albeit with a low probability.

The explanatory variables of the trip mode choice model include household and person variables, level-of-service between the trip origin and destination according to the time period for the tour leg, urban form variables, and alternative-specific constants segmented by tour mode.

TABLE 8: TENTATIVE RULES REGARDING TOUR MODE AND TRIP MODE CORRESPONDENCE

Trip Mode	Tour Mode									
	DRIVE ALONE	SHARED 2	SHARED 3+	WALK	BIKE	WLK ACCESS	PNR ACCESS	KNR ACCESS	RIDE HAIL	SCHOOL BUS
DRV ALONE				X	X	X	X	X	X	X
SHARE 2	X			X	X		X	X		
SHARE 3+	X	X		X	X		X	X		
WALK	X				X		X	X		
BIKE	X	X	X	X		X	X	X	X	X
WLK BUS	X	X	X	X	X		X	X	X	X
WLK METRORAIL	X	X	X	X	X		X	X	X	X
WLK MIX	X	X	X	X	X		X	X	X	X
WLK COM RAIL	X	X	X	X	X		X	X	X	X
PNR BUS	X	X	X	X	X	X		X	X	X
PNR METRORAIL	X	X	X	X	X	X		X	X	X
PNR MIX	X	X	X	X	X	X		X	X	X
PNR COM RAIL	X	X	X	X	X	X		X	X	X
KNR BUS	X	X	X	X	X	X	X		X	X
KNR METRORAIL	X	X	X	X	X	X	X		X	X
KNR MIX	X	X	X	X	X	X	X		X	X
KNR COM RAIL	X	X	X	X	X	X	X		X	X
TAXI	X	X	X	X	X					X
TNC - SINGLE	X	X	X	X	X					X
TNC - POOL	X	X	X	X	X					X
SCH BUS	X	X	X	X	X	X	X	X		

Note: X means that a trip mode is NOT allowed in a tour mode. For example, the drive-alone tour mode allows only one type of trip mode (drive alone).

Parking Location Choice

The parking location choice model is applied to tours with a destination in the urban area/city center with parking charges. This model is not currently implemented in ActivitySim due to concerns related to detailed data needs, therefore it will not be included in Phase I models.

The parking location choice model works in conjunction with the assignment to improve the realism of the auto component of assigned vehicle traffic. It is applied after destination and mode choice microsimulation has been completed. The destination end of auto-vehicle trips destined for the CBD are reallocated to parking location TAZs in accordance with model results for input to the assignment process. Two separate models are implemented -- one for work trips and one for non-work trips. The model is a two-step model where the first “choice” is whether the destination zone is the same as the parking zone and, if false, then the second choice is a location choice from 10 randomly selected CBD zones.

The parking models take advantage of the individual processing of records in micro-simulation. All records where a SOV trips is made to a CBD zone are individually re-processed. If the primary tour destination zone is not chosen for parking, then the record will be updated to indicate that the SOV trip had a different destination. Since the actual parking supply is used to regulate the allocation of parking locations, at least a rough balance between parking supply and demand is required.

4.4 | DESCRIPTION OF NEW ACTIVITYSIM FEATURES FOR MWCOG

We propose to modify the existing ActivitySim model system in the following ways:

1) Gen3, Phase I
    a. Implement a revised mode choice model structure, described above.
    b. Test for speed feedback convergence
2) Gen3, Phase II
    a. Replace aggregate accessibilities with disaggregate accessibilities
    b. Replace hourly time periods with half-hourly periods
    c. Estimate and implement transit pass ownership model
    d. Estimate and implement telework model
    e. Explicitly model internal-external transit travel
    f. Enhance models for autonomous vehicles
    g. Estimate and implement airport ground access models (optional depending on budget constraints)
    h. Estimate and implement an overnight visitor model (optional depending on data availability and budget constraints)

i. Estimate and implement a University student residential location choice model (optional subject to data availability and budget constraints)
j. Implement a Model Input Checker (optional)

We describe each Phase II enhancement in more detail below. Note that we discuss general approaches rather than a specific model design for many of these components since more effort will be required to analyze available data (the RSG team has not yet received household travel survey data), design, estimate, and implement a model. Such effort goes beyond the scope of a general model design document.

Disaggregate Accessibilities

The current ActivitySim accessibilities are relatively simple and aggregate. They essentially require running a simplified trip-based model prior to running ActivitySim. We plan to replace these accessibilities with a set of disaggregate accessibilities that utilize the tour mode and destination choice models implemented within the ActivitySim framework. This requires building a representative population that covers the market segments of interest, running tour mode and destination choice models for that synthetic population, and saving the logsums from these models in output files. The logsums are then used in the ActivitySim model for the resident population instead of zone-based accessibilities described above. The core functionality (saving logsums for tour mode and destination choice models) already exists in ActivitySim, so little additional development is required in order to implement this enhancement. The advantage of this approach is that any changes to the mode and destination choice models, like changes in alternative-specific constants or parameters, are automatically reflected in the accessibility terms used throughout the model system. This helps to ensure consistency across choices considered by the model.

Half-Hourly Periods

Currently hourly periods are used in the ActivitySim example model. We propose to replace hourly periods with half-hourly periods, which provides more consistency with potential DTA model development in Gen4. This enhancement is also proposed for the SEMCOG ActivitySim model, and was contributed to the ActivitySim project via the ARC ActivitySim implementation, so the work completed for these agencies can be leveraged for MWCOG.

Transit Pass Ownership Model

Transit pass ownership models have been applied in activity-based modeling systems as a mobility model, similar to parking pass and auto ownership models. These models attempt to represent the effect that ownership of a transit pass has on transit ridership and allow for the user to test different transit pass subsidy policies on transit ridership.[30] The choice has been structured as a binary ownership model (DaySim) or a joint choice with auto ownership (CT-RAMP). A similar model is planned for SEMCOG and could seed the MWCOG model development exercise.

The 2017-18 Regional Travel Survey asked workers who work outside the home which of the following benefits were provided by their employer:

- Free Parking
- Subsidized/Pre-Tax Benefit for Parking
- Subsidized/Pre-Tax Benefit for Transit Use
- Cash or other incentives for Carpool and Vanpool
- Cash or other incentives for walking or biking to work
- Electric vehicle charging station
- Secure bicycle parking facility
- None, employer doesn’t offer any transportation benefits
- Don’t know

It is also possible to identify federal workers from the employer type question asked:

1 Work for private for-profit firm/company

2 Work for nonprofit firm/organization

3 Work for federal government

4 Work for state or local government

5 Work for foreign governmental agency or international

governmental organization (e.g., World Bank, IMF, etc.)

6 Self employed

If the respondent traveled by transit, the survey collected information on how the respondent paid for their transit trip, with the following options:

1 SmarTrip (regular fare)

2 CharmCard (regular fare)

3 Single trip fare (cash, credit, ticket or token)

4 Round trip fare (cash, credit, ticket or token)

---

[30] see John Bowman, Mark Bradley, Puget Sound Regional Council, Resource Systems Group, and Cambridge Systematics. *SoundCast Activity-Based Travel Forecasting Model for PSRC Featuring DAYSIM—the Person Day Activity and Travel Simulator Model System Design*. September 25, 2014. and Parsons Brinckerhoff, Arizona State University. *MAG CT-RAMP Activity-Based Model Phase I Model Estimation Results*. August 2010.

5 Monthly pass (any transit operator or mode)

6 Weekly pass (any transit operator or mode)

7 Daily pass (any transit operator or mode)

8 TLC (Transit Link Card)

9 Senior/disabled pass, cash fare, or free fare

10 Youth/student pass, cash fare, or free fare

997 Other

998 Don't know

With these questions, it should be possible to model the probability of owning a transit pass, based on socio-demographic variables, origin-based transit accessibility, transit level-of-service between home and work (for workers), whether a subsidized/pre-tax benefit is offered by the employer, and the average cost of a transit pass. The choice set would be relegated to only persons who took transit; we would need to calibrate the model to the entire population, including persons who do not travel or who travel by other modes despite owning a transit pass. This would require data from transit agencies on total passes sold or would require borrowing relationships from household surveys from other large regions in which transit pass ownership was asked during the recruitment phase of the survey as opposed to the retrieval phase as was done for MWCOG.

To apply the model, we could simulate from the observed share of workers who have a transit pass provided by their employer by type (identified in synthetic population) for each person; this percentage could be modified for different scenarios in which there might be changes in the share of workers offered subsidized transit pass ownership. Those owning a transit pass would then be exposed to the cash transit fare in mode choice; note that this requires including the cost of a transit pass in the transit pass ownership model. Otherwise the model would not be responsive to transit cost for pass-holders.

Telecommute/telework Frequency Model

Telecommute or telework models are useful tools to analyze the impacts of telecommuting on travel frequency, complexity, and time-of-day. Recent model development work for San Diego Association of Governments indicates that workers who telecommute at least one day per week are more likely to go to work on their non-telecommute day, may be more likely to engage in some non-mandatory travel on days in which they do not go to work, and are more likely to have simpler non-work travel patterns when they do travel on those days.[31] These inter-relationships can be explicitly modeled, and the share of telecommuting workers can be modified to test potential changes in telecommute frequency for different future scenarios.

---

[31] RSG. *SANDAG Travel Model Enhancements To Support 2021 Long-Range Transportation Plan* Prepared For San Diego Association Of Governments, January 31, 2020.

The 2017-18 Regional Household Travel Survey collected information on the frequency of telecommute as follows:

1 Not eligible to telecommute
2 Eligible, but choose not to telecommute
3 Less than 4 weekdays per month
9 Weekends only
4 1 weekday a week
5 2 weekdays a week
6 3 weekdays a week
7 4 weekdays a week
8 5 weekdays a week

It should be possible to estimate a telecommute frequency model based on the response to this question, similar to the one estimated for SANDAG. The telecommute frequency variable would then be used as an explanatory variable in subsequent models including Coordinated Daily Activity Pattern model, Joint Tour Frequency, Non-Mandatory Tour Frequency, Time-of-Day Choice, and Intermediate Stop Frequency, depending on significance and reasonableness of parameters. To perform sensitivity and scenario tests, the alternative specific constant for different levels of telecommuting can be adjusted by the user. A similar model is planned for SEMCOG.

Internal/External and External/Internal Travel

In Phase I models, we plan to apply the existing IE/EI auto trip tables to the travel predicted for residents by ActivitySim. In Phase II, we suggest implementing an internal-external travel model for MWCOG residents. This model would predict the likelihood of a worker working outside of the region based on distance from the nearest external station and other explanatory variables. A Monte Carlo draw for each worker would indicate whether the worker works outside the region. If the worker works outside the region, an external station would be selected for the worker's regular workplace. Any tours that the worker generates would be sent to the external station instead of an internal zone. Work-based tours would be disallowed for the external worker. Note that this is essentially the same treatment for internal-external workers as is in the San Francisco County Transportation Authority CHAMP model.

Non-work external travel would be handled via a sub-model run before destination choice which predicts whether the primary destination on the tour is external. If so, a special destination choice model would be applied to predict the external station as the primary destination.

In order to develop these models, the household travel survey would need to be geocoded such that all external trip ends are allocated to the most likely external station. If the model considers transit, then transit networks would need to be extended to these external stations. Census journey-to-work data can be used to calibrate the flow of workers by internal county to external county. Traffic counts would be used to calibrate the external station choice model and transit on-board survey data would be used to calibrate and validate the tours and trips by mode and internal-external transit ridership on each route.

We propose to use factored observed external-internal trip tables for non-resident daytime trips into and out of the MWCOG region. The auto trip tables would be taken from the existing MWCOG model. Transit external-internal trip tables would be summarized from on-board survey data. Auto trip tables would be factored to forecasted changes in traffic counts at external stations. Transit trip tables would be held constant unless external forecasts of transit riders are available.

Autonomous Vehicles

RSG recently added Autonomous Vehicle (AV) to the San Diego Association of Governments (SANDAG) activity-based travel demand model system.[32] Some of these enhancements are based on previous work performed by RSG for a Federal Highway Administration project that analyzed AV demand using a combined Activity-based (DaySim) model and a Dynamic Traffic Assignment (DTA) model in Jacksonville Florida.[33] We propose to enhance ActivitySim to explicitly represent AVs in the Phase II deployment, in a manner that is similar to what is planned for SEMCOG. This will include the following:

1) Extension of the auto ownership model to consider autonomous (AV) versus human-controlled (HV) vehicles.
2) A simulation model that determines for each AV-owning household whether an AV is available for each tour.
3) Modifiers to the tour and trip mode choice model coefficients to represent AV-scenario assumptions.
4) Factors on trip tables to represent deadheading vehicle trips, or implementation of a the SANDAG AV intra-household vehicle sharing model if funds permit.

The AV models would not be estimated, calibrated, or validated due to lack of data. The coefficients used in forecasting the choice of mode for AV owning households are planned to be generic. The alternative-specific constants would be asserted (or set to 0, lacking any observed data).

Auto Ownership Enhancements

As shown in Figure 8, the auto ownership model would be extended to consider number of AVs versus HVs for each auto-owning choice. We assume that a household owning 4 vehicles would likely not own any autonomous vehicles and future scenarios that assume high levels of AV ownership would significantly reduce or even eliminate the share of 4+ vehicle owning households.

---

[32] RSG. *Model Enhancements to Support 2021 Regional Transportation Plan. San Diego Association of Governments*. January 31, 2020.

[33] Ben Stabler, Mark Bradley, Dan Morgan, Howard Slavin, Khademul Haque. Volume 2: Model Impacts of Connected and Autonomous/Automated Vehicles (CAVs) and Ride-Hailing with an Activity-Based Model (ABM) and Dynamic Traffic Assignment (DTA): An Experiment. Report FHWA-HEP-18-081. US Department of Transportation, Federal Highway Administration. April 2018.

Table 9 shows the coefficients to be added to the auto ownership model to capture likely socio-economic and mobility attributes related to AV ownership. The exponentiated values are also shown in order to illustrate the effect of the coefficient on the probability of AV ownership. These coefficients were adopted from recent AV scenario testing conducted by RSG using the Jacksonville DaySim model. They assume that younger and more wealthy households are more likely to own AVs, all else being equal. They also assume that households with longer work commutes would be more likely to own an AV. Although they are informed by current literature, *there is no way to statistically estimate these variables since there are no AV-owning households currently*.

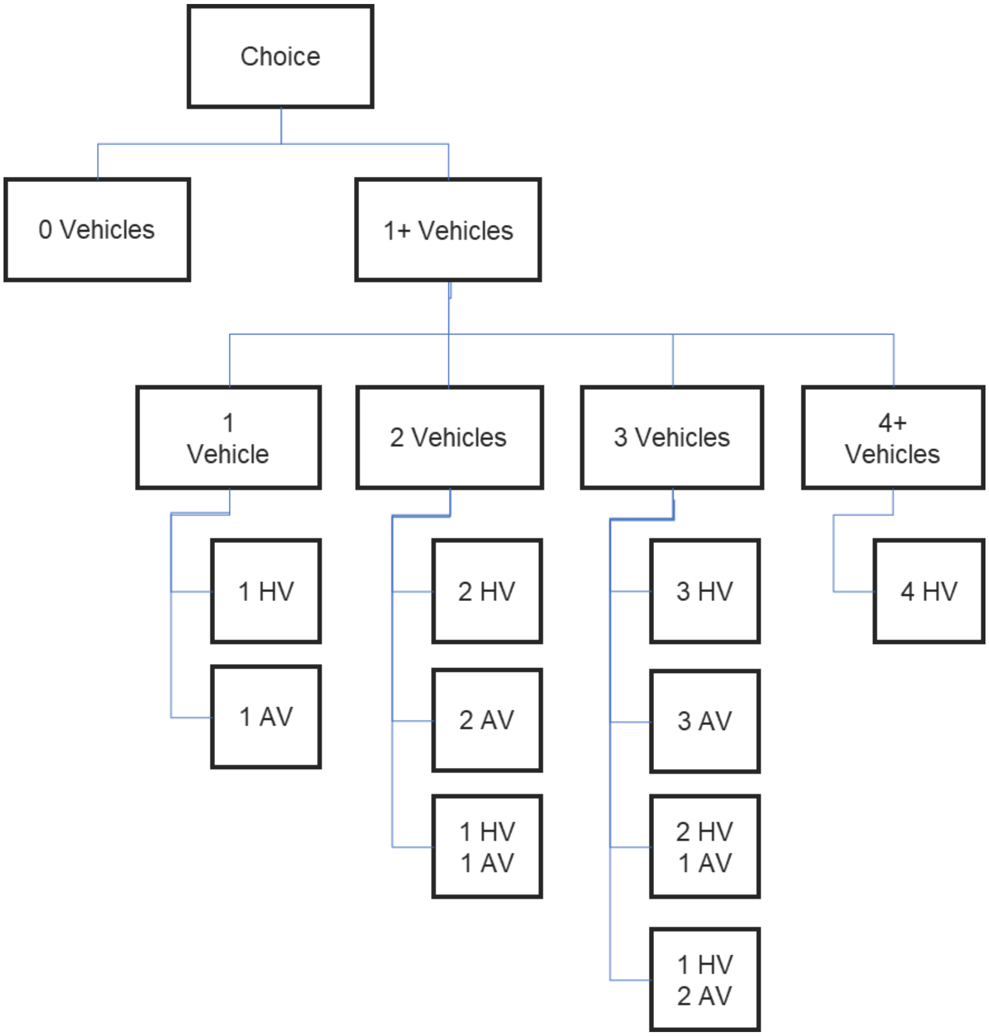
TABLE 9: AUTO OWNERSHIP VARIABLES AND COEFFICIENTS RELATED TO AUTONOMOUS VEHICLES

VARIABLE	COEFFICIENT	EXP (COEFFICIENT)
Household Income under $50k	-1.0000	0.37
Household Income 100k+	1.0000	2.72
Younger household (Number of persons 18 to 35 >= Number of Persons 65+	0.5000	1.65
Older household (Number of persons 18 to 35 < Number of Persons 65+	-1.0000	0.37
Hours of travel by auto for work, summed across all workers in household	0.2500	1.28

In addition to these variables, a set of alternative-specific constants can be applied and calibrated to reflect different levels of AV ownership according to scenario-specific targets. RSG has developed spreadsheets to calculate the target for each auto ownership choice based on a user-specified average vehicle ownership and a user-specified AV percentage of privately-owned vehicles. These constants will be calibrated for a few different levels of AV ownership (e.g. 20%, 50%, 80%) to be tested during sensitivity testing.

AV Tour Availability

FIGURE 8: AUTO OWNERSHIP MODEL WITH AV CHOICE



Households that own at least one of each type of vehicle (HV and AV) have a choice of which vehicle to use for each tour, which is taken into account not only when auto is the chosen mode but also when evaluating other modal options (walk, bike, transit, etc.). The MWCOG model will explicitly represent this decision, without introducing a full vehicle allocation model that would result in a much more complicated system. Instead, the AV availability model assumes that the starting point for the probability of an AV being available for the tour is equal to the share of AVs to total vehicles owned by the household. Since it is likely that the probability of an AV might be higher than the proportion of AVs owned by the household due to the flexibility offered by AVs in terms of repositioning, the user can set “probability boosts” based on the ratio of autos to drivers.

Mode Choice Enhancements

If the AV Tour Availability model indicates that an AV is available for the tour, a set of coefficient modifiers are applied to reflect differences in the actual or perceived travel time and cost of driving. These modifiers are specified for in-vehicle time, auto operating cost, parking cost, and terminal time, and are user-defined in the model properties file. Their base values are shown in Table 10. The in-vehicle time modifier is currently set to 0.75 to reflect the assumed increased comfort, productivity and reliability of driving in an AV. Parking cost is eliminated as it is assumed the vehicle would be sent to a free remote site for the duration of the activity or else sent home (the vehicle deadheading model reflects the actual decision). Auto operating cost modifier is 0.75 to reflect the increased fuel efficiency of an AV, and terminal time is eliminated from the utility of driving since it is assumed that an AV would provide curbside pick-up/drop-off service. All of these parameters can be modified by the user to test the effect of different assumptions regarding the operation and use of AVs.

TABLE 10: COEFFICIENT MODIFIERS

COEFFICIENT	MODIFIER
In-vehicle time	0.75
Parking cost	0
Auto operating cost	0.75
Terminal time	0

Empty Vehicle Trips and Assignment

There remains considerable uncertainty about the effects of autonomous vehicles on roadway capacity, safety, and congestion. However, research indicates that AV availability, if left unchecked, could lead to substantially higher vehicle miles of travel due to empty vehicle trips.[34] Therefore, we believe it is essential to account for empty vehicles in the traffic assignment process. This can be done one of two ways. A simple approach is to apply a user defined factor on the AV trip table. Of course, this factor must be defined by the user and is subject to much uncertainty. Further, the factor would apply to time periods, origins and destinations in the AV trip table, not necessarily reflect the origins, destinations, and time-of-day in which empty trips would occur. Another drawback of the approach is that such empty trips are not responsive to policy interventions such as the provision of remote parking lots or pricing.

---

[34] Mustapha Harb, Y Xiao, G Circella, P Mokhtarian, and J Walker, Projecting Travelers into a World of Self-Driving Vehicles: Estimating Travel Behavior Implications via a Naturalistic Experiment. Presented at the Transportation Research Board 97th Annual Meeting. January 2018.

A second, more sophisticated approach is to implement an AV routing model that explicitly sends empty vehicles between persons requiring their use. RSG developed such a model for SANDAG which could be adopted by MWCOG. However, the software is written in Java and compatible with the SANDAG ABM. It would either need a minor update to be consistent with ActivitySim output, or a major re-write to convert the code to Python. For this reason, we suggest discussing this further, and coordinate any related adoption of the code with SANDAG. The program would either be adopted in Phase II or else put off until Gen4.

Airport Ground Access Model (Optional)

Note that each of the “Optional” model features would require funding above the amount in the current, three-year contract (COG Contract 20-006).

The airport ground access model explicitly represents surface travel to and from each airport in the MWCOG region for arriving and departing passengers. The purpose of this model is to capture the demand of airport travel made by air passengers on transport facilities and allow MWCOG to test the impacts of various parking price and supply scenarios at each airport.

The model would be implemented in the ActivitySim framework, which includes implementing the model as a series of runnable submodel steps run via the data pipeline, implementing calculations/utilities as user input expression files, adding sufficient code test coverage, model re-estimation functionality, built-in user documentation, and multiprocessing if desired.

The airport ground access model would have the following features:

- A disaggregate micro-simulation treatment of air passengers, with explicit representation of duration of stay or trip in order to accurately represent costs associated with various parking and modal options
- The full set of modes within the MWCOG region, including auto trips by occupancy, transit trips by mode, consideration of TNCs versus traditional taxi, etc.
- Forecasts of airport ground access travel based upon official enplanement projections

The model flow and inputs are shown in Figure 9, and described in detail in the following sections. The model has been developed for the San Diego Association of Governments and is used to represent travel to/from San Diego International Airport and the Cross-Border Express facility.

Airport Model Trip Purposes

There are four trip purposes coded based on the resident status of air passengers and the purpose of air travel, as follows:

- **Resident Business:** Business travel made by MWCOG residents (or residents of neighboring counties who depart from the airport)
- **Resident Personal:** Personal travel made by MWCOG residents (or residents of neighboring counties who depart from the airport)
- **Visitor Business:** Business travel made by visitors to MWCOG region (or a neighboring county)

- **Visitor Personal:** Personal travel made by visitors to MWCOG region (or a neighboring county)

Airport Model Trip Mode

The model of airport ground access is trip-based rather than tour-based since the model applies to either arriving or departing airport passenger parties. If private auto is used to access the airport, the choice of parking versus curbside pickup/dropoff is explicitly represented. For travelers that park, the chosen lot (terminal, airport remote lot, private remote lot) is explicit as well. Also note that auto occupancy is not a choice for airport ground access trips. Auto occupancy is based upon travel party size, which will be simulated as part of the attribution of ground access trips.

Airport Model Inputs

The model system requires the following exogenously-specified inputs (note that three additional data sets are required in addition to the data currently input to the resident activity-based models):

- **Enplanement Forecasts at each airport:** The total number of yearly enplanements without counting transferring passengers (local originations) at each airport, and an annualization factor to convert the yearly enplanements to a daily estimate. This would be given for each simulation year.
- **Traveler characteristics distributions:** There are a number of distributions of traveler characteristics that are assumed to be fixed but can be changed by the analyst to determine their effect on the results. These include the following:
    - The distribution of travelers by purpose
    - The distribution of travelers by purpose and household income.
    - The distribution of travelers by purpose and travel party size.
    - The distribution of travelers by purpose and trip duration (number of nights).
    - The distribution of travelers by purpose, direction (arriving versus departing), and time period departing for airport.
- **Land-use data.** The population and employment (by type) in each TAZ, parking cost and supply, etc. This data provides sensitivity to land-use forecasts in the region. These are the same data sets as are used in the resident activity-based model.
- **Level-of-Service data.** Auto and transit network level-of-service between each transportation analysis zone. This provides sensitivity to auto network supply and cost. These are the same data sets as are used in the resident activity-based model.

Airport Model Description

This section describes the model system briefly.

1. **Trip Enumeration and attribution:** A total number of airport trips is created by dividing the input total enplanements (minus transferring passengers) by an annualization factor. The result will be divided by an average travel party size to convert passengers to travel parties. This will be converted into a list format that will then be exposed to the set of traveler characteristic distributions, as identified above, to attribute each travel party with the following characteristics:

- Travel purpose
- Party size
- Duration of trip
- Household income
- Trip direction (it will be assumed that 50% of the daily enplanements are arriving passengers and 50% are departing passengers)
- Departure time for airport

1. **Trip Models**
    1.1. Trip origin: Each travel party will be assigned an origin TAZ.
    1.2. Trip mode: Each travel party will be assigned a trip mode (see Figure 10).

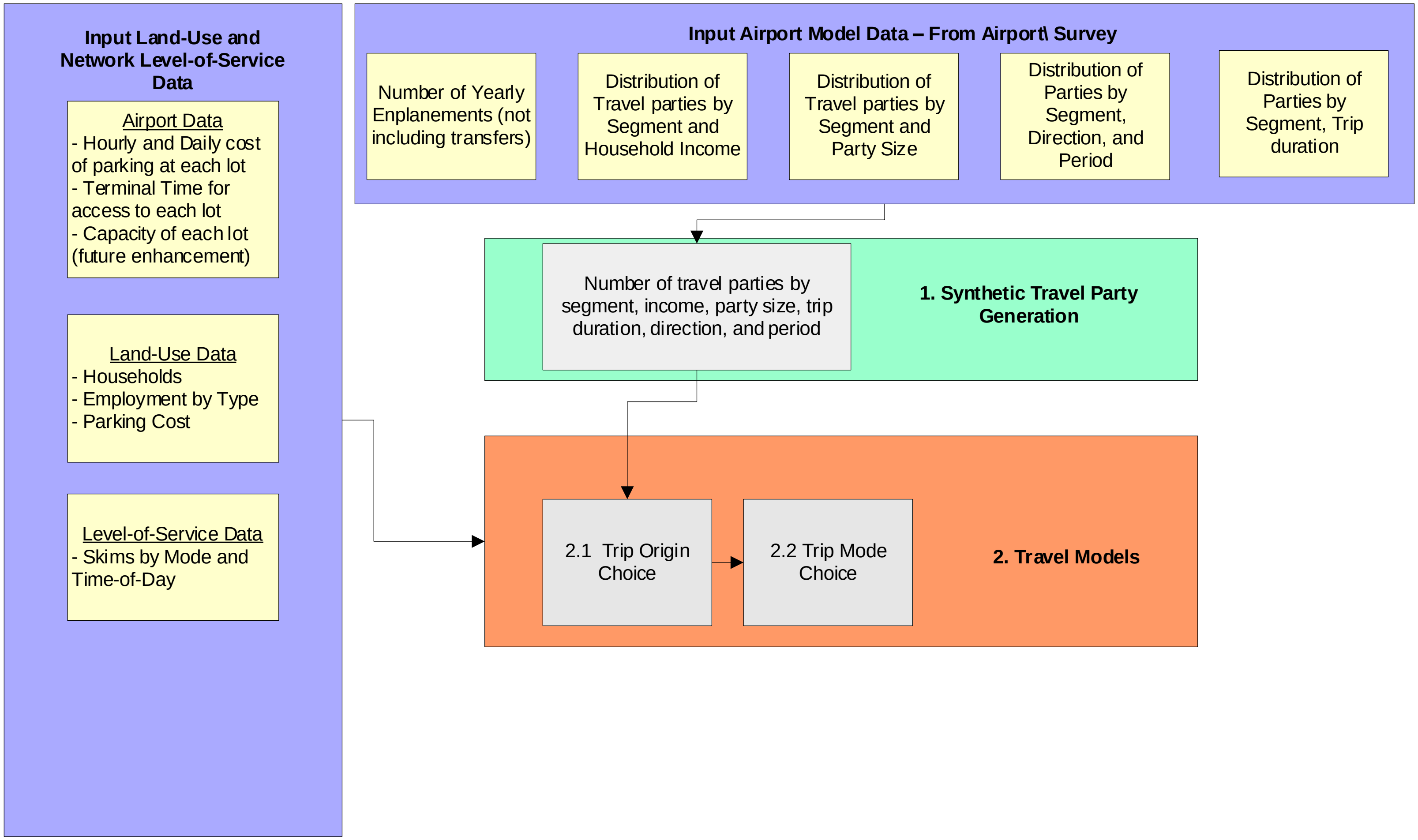


FIGURE 9: AIRPORT GROUND ACCESS TRAVEL MODEL

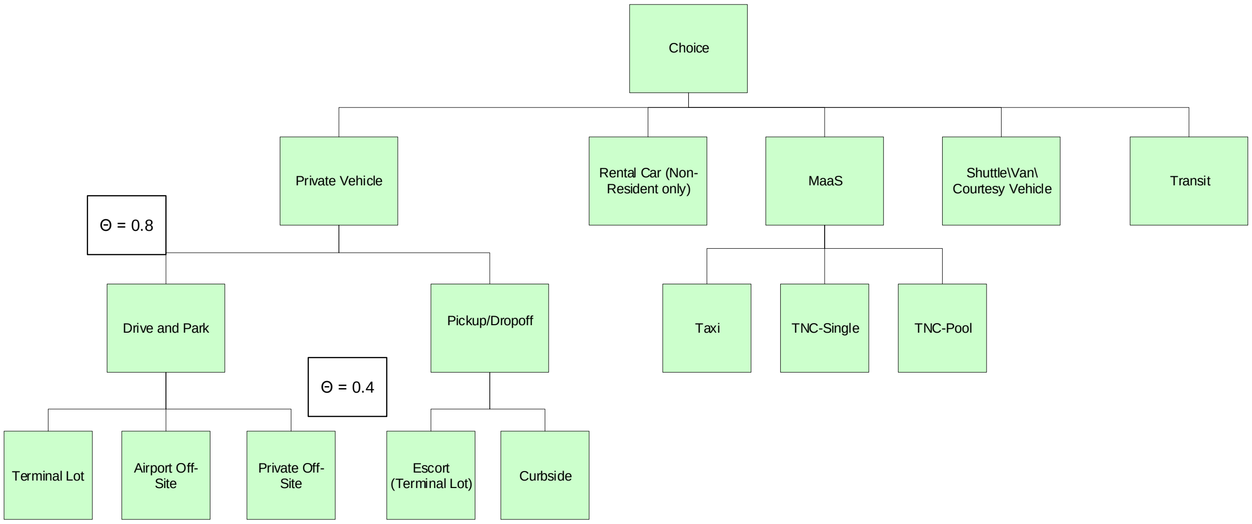


FIGURE 10: AIRPORT GROUND ACCESS MODE CHOICE STRUCTURE

Overnight Visitor Model (Optional)

The purpose of this model is to capture the demand of overnight visitor travel on transport facilities in the MWCOG region. Visitors whose stay does not involve spending a night are handled in the external-internal model. The visitor model has the following features:

- A disaggregate micro-simulation treatment of visitors by person type, with explicit representation of party attributes
- Special consideration of unique visitor travel patterns, including rental car usage and visits to MWCOG attractions like the capital, museums, etc.
- The full set of modes within the MWCOG region, including auto trips by occupancy, transit trips, non-motorized trips, and taxi and TNC trips.

The model would be implemented in the ActivitySim framework, which includes implementing the model as a series of runnable submodel steps run via the data pipeline, implementing calculations/utilities as user input expression files, adding sufficient code test coverage, model re-estimation functionality, built-in user documentation, and multiprocessing if desired.

Visitor Travel Parties

Visitors are generated for two visitor segment types:

- **Business:** Self-identified as business traveler, or self-identified as both business and personal but took at least one business purpose trip on travel day
- **Personal:** Self-identified as personal traveler, or self-identified as both business and personal but took no business purpose trips on travel day.

The model generates visitor parties by segment by applying separate occupancy rates to hotels and households, which must be obtained for the MWCOG region. The model then applies separate distributions of visitor parties by segment to hotel visitor parties and household visitor parties separately. Visitor parties are attributed with household income based upon the distribution of parties by visitor segment and income. Note that party size and auto availability are attributed on a tour-by-tour basis, since these attributes can change depending on which tour is undertaken and which day it is taken on.

Next, tours are generated by visitor parties and attributed with party size, auto availability, and income attributes. There are three tour purposes:

- **Work:** Business travel made by Business travelers
- **Recreational:** All other recreational purposes besides dining
- **Dining:** Travel to eating establishments

Tour purpose is coded according to a hierarchy of trip purposes, with work at the top and dining last. Each travel party can generate one or more tours of each purpose on any given day. The average size of the travel parties must be obtained from visitor survey data or from an external source.

If a visitor party drives into the region, either in a personal or rental vehicle, they are assumed to have access to a car during their stay. If the visitor flew into the region and rented a car, they were also assumed to have access to a car. Persons who do not fit into either of those categories are assumed to have no vehicle for use in mode choice.

The model flow and inputs are shown in Figure 11, and described in detail in the following sections.

Model Inputs

The model system requires the following exogenously-specified inputs (note that three additional data sets are required in addition to the data currently input to the resident activity-based models):

- **Traveler characteristics distributions**. There are a number of distributions of traveler characteristics that are assumed to be fixed but can be changed by the analyst to determine their effect on the results. These include the following:
    - Rates of visitor occupancy for hotels and separately for households
    - Shares of visitor parties by visitor segment for hotels and separately for households
    - The distribution of visitor parties by household income.
    - The distribution of business segment travel parties by number of tours by purpose
    - The distribution of personal segment travel parties by number of tours by purpose
    - The distribution of visitor tours by tour purpose and party size
    - The distribution of visitor tours by tour purpose and auto availability

  - The distribution of visitor tours by outbound and return time-of-day and tour purpose
  - The distribution of visitor tours by frequency of stops per tour by tour purpose, duration, and direction
  - The distribution of stops by stop purpose and tour purpose
  - The distribution of stops on outbound tour legs by half-hour offset period from tour departure period and time remaining on tour
  - The distribution of stops on inbound tour legs by half-hour offset period from tour arrival period and time remaining on tour
- **Land use data.** The population, employment (by type), and number of hotel rooms in each TAZ, parking cost, etc. This data provides sensitivity to land-use forecasts in the MWCOG region. These are the same data sets as are used in the resident activity-based model.
- **Level of service data.** Auto and transit network level-of-service between each TAZ. This provides sensitivity to transport network supply and cost. These are the same data sets as are used in the resident activity-based model.

It is possible that COG may want to add new questions to future regional air passenger surveys to collect this information. It is also possible that some of this information could be obtained from Big Data.

Model Description

This section describes the model system briefly.

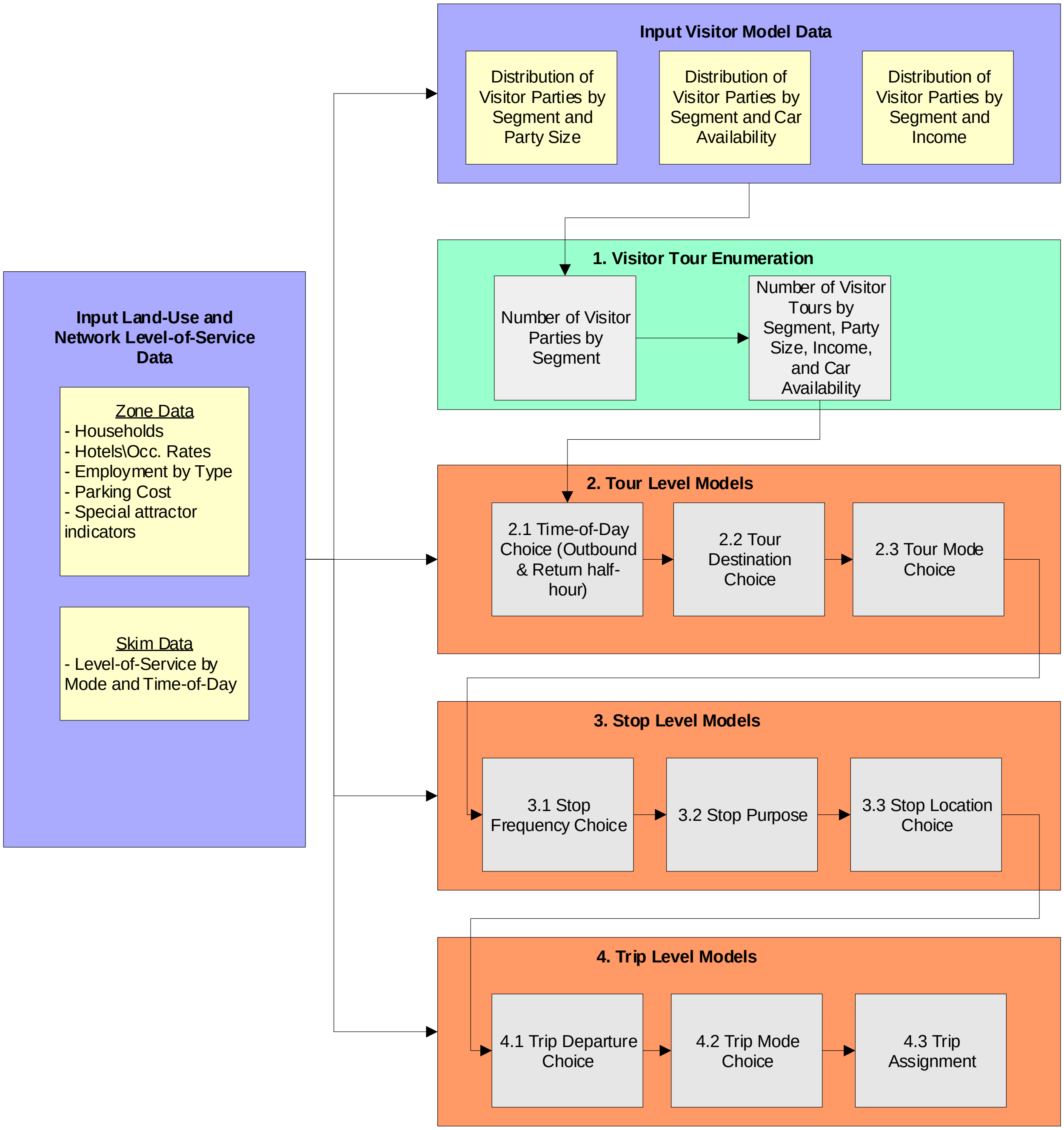


FIGURE 11: VISITOR MODEL DESIGN

1. **Visitor Tour Enumeration:** Visitor travel parties are created by visitor segment based upon input hotels and households. Travel parties are attributed with household income. Tours by purpose are generated for each party. Each tour is attributed with auto availability and party size. The tour origin TAZ is set to the TAZ where the tour was generated.
2. **Tour Level Models**
   2.1. Tour Time of Day: Each tour is assigned a time of day, based on probability distribution.

- 2.2. Tour Destination choice: Each tour is assigned a primary destination, based on the coefficients estimated through a multinomial logit model.
- 2.3. Tour Mode Choice: Each tour selects a preferred primary tour mode, based on an asserted nested logit model (the resident tour mode choice model).

3. **Stop Models**
   - 3.1. Stop Frequency Choice: Each tour is attributed with a number of stops in the outbound direction and in the inbound direction, based upon sampling from a distribution.
   - 3.2. Stop Purpose: Each stop is attributed with a purpose, based upon sampling from a distribution.
   - 3.3. Stop Location Choice: Each stop is assigned a location based upon a multinomial logit model (asserted based upon resident stop location choice models)
4. **Trip Level Models**
   - 4.1. Trip Departure Choice: Each trip is assigned a departure time period based upon sampling from distributions.
   - 4.2. Trip Mode Choice: Each trip within the tours selects a preferred trip mode, based on an asserted nested logit model.
   - 4.3. Trip Assignment: Each trip is assigned to the network.

University Student Residential Location Choice Model (Optional)

The university student residential location choice model is used to explicitly model residential location of students for major universities. These are universities where a large portion of the student body has relocated their residential location to be proximate to the university campus. Students living in group quarters would already be handled in the core population synthesizer. However, one key change to the synthesizer input would be to tag each university student housing unit with the university that it belongs to. This new model component specifically focuses on students living in non-group quarter households. These students are not readily identifiable from census data or other inputs. Therefore, a model is used to base their residential location on distance from the university that they attend. The model is a destination choice model, but unlike the destination choice models in ActivitySim, in which a non-home location is chosen based on the home location, the accessibility of each non-home location and the size or opportunities of activities in each TAZ, this model chooses a home location based on the non-home location, and the size or number of available housing opportunities in each TAZ.

The model would be implemented in the ActivitySim framework, which includes implementing the model as a series of runnable submodel steps run via the data pipeline, implementing calculations/utilities as user input expression files, adding sufficient code test coverage, model re-estimation functionality, built-in user documentation, and multiprocessing if desired.

The model requires the following inputs

- **University data**: For each major university
    - The number of students enrolled by undergraduate vs graduate degree (useful for segmenting distance and/or size terms in destination choice)
    - The number of students housed in group quarters
    - Total square footage of classroom space in each TAZ are a useful input for universities that span multiple zones.
    - Quantity of university parking by TAZ is useful if applying a parking location choice model for university students and/or faculty. Also, university parking policies regarding pricing, accessibility, and transit pass ownership are useful.
- **Land use data:** A central TAZ for each major university, group quarters housing units for each university, total households and/or a count of family and non-family households in each TAZ.
- **Level-of-service data:** Auto and transit network level-of-service between each TAZ. This provides sensitivity to transport network supply and cost. These are the same data sets as are used in the resident activity-based model.

The model works by first determining the number of students living in non-group quarters housing in each TAZ. This data is then introduced as a person constraint in population synthesis. Persons generated by this constraint are “tagged” with the major university that they attend. University tours generated by these persons are constrained to only the set of TAZs associated with their school.

ActivitySim has a college/university student person type that is used to generate university tours and influence other non-work travel. However, university students often have special mode and mobility options, including lower rates of auto ownership, higher rates of walking and biking, free or discounted transit fares, designated parking lots, access to university shuttles, etc. All of these unique characteristics can be taken into account in the model but may require specialized inputs (such as parking data) or sub-models (such as parking location choice).

Model Input Checker (Optional)

Activity-based travel models rely on data from a variety of sources (zonal data, highway networks, transit networks, synthetic population, etc.). A problem in any of these inputs can affect the accuracy of model outputs and/or can result in run time error(s) during the model run. It is important that the analyst carefully prepare and review all inputs prior to running the model. However, even with the best of efforts, sometimes errors in input data remain undetected. In order to aid the analyst in the input checking process, an automated Input Checker Tool can be implemented, similar to work already performed for Oregon Department of Transportation for use with the Southern Oregon activity-based model.

The Input Checker Tool (inputChecker) was implemented in Python and makes heavy use of the Python pandas and numpy packages. The main inputs to inputChecker are a list of ABM input tables, a list of QA/QC checks to be performed on these input tables and the actual ABM inputs in CSV format. All CSV inputs are read as pandas DataFrames (2-dimensional data tables). The input checks are specified by the user as pandas expressions which are solved by the inputChecker on the input pandas DataFrames. The input checks can be completely customized by the user. Examples of input checks include:

- The synthetic population is internally consistent
- The number of workers in the synthetic population is consistent with total employment in the land-use data
- The number of zones in the input TAZ file is consistent with the network
- The network has all fields coded
- etc.

The inputChecker generates a log file summarizing the results of all of the input checks.

The program executes the following steps:

1. Read Inputs:

First, inputChecker reads all the inputs specified in the list of inputs and copies them to the `inputChecker/inputs` directory. After assembling all inputs in the `inputChecker/inputs` directory, all the inputs are loaded as pandas DataFrames.

2. Run Checks

Next, the list of input checks is read. inputChecker loops through the list of input checks and evaluates the checks. The result of each check is sent to the logging module. The user must specify the severity level of each check as - Fatal, Logical or Warning.

3. Run Self Diagnostics

Besides the checks specified by the user, inputChecker also performs self-diagnostics to check for missing values in inputs. The severity level for the automated missing value checks is set via the `config/settings.csv` file.

4. Generate LOG File and Return Error Status

The final step is to generate the inputChecker log file. The inputChecker log includes results of all checks. The checks that failed are moved up in order of the severity-level specified for the test. A summary of inputChecker results is also generated to be read by the RunModel.bat DOS batch file to generate a reminder message for the user at the end of the ABM model run. An appropriate exit code is returned depending on the outcome of the inputChecker run. The table below describes the various outcomes and the associated exit codes:

- 0: inputChecker ran successfully with no fatal checks fails
- 1: inputChecker did not run successfully due to errors
- 2: inputChecker ran successfully with at least one fatal check fail

With a return code of 0, the model continues to run. A reminder message is generated at the end to check the inputChecker log file. In the case where the inputChecker results in an error, the model run is aborted. If the inputChecker completes with at least one fatal check fail, the model run is aborted and the user is directed to check the inputChecker log file.

5.0 MODEL DEVELOPMENT PLAN

Below we outline our plan for model development, by phase.

5.1 | GEN3 MODEL PHASE I

In the initial phase of deployment, the RSG team will undertake the following tasks. Note that the budget and schedule spreadsheet contain tasks for project management, including hours for bi-weekly conference calls and invoicing. These tasks are not described below.

Task 1: Population Synthesizer Deployment

The RSG team will finalize controls for the population synthesizer and assemble those controls. The RSG team will implement an automated procedure to scale controls to base and future year inputs. The RSG team will implement the population synthesizer software (PopulationSim) and automated procedures to prepare ActivitySim input household and person files. The RSG team will validate the ActivitySim inputs and parameters to improve goodness-of-fit to controls if deemed necessary by the model development team. The RSG team will develop and implement methods to create future year controls. The RSG team will document the population synthesis procedures, controls, base-year and future-year goodness of fit, inputs and outputs in a Technical Memorandum.

Deliverables

- Base-year population synthesis inputs, software, python scripts, and outputs
- Future-year population synthesis inputs, software python scripts and outputs
- MWCOG Gen3 Model Development Technical Report or Memorandum: Base and Future Year Synthetic Population

Task 2: Data Development

The RSG team or the MWCOG team will re-expand the 2017-18 Regional Travel Survey (RTS) consistent with Census controls assembled in Task 1. The RSG team will ensure that the expansion meets the objectives of the project and results in reasonable expansion factors. This will be achieved with careful use of PopulationSim survey data expansion functionality (since it does survey re-weighting in additional to population synthesis). The RSG team will code RTS tours, stops, etc. consistent with ActivitySim.

The RSG team will review, analyze, and code on-board survey data for use in travel demand model development. The RSG team will expand on-board survey data to result in a consistent dataset that replicates observed transit boardings by operator, route (if available) and transfer rates. The expansion will be performed using PopulationSim survey data expansion functionality. The RSG team will create transit on-board survey trip tables and assign those trip tables to the transit network, to determine the quality of the transit on-board survey data, explore potential transit network coding issues, and calibrate PT transit assignment parameters. An additional analysis will explore model behavior and goodness-of-fit when enabling PT link-level transit capacity restraint. The RSG team will document the activities in this task in a technical memorandum.

Deliverable(s)

- 2017-18 Regional Travel Survey re-expansion setup
- Re-expanded and coded 2017-18 Regional Travel Survey
- Expanded and coded transit on-board survey data
- Revised PT parameters
- MWCOG Gen3 Model Development Technical Report or Memorandum: Data Development

Task 3: Phase I ActivitySim Deployment

The RSG team will implement the Phase I ActivitySim model. The implementation will involve a transfer of the existing ActivitySim model to the maximum extent possible. This will require revisions to the existing skimming and assignment procedures to build the required skims by time of day, implement transit skims with drive as an egress option, and convert output skims to Open Matrix (.omx) format. The team will revise model implementation code to read output trip tables from ActivitySim and eliminate unneeded trip-based model code from the run process. The RSG team will implement internal/external transit trip tables in transit assignment. The RSG team will modify ActivitySim Utility Expression Calculator (UEC) files to be consistent with MWCOG land-use input data and the MWCOG mode choice model structure. The RSG team will estimate work size terms by cross-classifying Census Public Use Microdata Sample (PUMS) workers by household income and North American Industry Code (NAICS) grouped into MWCOG employment categories. In the initial deployment, we will assert size terms consistent with existing ActivitySim coefficients but applied to MWCOG data. This will allow us to get the ActivitySim model up and running quickly and assess model performance. We will deploy the activity-based model output visualizer to compare model outputs to observed data.

Deliverable(s)

- Gen3 Phase I Initial Deployment Model and model documentation
- ActivitySim Visualizer HTML file(s)

Task 4: Phase I Model Estimation

In the Phase I model deployment, we will estimate only two key models: tour mode choice and tour destination choice. These key models play a key role in the ability of the model to replicate observed behavior of MWCOG residents, and previous research indicates that destination choice models are the least transferable component between regions. The RSG team will assemble required datasets and utilize the model estimation functionality currently being implemented in the ActivitySim platform to estimate models by tour purpose. The RSG team will implement the estimated coefficients in the UEC spreadsheets and test the revised model system. The RSG team will document the estimation results in a technical memorandum.

Deliverable(s)

- Model estimation datasets
- Revised Gen3 Phase I Model
- MWCOG Gen3 Model Development Technical Memorandum: Tour Mode and Destination Choice Model Estimation Results

Task 5: Phase I Model Calibration & Validation

The Phase I models will be calibrated and validated to observed data.

Model Calibration

Following is a list of calibration summaries to be performed for each model component as well as the source of the observed data. The goal of this process in Phase I is to have a model suitable for testing purposes. This model will be suitable for sensitivity analysis and provide the project team with a good understanding of what aspects of calibration and validation to refine in Phase II. While all of the summaries listed below will be created for the Phase I model system, we would accept higher levels of error in Phase I calibration compared to Phase II calibration. For example, we may accept matching transit trips roughly by tour mode in Phase I, while in Phase II we would expect to match on transit trips by technology/transit sub-mode and a closer match to route-level boardings.

Auto ownership:

- Households by vehicles available and workers (Census, ACS PUMS)
- Households by vehicles available and household income (Census, ACS PUMS)
- Households by vehicles available and district (CTPP, ACS 5-year summaries)
- Share of 0-auto households by Census Tract (Census)

Free Parking Eligibility:

- Workers by free parking available and district, if available (household survey)

Work Location Choice Model:

- Share of workers who work from home (household travel survey)
- Home to work average distance and trip length frequency distribution (household travel survey)
- Workers by place of residence and place of work, district level (household travel survey, CTPP, ACS 3 or 5-year summaries)

University, school location choice:

- Home to school average distance and trip length frequency distribution (household travel survey)
- Students by place of residence and place of school, district level (household travel survey)

Coordinated Daily Activity Pattern Model:

- Share of persons by person type and daily activity pattern (household survey)

- Share of households by presence of fully joint tours and household size (household survey)

Mandatory Tour Generation Model:

- Share of mandatory tour generation model alternatives by person type (household survey)

Fully Joint Tour Generation/Composition and Participation Models:

- Share of fully joint tour generation/composition alternatives (household survey)
- Share of fully joint tours by number of persons participating (household survey)

Individual Non-Mandatory Tour Generation Model:

- Share of non-mandatory tours by purpose, number, and person type (household survey)
- Total number of individual non-mandatory tours by person type (household survey)

At-work subtour Frequency Model:
Share of work tours by at-work subtours

Non-mandatory Tour Location Choice:

- Home to primary destination average distance and trip length frequency distribution (household travel survey)
- Tours by origin and primary destination district (household travel survey)

Tour Time-of-Day Choice:

- Share of tours by departure, arrival, and duration half-hour period and purpose (household survey)

Tour Mode Choice:

- Tours by tour purpose, mode and auto sufficiency (household survey and transit on-board survey)
- Transit tours by mode of and origin/destination district (transit on-board surveys and/or other data)

Intermediate Stop Frequency:

- Share of tours by number of outbound and inbound intermediate stops and tour purpose (household survey)
- Number of trips per tour by tour purpose (household survey)

Intermediate Stop Location Choice:

- Intermediate stops by tour purpose and out-of-direction distance (household survey)

Trip Mode Choice:

- Trips by tour purpose, tour mode, and trip mode (household survey and transit on-board survey)

- Transit trips by access mode and trip distance (household survey and transit on-board survey)
- Transit trips by district and line-haul mode

Model Validation

The model will be validated to traffic counts and transit boardings. Model validation is an iterative task with model calibration; often assignment reveals issues with underlying travel model calibration or specification that must be addressed in order to improve assignment results. An example would be development of district-level adjustment factors in destination choice, to improve screenline volumes, though we prefer not to include district constants in destination choice models since they can affect model sensitivities. This is a matter of judgement which will be based on model goodness-of-fit. As noted above, we will accept higher levels of error in assignment validation in the Phase I models; we will focus validation efforts on Phase II models. As part of this task, we will assess the need for performing sequential assignment of non-HOV3 and HOV3+ vehicles.

The following are some examples of network validation summaries.

Auto assignment:

- Vehicles by facility type, area type, and district or county (traffic counts)
- Percent route mean square error by facility type, area type, and district or county (traffic counts)
- Screenline, bridge, and key location summaries (traffic counts)
- Travel time data compared to congested speeds (subject to data availability)
- Comparison of trip tables to Location Based Services (LBS) data (subject to data availability)

Transit assignment:

- Transit boardings by route and operator (on-board survey or transit operator passenger counts)
- Transit passengers by transit screenline (transit ridership data)
- Metrorail boardings and alightings by station pair (transit ridership data)
- Total transfers by operator (transit on-board survey data)

Deliverables

- Model Calibration and Validation Plan, Gen3 Model, Phase I
- Updated Gen3 Phase I Model with Calibrated Parameters
- MWCOG Gen3 Model Development Technical Report or Memorandum: Phase I Model Calibration and Validation Results

Task 6: Sensitivity Testing

Sensitivity testing is a fundamental component of the development of a new modeling system. Although the activity-based model being deployed for MWCOG has been applied in a number of other regions, the project team is interested in analyzing model sensitivities specific to the land-use data, network, and policies of interest to MWCOG. Sensitivity testing involves systematically varying one or more model inputs to understand how the model responds to those changes. It is fundamentally different from model calibration, which involves comparing goodness-of-fit of model output against observed data using a fixed set of inputs. The purpose of sensitivity testing is to understand model response to *changes* in inputs.

In order to be a useful metric of model sensitivity, sensitivity tests must be carefully formulated in order to isolate model responses to the inputs that vary. This requires limiting input changes to only those directly related to the sensitivity test and comparing the outputs against a baseline scenario. For this reason, typically only one variable is varied for each sensitivity test. If elasticity measurements are of interest, the model must be run several times, each with a different level of input variable. We will specify and test three sensitivity tests based on the following discussion in consultation with MWCOG. We assume that MWCOG staff will take responsibility for running one of these sensitivity tests internally. That will give MWCOG staff the opportunity to learn how to run and analyze output from the Gen3 Phase I Model.

Land-Use Scenarios

This is a broad group of sensitivity tests involving the analysis of changes in land-use on model outputs. Proposed land-use scenario tests for MWCOG include the following:

- ***Analysis of a new major employment center.*** This sensitivity test would entail locating a new job center somewhere in the ***MWCOG*** region and adding a significant number of jobs to those MAZs or TAZs. Assuming the same synthetic population (number of workers) as the baseline scenario, it may be reasonable to reduce employment in other TAZs/MAZs. Varying the number of total jobs would provide a range of sensitivities to changes employment. It would also allow the team to understand the implications of different levels of shadow pricing.
- ***Analysis of changes in parking cost.*** Parking costs can be systematically varied for a selected group of TAZs/MAZs. This test is relatively simple to analyze as it involves only changes to the land-use input file. Typically, mode shares to the selected zones would be measured against the baseline scenario as the key output of interest, along with transit boardings/alightings on routes serving the changed area.
- ***Analysis of a new mixed-use or transit-oriented development.*** Mixed-use and transit-oriented developments would typically involve a combination of several of the land-use changes described above for a selected set of zones. This can include changing the number and type of households (e.g. synthetic population) and employment, changes in parking cost, as well as introducing network changes such as transit and/or bicycle lanes. Therefore, this would be a more complicated scenario to model and isolate model responses.

Network Scenarios

Network scenarios involve changes to the road, transit, and/or non-motorized network.

- ***Major new transit system expansion.*** This could include one or more of the following: new routes, reduced headway on existing routes, reduced fares, and/or longer service hours. All of these, with the possible exception of specification of new routes, would be relatively easy to code and analyze. Key outputs of interest include impacts on auto ownership, mode share, and transit boardings.
- ***Road capacity expansion.*** An analysis of base-year level-of-service would inform the selection of a congested corridor. Capacity would be added to the corridor to reduce congestion. Key metrics would be travel time changes in the corridor, Vehicle Miles of Travel (VMT), Vehicle Hours of Travel (VHT), and Vehicle Hours of Delay (VHD).

MWCOG staff have conducted various sensitivity tests in the past, such as added capacity on a Potomac River bridge.[35]

Demographic Scenarios

Demographic scenarios involve systematically changing the controls to population synthesis, to change the characteristics of the synthetic population (with the same number of total households by zone as the baseline scenario). Examples of demographic scenarios include:

- ***Aging households.*** Age distributions used in the synthetic population controls will be shifted to generate an older age distribution. The household size, worker per household, and workers by occupation distributions may also need to shift to reflect a consistent number of workers and average household size with the older population. A new synthetic population will be generated using these controls and model results will be compared to baseline results. All model results would be analyzed including tour and stop generation, destination, time-of-day, mode choice, VMT, VHT, and VHD.
- ***Income shifts.*** The income distributions used in the population synthesizer will be shifted to generate a much lower and/or higher income population. It may also be useful to change the worker per household and workers by occupation distributions to ensure consistency across population synthesis inputs. The same sorts of outputs described above will be analyzed.

Deliverables

- MWCOG Gen3 Model Development Technical Report or Memorandum: Phase I Sensitivity Testing Plan

---

[35] See, for example, Ron Milone, “Ver. 2.5 Travel Model Development and Evaluation” (July 2018 meeting of the COG/TPB Travel Forecasting Subcommittee, held at the Metropolitan Washington Council of Governments, Washington, D.C., July 20, 2018) or; Ronald Milone and Mark S. Moran, “TPB Version 2.3 Travel Model on the 3,722-TAZ Area System: Status Report and Sensitivity Tests” (July 2011 meeting of the COG/TPB Travel Forecasting Subcommittee, held at the Metropolitan Washington Council of Governments, Washington, D.C., July 22, 2011).

- MWCOG Gen3 Model Development Technical Report or Memorandum: Phase I Sensitivity Testing Results
- Sensitivity Test 1 Model Inputs and Outputs
- Sensitivity Test 2 Model Inputs and Outputs
- Sensitivity Test 3 Model Inputs and Outputs

5.2 | GEN3 MODEL PHASE II

In the second phase of deployment, the RSG team will undertake the following activities:

Task 1: Model Estimation

The RSG team will prepare data and estimate the following models:

- Transit Pass Ownership Model
- Telecommute Frequency Model
- Auto Ownership Model
- Coordinated Daily Activity Pattern Model
- Mandatory Tour Frequency Model
- Non-Mandatory Tour Frequency Model
- Tour Time-of-Day Choice Models
- Intermediate Stop Frequency Model

All models will be estimated using the estimation functionality in ActivitySim. The RSG team will ensure that model parameters are reasonable. The RSG team will document model estimation results.

Deliverables

- Model estimation datasets
- MWCOG Gen3 Model Development Technical Report or Memorandum: Phase II Model Estimation Results

Task 2: Phase II ActivitySim Deployment

The RSG Team will implement updated models estimated in Phase II Task 1. The RSG team will implement the following new models in ActivitySim. Many of these components will also be implemented in the SEMCOG models, so we expect leveraging those investments for MWCOG.

a. Transit pass ownership model
b. Telecommute frequency model

c. Auto ownership model extension for AVs
d. AV tour availability model
e. AV tour and trip mode parameters
f. AV trip table parameters

The RSG team will also implement transit capacity restraint if agreed to after testing in Phase I.

As with all new models, there will be areas of improvement in the initial implementation that will not be evident until the model system is up and running. In cooperation with the project team, select usability, input checking, and performance improvements will be identified and implemented as the project can afford.

Deliverables

- Gen3 Phase II model implementation

Task 3: Model Calibration and Validation

The RSG team will calibrate and validate the Gen 3 Phase II MWCOG model system. See above Phase I Task 5 for a description of model calibration and validation tasks. For Phase II, we will tighten calibration and validation to meet or exceed validation of current MWCOG Gen2/Ver. 2.3 Travel Model. Based on the findings from the Phase 1 calibration/validation work, there may be a need to develop a calibration/validation plan for Phase 2, so that work can be allocated across the development team. Ideally, such a plan would specify calibration/validation benchmarks.

Deliverables

- Updated Gen3 Phase II Model Implementation
- MWCOG Gen3 Model Development Technical Memorandum: Phase II Final Model Calibration and Validation Results

Task 4: Sensitivity Testing

The RSG team will re-run the sensitivity tests agreed to in Phase I. MWCOG will take a lead role for running and documenting one of the tests. The RSG team will make modifications to the Phase II model to correct any issues discovered during the sensitivity test runs and ensure reasonable elasticities to inputs. In addition to the three sensitivity tests proposed below, the consultant and COG staff could develop a list of additional sensitivity tests which could be conducted by COG staff.

Deliverables

- MWCOG Gen3 Model Development Technical Memorandum: Phase II Sensitivity Testing Results
- Phase II Sensitivity Test 1 Model Inputs and Outputs
- Phase II Sensitivity Test 2 Model Inputs and Outputs

- Phase II Sensitivity Test 3 Model Inputs and Outputs

Task 5: Documentation and Training

The RSG team will prepare a final model report and a model user guide. The final report will assemble all the technical memorandum previously developed into one document. The RSG team will prepare a user guide that describes model setup, running the model, inputs, and outputs. The RSG team will prepare and deliver a 1.5 day in-person or online model use training.

Deliverables

- MWCOG Gen3 Model Development Final Report
- MWCOG Gen3 Model User's Guide
- MWCOG Gen3 Model Use Training Materials
- MWCOG Gen3 Model Use Training Delivery

Resource requirements

We provide a draft budget and schedule in a spreadsheet incorporated herein by reference. For the purposes of budgeting the model development effort, we assume that the RSG team will take full responsibility for all tasks with the exception of the model sensitivity testing (one scenario) that is identified as the responsibility of COG staff in each model development phase. We also identify the following model features as "optional" components that are not within the current budget constraints:

- Internal-external model for residents. This model would explicitly model internal resident travel to external stations. The components include IE travel generation, destination choice, and mode choice. Trip lists output by the model would be integrated into the model system and replace the existing IE trip tables. The model can be developed from existing data.
- Airport ground access model. This model would explicitly represent travel to the region's three airports (Ronald Reagan Washington National Airport, Washington Dulles International Airport, and Baltimore/Washington International Thurgood Marshall Airport). It will take the form described above. The survey data currently exists to build this model. It is possible that the model could be funded either wholly or in part by an ActivitySim consortium member, in which case the implementation cost could be reduced or mostly eliminated.
- Overnight visitor model. This model would explicitly represent overnight visitor travel in the MWCOG region and is described above. It would require collection of several datasets (hotel/motel room inventory, information on total visitors by purpose, length of stay, and travel patterns). This model may also be funded by an ActivitySim consortium member, in which case the implementation cost could be reduced or mostly eliminated.

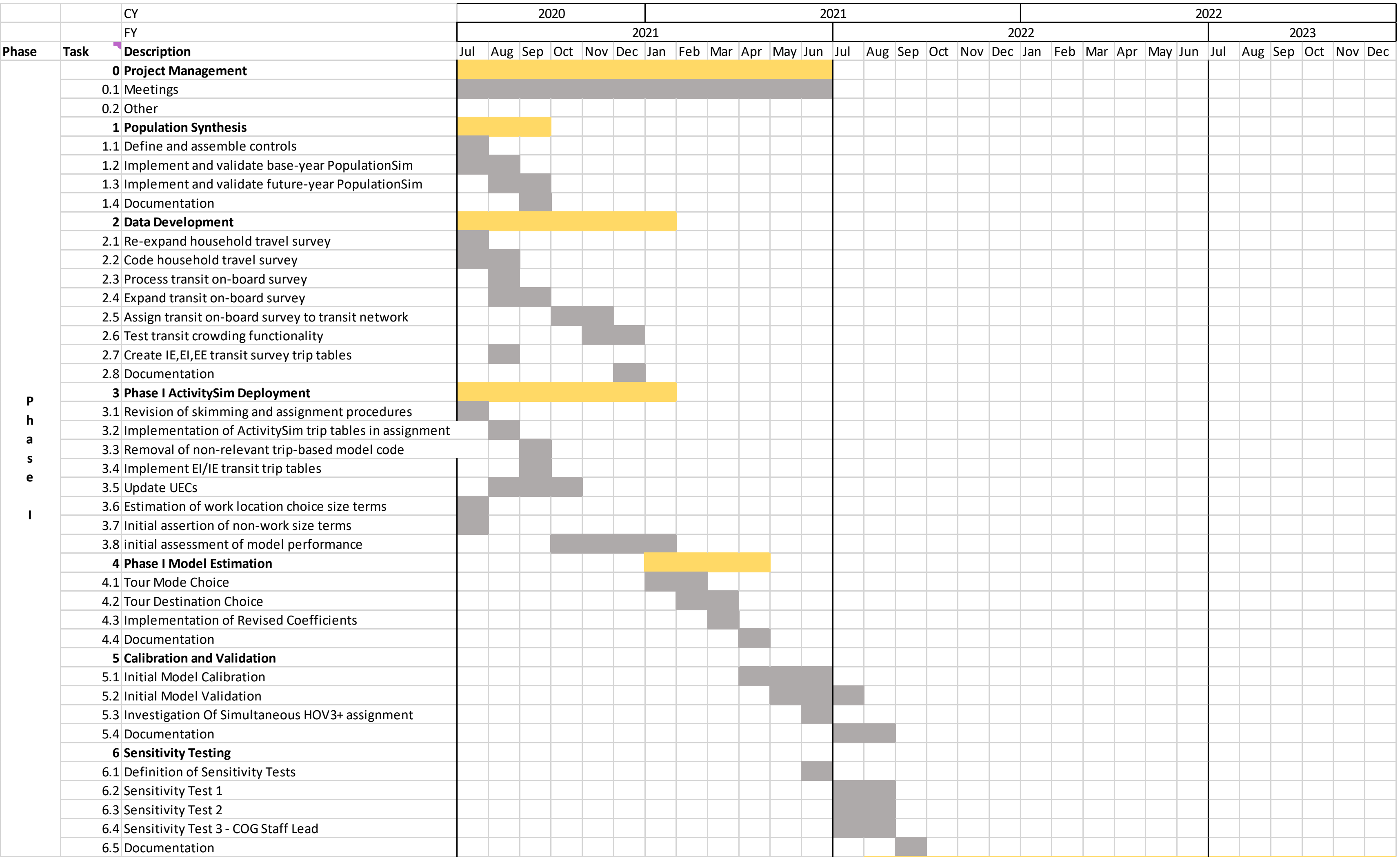
- University residential location choice. This model would explicitly model the residential location of students housed in non-group quarters residences, for each major university. The model would be used as an additional control to the synthetic population, to improve the location choice of university students. It would require collection of data on university student housing location and estimation of a location choice model.

However, we identify the following potential tasks that COG staff may wish to take on, which would free up model development resources to devote to one or more of the optional model components listed above. Note that if COG staff takes on responsibility for some of these tasks, we would need to ensure that Baseline Mobility continues to receive the required share of project budget to meet Disadvantaged Business Enterprise (DBE) requirements.

- Population Synthesis Deployment. Several of our client agencies have taken on the development of a synthetic population. The RSG team would retain hours for oversight of this task but the work would be completed by MWCOG staff. This would save approximately $20k of RSG team budget.
- Model Estimation. Model estimation is one of the more challenging aspects of model development, so we suggest only taking on estimation of simpler models such as auto ownership and/or mandatory tour frequency. There could be savings of $20k or more depending on which models are estimated by MWCOG staff.
- Model Calibration: MWCOG staff could take responsibility for initial calibration of mode choice models, for example. RSG would provide oversight but COG staff would be responsible for running the model, importing results into calibration spreadsheets, adjusting constants, re-running models, etc. Note that there is no auto-calibrate functionality currently in ActivitySim. This could result in savings of $20-$30k depending on the extent of task sharing.
- Sensitivity Testing. MWCOG staff is currently responsible for running one of the three sensitivity tests. COG could take on an additional task in Phase I and Phase II for a savings of approximately $8k.

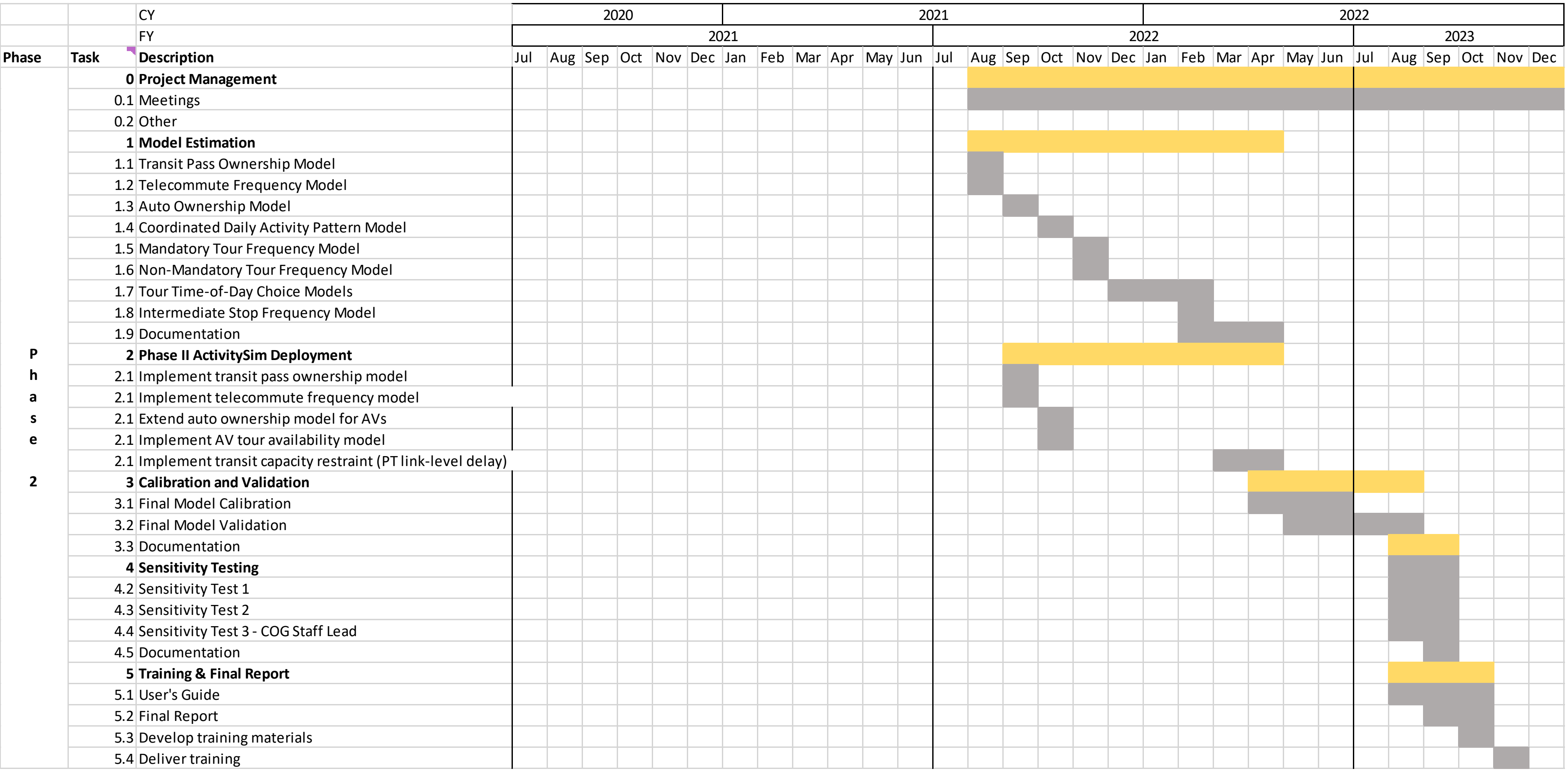
Model Development Schedule

The model development schedule is shown in Figure 12 (Phase I) and Figure 13 (Phase II). Phase I model development will begin in July 2020 and expected to be completed by September 2021. Phase II model development is expected to begin in August 2021 and be completed by end of December 2022. The second half of 2022 is planned for sensitivity testing of the final Gen3 Model, training, and documentation.



		CY	2020						2021												2022											
		FY	2021												2022												2023					
Phase	**Task**	**Description**	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Phase I	**0**	**Project Management**	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨																		
	0.1	Meetings	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨																		
	0.2	Other																														
	1	**Population Synthesis**	🟨	🟨	🟨																											
	1.1	Define and assemble controls	▨																													
	1.2	Implement and validate base-year PopulationSim	▨	▨																												
	1.3	Implement and validate future-year PopulationSim		▨	▨																											
	1.4	Documentation			▨																											
	2	**Data Development**	🟨	🟨	🟨	🟨	🟨	🟨	🟨																							
	2.1	Re-expand household travel survey	▨																													
	2.2	Code household travel survey	▨	▨																												
	2.3	Process transit on-board survey		▨																												
	2.4	Expand transit on-board survey		▨	▨																											
	2.5	Assign transit on-board survey to transit network				▨	▨																									
	2.6	Test transit crowding functionality					▨	▨																								
	2.7	Create IE,EI,EE transit survey trip tables		▨																												
	2.8	Documentation						▨																								
	3	**Phase I ActivitySim Deployment**	🟨	🟨	🟨	🟨	🟨	🟨	🟨																							
	3.1	Revision of skimming and assignment procedures	▨																													
	3.2	Implementation of ActivitySim trip tables in assignment		▨																												
	3.3	Removal of non-relevant trip-based model code			▨																											
	3.4	Implement EI/IE transit trip tables			▨																											
	3.5	Update UECs		▨	▨	▨																										
	3.6	Estimation of work location choice size terms	▨																													
	3.7	Initial assertion of non-work size terms	▨																													
	3.8	initial assessment of model performance				▨	▨	▨	▨																							
	4	**Phase I Model Estimation**							🟨	🟨	🟨	🟨																				
	4.1	Tour Mode Choice							▨	▨																						
	4.2	Tour Destination Choice								▨	▨																					
	4.3	Implementation of Revised Coefficients									▨																					
	4.4	Documentation										▨																				
	5	**Calibration and Validation**																														
	5.1	Initial Model Calibration										▨	▨	▨																		
	5.2	Initial Model Validation											▨	▨	▨																	
	5.3	Investigation Of Simultaneous HOV3+ assignment												▨																		
	5.4	Documentation													▨	▨																
	6	**Sensitivity Testing**																														
	6.1	Definition of Sensitivity Tests												▨																		
	6.2	Sensitivity Test 1													▨	▨																
	6.3	Sensitivity Test 2													▨	▨																
	6.4	Sensitivity Test 3 - COG Staff Lead													▨	▨																
	6.5	Documentation															▨															

FIGURE 12: PHASE I MODEL DEVELOPMENT SCHEDULE



		CY	2020						2021												2022											
		FY	2021												2022												2023					
Phase	**Task**	**Description**	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
	0	**Project Management**														🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨
	0.1	Meetings														▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨
	0.2	Other																														
	1	**Model Estimation**														🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨								
	1.1	Transit Pass Ownership Model														▨																
	1.2	Telecommute Frequency Model														▨																
	1.3	Auto Ownership Model															▨															
	1.4	Coordinated Daily Activity Pattern Model																▨														
	1.5	Mandatory Tour Frequency Model																	▨													
	1.6	Non-Mandatory Tour Frequency Model																	▨													
	1.7	Tour Time-of-Day Choice Models																		▨	▨	▨										
	1.8	Intermediate Stop Frequency Model																				▨										
	1.9	Documentation																				▨	▨	▨								
P	**2**	**Phase II ActivitySim Deployment**															🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨								
h	2.1	Implement transit pass ownership model															▨															
a	2.1	Implement telecommute frequency model															▨															
s	2.1	Extend auto ownership model for AVs																▨														
e	2.1	Implement AV tour availability model																▨														
	2.1	Implement transit capacity restraint (PT link-level delay)																					▨	▨								
2	**3**	**Calibration and Validation**																						🟨	🟨	🟨	🟨	🟨				
	3.1	Final Model Calibration																						▨	▨	▨						
	3.2	Final Model Validation																							▨	▨	▨	▨				
	3.3	Documentation																										🟨	🟨			
	4	**Sensitivity Testing**																										▨	▨			
	4.2	Sensitivity Test 1																										▨	▨			
	4.3	Sensitivity Test 2																										▨	▨			
	4.4	Sensitivity Test 3 - COG Staff Lead																										▨	▨			
	4.5	Documentation																											▨			
	5	**Training & Final Report**																										🟨	🟨	🟨		
	5.1	User's Guide																										▨	▨	▨		
	5.2	Final Report																											▨	▨		
	5.3	Develop training materials																												▨		
	5.4	Deliver training																													▨	

FIGURE 13: PHASE II MODEL DEVELOPMENT SCHEDULE

6.0 CURRENT AND FUTURE DATA

6.1 | MODEL INPUTS

The activity-based model inputs can be grouped into several categories:

1. **Model parameters and other settings.** These include coefficients of statistically estimated model such as logit models, or empirical distributions that are created from survey data. These are typically static across scenarios (unless the objective of the scenario is to specifically test for example increased sensitivity to particular model parameter). Most parameters will be estimated using the 2017-18 Regional Travel Survey and available on-board transit survey data.

2. **Synthetic Population.** The synthetic population consists of a household file and a person file which represent the population for the scenario and year to be modeled. These files will be created using PopulationSim. Inputs to the population synthesizer include census data and land-use data. We recommend using household size, household income, workers per household, and population by age category as controls. Final controls for PopulationSim will be determined as part of model development. Formats for synthetic households and person tables are shown in the tables below. ActivitySim's table annotation functionality will be used to calculate derived data columns that are required by the models but not included in the input household or person tables. The files must be in either Python/Pandas HDF5 format or CSV format. The final synthetic population will contain both residential and group quarters population. The group quarters population is generated using the same approach and tools as the residential population. There is much less information available to control the group quarter population generation. The population synthesizer controls for group quarters and residential populations are independent of each other. Therefore, the population synthesizer for group quarters can be run independently to produce a household file and person file in the same format as the residential synthetic population. The residential and group quarters files are combined to generate the final synthetic population files.

   It should be noted that institutionalized group quarters do not need to be generated since their travel is not modeled by the activity-based model. Group quarters are treated as single person households in the activity-based model framework. The group quarters person can either be a university student, active military person or other non-institutional group quarter type.

3. **Zonal Data.** TAZ-level data includes total households, total persons, employment by type, school enrollment, total zonal area, and other relevant information. The files must be in either Python/Pandas HDF5 format or CSV format. Zonal data is described in greater detail below.

4. **Network Data.** Network data is used to create skims and is described in more detail above under Trip Modes and below. Our review of network attributes suggests that necessary attributes are available for base-year modeling.

TABLE 11: SYNTHETIC HOUSEHOLD TABLE

Field	Description	Scale
HHID	Unique household ID number	Integer, any value
TAZ	Transportation analysis zone of home location	Integer
TYPE	Type of unit	Integer, 1 – housing unit; 3 – non-institutional groups quarters
HINCP	Household income	Numeric, dollars in ACS year
ADJINC	2015 adjustment factor for dollar amounts	Numeric
NP	Number of persons in the household	Integer, 1 and up
HHT	Household/family type	Integer, 1 - family household: married-couple; 2 - family household: male householder, no wife present; 3 - family household: female householder, no husband present; 4 - non-family household: male householder living alone; 5 - non-family household: male householder, not living alone; 6 - non-family household female householder, living alone; 7 - non-family household: female householder, not living alone
VEH	Number of vehicles available	Integer, 1 - one vehicle; 2 - two vehicles; 3 - three vehicles; 4 - four vehicles; 5 - five vehicles; 6 - six or more vehicles

TABLE 12: SYNTHETIC PERSON TABLE

Field	Description	Scale
HHID	Unique household ID number	Integer, any value
PERID	Unique person ID number	Integer, any value
AGEP	Person's age in years	Integer, 0 and up
SEX	Gender	Integer, 1 - male; 2 - female
ESR	Employment status recode	Integer, 0 . N/A (less than 16 years old) 1 . Civilian employed, at work 2 . Civilian employed, with a job but not at work 3 . Unemployed 4 . Armed forces, at work 5 . Armed forces, with a job but not at work 6 . Not in labor force
WKHP	Usual hours worked per week past 12 months	Integer, 0 .N/A (less than 16 years old/did not work during the past .12 months) 01..98 .1 to 98 usual hours 99 .99 or more usual hours
WKW	Weeks worked during past 12 months	Integer, 0 .N/A (less than 16 years old/did not work during the past 12 .months) 1 .50 to 52 weeks worked during past 12 months 2 .48 to 49 weeks worked during past 12 months 3 .40 to 47 weeks worked during past 12 months 4 .27 to 39 weeks worked during past 12 months 5 .14 to 26 weeks worked during past 12 months 6 .less than 14 weeks worked during past 12 months
SCHG	Grade level attending	Integer, 0 .N/A (not attending school)

		1 .Nursery school/preschool 2 .Kindergarten 3 .Grade 1 4 .Grade 2 5 .Grade 3 6 .Grade 4 7 .Grade 5 8 .Grade 6 9 .Grade 7 10 .Grade 8 11 .Grade 9 12 .Grade 10 13 .Grade 11 14 .Grade 12 15 .College undergraduate years (freshman to senior) 16 .Graduate or professional school beyond a bachelor's degree
MIL	Military service	Integer, b .N/A (less than 17 years old) 1 .Now on active duty 2 .On active duty in the past, but not now 3 .Only on active duty for training in Reserves/National Guard 4 .Never served in the military

Coding Person Type

As mentioned above, person type is an important variable in ActivitySim; in some cases, it is used to stratify models; for example, there is a separate individual non-mandatory tour generation model for each person type category. In other cases, it is an explanatory variable. For example, the coordinated daily activity pattern model represents interactions between household members as a function of their person type. The person type coding logic is a three-step process. First, an employment category is assigned to each person. Next, a student category is assigned to each person. Finally, each person is assigned one of the eight person types. Person type is coded based on the following fields: AGEP, ESR, WKHP, WKW, and SCHG.

Employment category is coded as follows:

1. If AGEP<16, employment category is "Under 16"

2. Else if ESR={3,6}, employment category is "Not employed"
3. Else if WKHP>35 and WKW>={1,2,3,4} employment category is "Employed Full-Time"
4. Else the employment category is "Employed Part-Time"

Student category is coded as follows:

1. If AGEP<16, student category is "High school or less"
2. Else if employment category is "Employed Full-time" or SCHG=0, student category is "Not a student"
3. Else if SCHG>=6 or AGEP>19, student category is "College or higher"
4. Else student category is "High school or less"

Then person type is coded as follows:

1. If employment category is "Employed Full Time", person type is "Full-time worker" (1)
2. Else if student category is "Not a student"
    a. If employment status is "Part-time worker" person type is "Part-time worker" (2)
    b. Else if AGEP>=65, person type is "Non-working senior" (5)
    c. Else if AGEP<6 person type is "Pre-School" (8)
    d. Else person type is "Non-worker and Non-student" (4)
3. Else if student category is "College or higher" person type is "College student" (3)
4. Else if AGEP<6 person type is "Pre-school" (8)
5. Else if AGEP>=16 person type is "Driving age student" (6)
6. Else person type is "Non-driving student" (7)

There may be other household and person level variables used in ActivitySim that are combinations of the above household and person variables, such as single-parent indicators, or auto sufficiency (defined as a combination of autos compared to drivers or persons age 16+). These variables vary between implementations and are best described in model estimation documentation after the models are finalized.

Zonal Data

Zonal data is a key input to ActivitySim, since zonal data provides information on the opportunity to engage in out-of-home activities in each zone. We list required TAZ data fields in Table 13. We recommend that MWCOG staff collect and geocode base and future year school enrollment data. The data is available from the Institute of Education Sciences National Center for Education Statistics website at https://nces.ed.gov/programs/edge/Home. College and university enrollment data can be found from a web search for each school. For college campuses that span multiple zones, enrollment should be apportioned across TAZs based on square footage of active space, classroom space, and/or student parking.

Parking cost and terminal time are estimated from an existing model in Ver. 2.3. We recommend that MWCOG staff collect more up-to-date parking cost information for the Gen3 Model and re-estimate the parking cost model. The re-estimated parking costs can be used for Phase II model deployment. We suggest that COG collect disaggregate parking cost data, including the XY location of each surface and structured lot and pricing policy of each site so that parking costs can be re-calculated based on the raw data. Note that parking costs should be specified as an average cost across all zones whose visitors are likely to pay for parking, rather than the cost specific to the actual destination zone, and should be the true cost of parking rather than an average cost that considers non-payers.

We also recommend collection of "active" park space and open space total acres, to be used in destination choice model estimation. Such data is very helpful for discretionary activity size terms, since parks, playgrounds, cemeteries, and other types of parks and public areas often attract trips but have little to no on-site employment associated with them. San Diego Association of Governments includes the following in active park space: cemeteries, golf courses (public and private) campgrounds and retreat centers, rifle and archery ranges, recreation areas and centers including basketball courts, baseball diamonds, soccer fields, and neighborhood parks. Open spaces include large parks and public areas such as nature trails, nature preserves, and wilderness areas. If one of the special market models is developed, there may be other zonal data to be collected. For example, if the visitor model is developed, visitor attractions - museums, national monuments, etc. - are important size term variables. Square footage of classroom space is a useful variable for a university model size term. These data will be specified once a determination has been made on special market model development.

TABLE 13: TAZ DATA TABLE

Field	Description
TAZ	TAZ number.
HH	Households
HHPOP	Household population
GQPOP	Group quarters population
TOTPOP	Total population
JURCODE	Jurisdiction Code (0-23)
LANDAREA	Gross land area (square miles)
ADISTTOX	Airline distance to the nearest external station (whole miles)
TAZXCRD	TAZ X-coordinate (NAD83, whole feet)
TAZYCRD	TAZ Y-coordinate (NAD83, whole feet)
HHINCIDX	Ratio of zonal HH median income to regional median HH income

Field	Description
TOTEMP	Total employment
RETEMP	Retail employment
OFFEMP	Office employment
OTHEMP	Other employment
INDEMP	Industrial employment
k_8_enroll	K-8 school enrollment
9_12_enroll	9-12 school enrollment
univ_enroll	College/university enrollment
WrkPrkCost	All day parking cost applied to commute trips; note that these costs should not include workers whose parking is reimbursed.
NonWrkPrk	Short-term parking cost applied to non-work trips
TERMINAL	Terminal time in minutes, used for parking access/egress time for mode choice.
ACTIVEPARK	Acres of active park space
OPENSPACE	Acres of open recreational space

6.2 | MODEL OUTPUTS

ActivitySim model outputs include the following tables written to Python/Pandas HDF5 format or CSV format. The trip data files are used to create trip tables by period and occupancy or transit mode for assignment.

1. **Household data.** This table contains outputs from the travel model at a household level including auto ownership and coordinated household daily activity pattern type.
2. **Person data.** This table contains outputs from the travel model at a person level including daily activity pattern type, mandatory tour choice, and non-mandatory tour choice.
3. **Tour data.** This table contains individual tours (one record per individual tour) and fully joint tours (one record per joint tour) with departure and arrival period, tour purpose, tour mode, and other relevant data.
4. **Trip data.** This table contains individual trips (one record per individual trip) and fully joint trips (one record per joint trip) with time period, origin purpose, destination purpose, trip mode, and other relevant data.

.

6.3 | NETWORK DATA

Network data includes highway and transit network and is used for creating skims. Table 5 described the skims required for the current ActivitySim design. Current ActivitySim design requires the skims be copied into a single Open Matrix Format (OMX) file, although this constraint is expected to be lifted by the time the MWCOG model is developed. The base highway network of the Ver 2.3 Model has all the necessary variables available for generating the required highway skims. These variables include distance, facility type codes, lanes by period, toll value fields and period-wise, user-market enable/disable codes (vehicle occupancy, transit, trucks, etc., also known as LIMIT codes). The Ver. 2.3 Model's base highway network will be used for Gen3 Model development.

The transit network consists of the highway network (used by buses), transit-only network (stations and rail links) and transit service details. The transit network for the Ver. 2.3 Model has all the necessary elements required for the current effort. The link-level transit crowding in Cube PT will be tested for the current year using observed transit trip tables from on-board survey data and for future-year scenarios using the Phase I future-year model trip tables and networks. The Cube PT link-level crowding testing will require the following additional inputs:

- Load distribution factor: the percentage of seats that are occupied before the transit travelers begin to start standing.
- Seats: number of seats in the vehicle.
- Total vehicle capacity.
- Crowding Curve: crowding factor distribution segmented by vehicle type and user class.

A key enhancement planned under the Gen4 Model development is the incorporation of an all-streets network. The all-streets network will be used for generating skims for the walk and bike modes. The representation of bicycle facilities in the all-streets network can further improve the model's ability to assess the benefits of investments in bicycle infrastructure. For this purpose, a link-level bicycle facility type code must be included in the all-streets network. The facility type code must represent various levels of bicycle infrastructure. Figure 14 presents the graphical depiction of various bicycle facility classes. The bike skimming process can be updated to use a generalized cost representing perceived bike time on each bicycle facility type class.

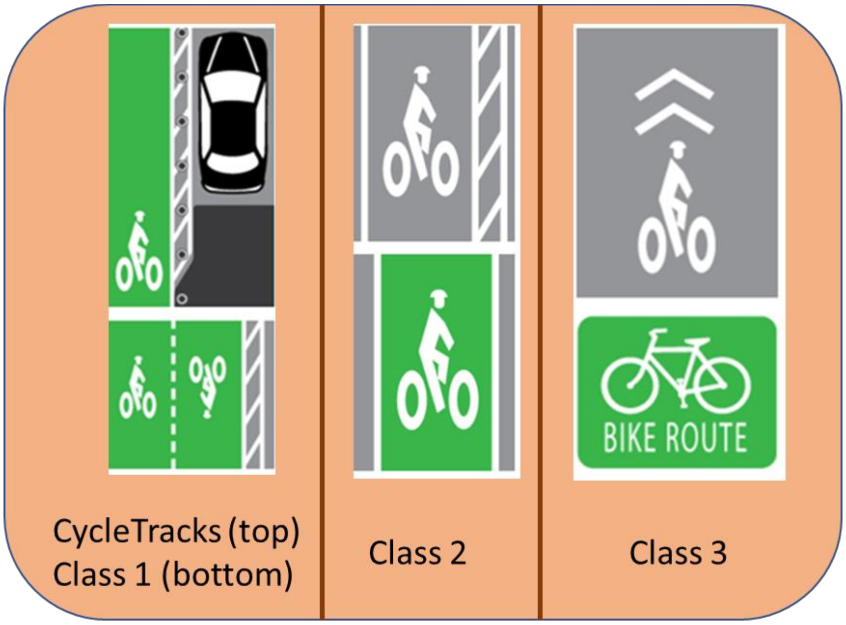


FIGURE 14: GRAPHICAL DEPICTION OF BICYCLE FACILITIES

Another potential enhancement under Gen4 Model development is ABM-DTA integration. A DTA model generally can be used to model signal synchronization, lane-based effects, various intelligent transportation system (ITS) elements, and ramp metering. Thus, DTA models typically require more inputs on the network side. There are many DTA software packages that differ in their modeling capability and input requirements, so we recommend selecting a DTA modeling package before collecting data and structuring a data collection program based on model needs. Below is a list of potential network attributes and network enhancements needed for a DTA application, but the actual list would depend on the choice of the DTA software and the level of analysis:

- Acceleration-deceleration lanes, reserved lanes and other auxiliary lanes.
- Lane movements and lane connectivity.
- Intersection geometry including turn bays.
- Traffic signal locations and timings. The representation of signals varies by DTA software and may need parameters such as green time per phase.
- Standard uncontrolled intersections – all-way stop controlled, two-way stop controlled, roundabouts, freeway merges and yield signs. Please note that only some DTA software provide for the specification of uncontrolled intersections.
- Ramp meter controls.
- Freeway lane controls signs and other ITS elements.

6.4 | SURVEY DATA FOR ESTIMATION AND CALIBRATION

Regional Travel Survey

The TPB staff has been conducting a household travel survey (HTS) approximately every ten years since 1968. The latest completed and available HTS was conducted in 2007-08[36] and the dataset was used for the last major calibration of the Ver. 2.3 Model. In 2017-18, TPB staff collected an HTS called the 2017-18 Regional Travel Survey (RTS).[37] The 2017-18 RTS dataset will be the primary data source for estimation and calibration of the ActivitySim sub-models for the Gen3 Model under the current effort. The 2017-18 RTS was collected from Fall of 2017 to the end of 2018. The dataset is divided into household, person, vehicle and trip files. It has around 16,000 household records and around 122,000 trip records. The MWCOG staff is currently cleaning and editing the 2017-18 RTS dataset. As described earlier, the ActivitySim model system works with data structures such as stops, linked trips, tours and joint tours. The 2017-18 RTS dataset will be processed using Oregon Department of Transportation 's (ODOT) Python-based Survey Processing Application (SPA). The SPA takes the survey inputs in the required format and produces linked trips, tours and joint tours files in the ActivitySim format. This processed RTS data will be used for creating estimation datasets and calibration targets for various ActivitySim sub-models.

Transit On-board Surveys

Most regional travel surveys underestimate transit ridership and underrepresents some transit markets. Therefore, the regional travel survey is typically supplemented with a transit on-board survey. In 2008, MWCOG had conducted a regional bus survey that included all bus operators.[38] Similarly, in 2008, WMATA and the Maryland Transit Administration (MTA) conducted a Metrorail survey.[39] These and other transit on-board surveys conducted between 2005 and 2008 were used for calibration of the Ver. 2.3 Model. A more recent dataset will be required for the current effort. A systemwide transit on-board survey covering all transit operators in the modeling region is desirable but would be practically challenging to conduct. MWCOG is currently reaching out to local transit agencies that have conducted their own transit on-board surveys within the last few years. 24 agencies have been contacted in this regard. A

---

[36] National Capital Region Transportation Planning Board, Metropolitan Washington Council of Governments, “2007/2008 TPB Household Travel Survey: Technical Documentation,” Draft report (Washington, D.C.: National Capital Region Transportation Planning Board, Metropolitan Washington Council of Governments, August 27, 2010), http://www.mwcog.org/uploads/committee-documents/Zl5YWV5W20100903131244.pdf.

[37] Kenneth Joh, “2017-2018 Regional Travel Survey Briefing: Demographic Changes and Typical Commute” (April 2020 meeting of the Technical Committee of the National Capital Region Transportation Planning Board, held at the Metropolitan Washington Council of Governments, Washington, D.C., April 3, 2020), https://www.mwcog.org/events/2020/4/3/tpb-technical-committee/.

[38] NuStats, “2008 Regional Bus Survey: Draft Report” (Washington, D.C.: Metropolitan Washington Council of Governments, June 2009).

[39] WB&A Market Research, “2008 Metrorail Passenger Survey,” Final Report (Maryland Transit Administration and Washington Metropolitan Area Transit Authority, July 10, 2009).

final systemwide transit on-board survey dataset will be compiled from the individual agency surveys.

The transit operators in the COG/TPB modeled region report route-wise average weekday transit ridership to the TPB for each month in a fiscal year (July to June). The final transit on-board survey will be expanded to replicate observed transit ridership by operator, route and transfer rates. The consolidated transit on-board survey data will be coupled with the RTS dataset to develop tour and trip mode choice targets for ActivitySim calibration. As mentioned earlier, trip tables will be developed from the transit on-board survey and assigned to the transit network to determine the quality of the transit on-board survey data, explore potential transit network coding issues, and calibrate PT transit assignment parameters. The transit on-board survey dataset will also be used to develop external-internal transit trip tables.

Other Data Collection

The COG/TPB modeled region has 7 major universities (see Table 1 for a preliminary list). An explicit university student travel model can be potentially developed in Phase II or Gen4 depending on data availability and project resources. The current representation of university students in ActivitySim requires Census counts of group quarters (GQ) households by type (including university students). This data can be obtained from American Community Survey (ACS) Public Use Microdata Sample (PUMS). In addition, the current ActivitySim design also requires base-year university student enrollment data. If the campus spans multiple zones, building square footage or student attraction estimate by zone is also needed. The next logical enhancement in modeling university student would be to add non-GQ university student segment in population synthesis. This would require residential location data for off-campus university students. Many universities collect residential location and travel data from their students. Universities can be contacted to obtain student residential location and travel data. In case of privacy concerns, residential data can be obtained at an aggregate level (Zone, Tract or ZIP Code). In order to test university specific parking policies, a university parking location choice model would be needed. A parking location choice model would require data on on-campus parking location and utilization. A more explicit treatment of university students can be considered if disaggregate data on student travel is available. This will potentially be contemplated under the Gen4 Model development and will require a student travel survey. TPB can consider adding a university student survey sub-sample for the next regional travel survey. The university student survey must include sample for all major university students and collect demographic information, residential location data and travel-related details.

The current ActivitySim model does not have a visitor model. An explicit overnight visitor travel model can be potentially developed in Phase II or Gen4 depending on data availability and project resources. For an aggregate treatment, visitor data can be collected from public sources, hotel/motel room inventory, and trip attraction data. A disaggregate treatment of visitors would require disaggregate visitor travel survey data. TPB should consider adding an overnight visitor travel survey component to the next airport survey. The survey must provide demographic information of the visitors and all travel-related details including information for all locations visited, trip purpose, travel mode, etc.

6.5 | TRAFFIC AND BOARDING COUNTS

COG/TPB staff gather traffic count data from the District Department of Transportation (DDOT) and the DOTs of Maryland, Virginia, and West Virginia. These traffic counts are collected from various permanent and short-term count stations across the TPB modeled region. The count datasets are packaged into a web-based application, called the Regional Transportation Data Clearinghouse[40] (RTDC) for easy access and data sharing across partner agencies. In addition to traffic counts, the TPB staff also compiles the observed VMT data at the jurisdiction level from the HPMS summaries reported by local state DOTs. The observed traffic counts and VMT data in the RTDC system will be used for the Gen3 Model validation. The model VMT will be compared against the observed VMT for the 22 jurisdictions in the TPB modeled region. The modeled volume will be compared against the ground counts and percent RMSE will be computed by facility type. It has been a challenge for COG/TPB staff to get many hourly traffic counts, since, as of 2012, there were only about 112 permanent count stations (hourly counts) in the region and none in the District of Columbia.[41]

A screenline analysis is an important part of the model validation effort. Screenline analysis compares estimated and observed daily link volumes across the pre-defined screenlines in the model network. For the current effort, the screenlines defined in the Ver. 2.3 Model network will be used for screenline analysis. The Ver. 2.3 Model had a total of 38 screenlines across major roads in each jurisdiction within the TPB modeled region. Only 57% of all links crossing regional screenlines were coded with a ground count in the Ver. 2.3 Model. Depending on data availability, more links can be coded with a ground count under the current effort. Transit boarding counts are obtained from the Washington Metropolitan Area Transit Authority (WMATA or Metro), which operates the Metrorail and Metrobus systems, and from other transit agencies, which operate commuter rail, bus, light rail, and bus rapid transit (BRT) services.

6.6 | BIG DATA

The project team recommends passively collected origin-destination (OD) data for three uses in MWCOG's travel model development:

- External travel patterns.
- Visitor travel patterns.
- Special attraction factors

---

[40] Metropolitan Washington Council of Governments. “TPB Regional Transportation Data Clearinghouse,” 2020. http://rtdc-mwcog.opendata.arcgis.com/

[41] Mary Martchouk to Ronald Milone, “2010 Hourly Counts,” Memorandum, April 13, 2012.

In 2014, MWCOG purchased and analyzed passive cellular phone data from AirSage and compared the data to model trip tables and traffic counts.[42] COG found that trip lengths and flows compared reasonably well to model results, though there were some differences between total trips by purpose, likely due to definitional differences. Spatially, COG staff found better matches to model data at higher levels of aggregation. Travel patterns for non-residents appeared reasonable, with clustering of activities around airports and tourist attractions. Finally, they found that the distribution of travel patterns at external stations was reasonable but the total magnitude of external travel was over-estimated.

Passive data from smartphone applications using location-based services (LBS) is now available. Although LBS data offers more variables and—on average—somewhat less precision than GPS data, it is far more precise than cellular data. Although it offers slightly lower sample penetration than cellular data, the difference is increasingly marginal, and it is equally capable of being expanded to represent all travel. Like cellular data, it also offers good device identifier persistence; however, unlike cellular data, it can be acquired at either the aggregate or disaggregate level. In these ways, LBS data offers a particularly attractive combination of attributes and may provide the best source of passive data for MWCOG model development activities listed below. RSG has used passive data to build models of visitor and external travel for the Michegan Statewide model[43] and the Charleston Area Transportation Study.[44]

External Travel Patterns

Passive OD data are recommended for use in determining external (IX, XI, XX) travel patterns for the MWCOG region and the refinement of the MWCOG travel model’s external travel modules. In recent years, fully passive data collection methods rely on processing datasets collected/produced from mobile devices or in-vehicle devices. Compared to older methods like traditional surveys and semi-passive methods (e.g., LPC surveys and Bluetooth), these newer methods are more cost-effective. Moreover, unlike semi-passive methods, fully passive methods provide observations not only on external-external trips passing through the region, but also on the internal origins and destinations of inbound and outbound trips.

The development of external travel models is one of the most widespread uses of passive OD data in travel modeling and has become commonplace. Additionally, expanding passive external OD data is often a simpler process than other OD data. This is because Iterative Proportional Fitting to traffic counts at external stations is often (but not always) sufficient to properly expand the data.

To support the full range of external travel modeling required and of interest to MWCOG it is recommended that the passive data source(s) be segmented by vehicle class (at least light

---

[42] Milone, Ronald. “Preliminary Evaluation of Cellular Origin-Destination Data as a Basis for Forecasting Non-Resident Travel.” In 15th TRB National Transportation Planning Applications Conference, May 17-21, 2015. Atlantic City, New Jersey, 2015. https://www.trbappcon.org/oldsite/2015conf/program.html.

[43] RSG. *Michigan Statewide Passenger and Freight Travel Demand Model*. Michigan Department of Transportation. March 31, 2019.

[44] RSG. *CHATS Travel Demand Model Update*. Berkeley-Charleston-Dorchester Council of Governments. March 19, 2019

vehicles and trucks), TOD, residence (within or outside the region), and purpose (work and nonwork). This may require purchasing two datasets: 1) a GPS OD dataset specific to commercial vehicles; and 2) a cellular or LBS dataset, which can provide segmentation based on residency and purpose.

Visitor Travel Patterns

Passive OD data are recommended for use in determining visitor travel patterns for the overnight visitor model described above. Visitor travel is significant in the region and must be accounted for to properly represent travel patterns and traffic. However, MWCOG has not conducted a visitor survey or collected other data on visitor travel patterns in some time. MWCOG should purchase a passive dataset that can provide OD data (and perhaps other information) specific to visitors to the region.

The use of passive OD data for visitor modeling is more recent, but quickly growing, with recent/ongoing studies in several states. Passive OD data for visitors to a region can be acquired at less cost than a specific visitor survey can be conducted and can provide much of the information needed for visitor modeling. Although it typically cannot provide party size, purpose, or mode, passive OD data can provide rich information including the entry/exit mode (auto, rail, air), duration of stay within the region (less than a day, overnight, multiple night), visitor trip/tour rates and attraction rates, trip/tour lengths, and general OD patterns to support either trip-based or basic tour-based simulation models of visitor travel.

Only passive datasets with significant device identifier persistence to allow residence location imputation can provide information specific to visitors. These datasets include LBS and cellular data. An aggregate passive dataset would be adequate to support trip-based visitor modeling. However, a disaggregate (i.e., trace-level) dataset could support tour-based modeling; this may be an option for LBS data, but it is not an option for cellular.

Attraction Rate Estimation

Passive OD data can be used to develop destination choice size terms for the MWCOG region. While size terms have traditionally been estimated from household and/or establishment surveys, both new multiday smartphone household surveys and passive OD data provide more cost-effective information for estimating size terms, specifically for unique travel destinations such as hospitals, major shopping centers, and recreational destinations. Passive data provides observations of far more attractions than establishment surveys or even multiday household surveys and is believed to be a superior data source for size term estimation, provided it is properly expanded.

7.0 TRAFFIC AND TRANSIT ASSIGNMENT

7.1 | DISCUSSION OF STATIC ASSIGNMENT ALGORITHMS

Synopsis of Current Practices

Static traffic assignment (STA) models allocate traffic among competing highway routes. The word “static” means that these STA models are flow-based and do not incorporate the notion of time. This class of models is typically macroscopic and deterministic.

The principle of User Equilibrium (UE), first proposed by Wardrop in 1952,[45] is the fundamental concept used in static highway traffic assignment method. In this principle, a User Equilibrium is reached when no individual travelers can reduce their travel time by changing routes. In other words, all used routes for an origin-destination pair have the same minimum cost. The underlying assumptions in this UE principle are that: 1) all travelers have full knowledge on available routes, 2) all travelers choose routes that minimize travel time or costs, and 3) all travelers have the same value of time. These UE assumptions are obviously simplifications of how people make their route or path choices.

The traffic assignment algorithm that was proposed by Frank and Wolfe in 1956 became known as the Frank-Wolfe (FW) algorithm[46] and was widely adopted as the standard traffic assignment algorithm for many years. This FW algorithm is a link-based algorithm and, as such, computationally simple to implement. Many researchers have developed derivatives of the FW algorithm such as the Conjugate Frank-Wolfe (CFW) or Bi-Conjugate Frank-Wolfe (BiFW) to improve the convergence efficiency. The main criticism about these link-based algorithms is that they discard information about the origin and destination of travelers.

In 1977, Daganzo and Sheffi[47] proposed the Stochastic User Equilibrium (SUE) algorithm, which is also a link-based algorithm. In this algorithm, a Stochastic User Equilibrium is reached when no travelers believe that they can change their expected utility by changing routes. This SUE algorithm is more consistent and intuitive than the FW algorithm in terms of how people make route choices in large-scale congested urban area road networks. The SUE algorithm is typically implemented using the Method of Successive Averages (MSA) that takes a large number of iterations to reach convergence. Cube Voyager provides two methods, namely Burrell’s method and Probit method, for SUE assignment options.

---

[45] John Glen Wardrop, “Some Theoretical Aspects of Road Traffic Research.” Proceedings of the Institution of Civil Engineers 1, no. 3 (January 1952): 325-62, doi: 10.1680/ipeds.1952.11259.

[46] Frank, M., and P. Wolfe. 1956. An algorithm for quadratic programming. Naval Research Logistics Quarterly 3(1-2): 95-110.

[47] Carlos F. Daganzo, Yosef Sheffi, (1977) On Stochastic Models of Traffic Assignment. Transportation Science 11(3):253-274. https://doi.org/10.1287/trsc.11.3.253

In recent years, more advanced UE traffic assignment algorithms have been developed that generally belong to two categories: Path-based and origin-based UE algorithms. These advanced algorithms improved on finding more robust and realistic equilibrium solutions at higher convergence speed. More importantly, these path-based or origin-based algorithms keep track of all used routes and the flow on each of those routes to achieve higher accuracy. Consequently, these advanced algorithms require more computational memory.

Several researchers developed path-based methods to compute a user equilibrium (path-based UE) assignment. One such path-based UE was developed by Dial[48] in 2006 that proved to be very robust and efficient. In this Dial's method, a path-based UE solution is efficiently computed with tighter convergence because it does not require explicit enumeration of the paths between different origin-destination pairs. The select-link analysis for path-based UE assignment is typically computed based on the most likely path flows as outlined in the algorithm by Bar-Gera et al.[49] in 2012. These path-based algorithms are more advanced than the link-based algorithms because they provide a more realistic portrayal of traffic flows among different routes between each origin-destination pair.

The travel demand modeling software that we reviewed, namely Cube Voyager, TransCAD, VISUM, and Emme, all provide the legacy link-based traffic assignment methods using the FW equilibrium assignment algorithms and several variants of the FW algorithm.

For advanced path-based UE traffic assignment methods, TransCAD has implemented a variation of the Dial's path-based UE assignment algorithm mentioned before. Similarly, Emme's advanced assignment method also use a path-based UE assignment method known as the Projected Gradient (PG) method developed by Florian[50] in 2009. Cube Voyager's advanced traffic assignment method also uses a path-based UE assignment method known as the Gradient Projection (GP) algorithm developed by Jayakrishnan et al.[51] in 1994. However, the select-link functionality is not available in Cube Voyager with path-based UE assignment. It should be noted that the PG method in Emme and the GP method in Cube Voyager are similar in computational steps, whereby in each iteration the origin-destination flows are shifted from the maximum cost paths to shortest path in such amounts that the overall projections remain feasible.

---

[48] Dial, R. B., 2006. A path-based user-equilibrium traffic assignment algorithm that obviates path storage and enumeration. Transportation Research Part B: Methodological 40 (10), 917–936.

[49] Bar-Gera, H., Boyce, D., Nie, Y. M., 2012. User-equilibrium route flows and the condition of proportionality. Transportation Research Part B: Methodological 46 (3), 440–462

[50] Florian, M., I. Constantin, and D. Florian. 2009. A new look at the projected gradient method for equilibrium assignment. In Proceedings of the Transportation Research Board Annual Meeting, Washington, D.C.

[51] Jayakrishnan, R., W. K. Tsai, J. N. Prashker, and S. Rajadhyaksha. 1994. A faster path-based algorithm for traffic assignment. In Proceedings of the TRB, National Research Council, Washington D.C.: 75-83.

The first advanced origin-based assignment algorithm was developed by Bar-Gera[52] in 2002. Bar-Gera's 2002 algorithm focused on origin-based link flows aggregated over all destinations. In 2010, Bar-Gera proposed the TAPAS algorithm,[53] or the Traffic Assignment by Paired Alternative Segments, which focuses on pairs of alternative segments as the key building block to the UE solution. A condition of proportionality, similar to the concept of entropy maximization, is used to choose one stable route flow solution. Overall, the family of origin-based UE algorithms produce comparable traffic flow details as the path-based UE algorithms but by using approach proportions and additional computational steps, such as route flow entropy maximization. Research suggests that origin-based UE assignment algorithms are computationally efficient for large urban area networks. VISUM software offers a Linear User Cost Equilibrium (LUCE) assignment option, which is an origin/bush based equilibrium algorithm.

Another refinement to making static traffic assignment more realistic is the multi-class equilibrium models proposed by Dial[54] and Leurent[55] in 1996. In this multi-class equilibrium assignment algorithm, different values of travel time and reliability can be used for different road user classes such as Single Occupant Vehicles (SOVs) and High-Occupancy Vehicles (HOVs). This multi-class equilibrium assignment algorithm is suitable for modeling managed lanes, toll roads and commercial vehicles. These multi-modal multi-class assignment features are available in all commercial transportation software packages.

The standard practice in measuring convergence in traffic assignment is the use of the Relative Gap (RG) measure. For each traffic assignment iteration, gap is measured first as the difference between the total vehicle travel cost and the total vehicle travel cost by all-or-nothing (AON) at that iteration. Then, a ratio is calculated between the calculated gap for that iteration and the maximum gap of all previous iterations. The current modeling practice is to use an RG value of 0.0001 (10^-4) for checking model convergence. At this convergence level, link flows typically become stable from one iteration to the other.

The Gen2 family of MWCOG's trip-based models (e.g., Ver. 2.3.70 or Ver. 2.3.75) uses the static User Equilibrium (UE) traffic assignment algorithm available in the Cube Voyager software. The solution algorithm used is Bi-Conjugate Frank-Wolfe, run to a relative gap of 0.01 for the first through third iteration of the model, a relative gap of 0.001 (10^-3) for iteration four, and a relative gap of 0.0001 (10^-4) for the last iteration of the model.

The UE procedure is run for the following six user classes for each of the modeled time of day periods (i.e., AM from 6 am to 9 am, PM from 3 pm to 7 pm, Mid-Day from 9 am to 3 pm and Night from 7 pm to 6 am):

---

[52] Bar-Gera, H. 2002. Origin-based algorithm for the traffic assignment problem. Transportation Science 36: 398--417.

[53] Bar-Gera, H. 2010. Traffic assignment by paired alternative segments. Transportation Research Part B

[54] Dial R. B. (1996) “Bicriterion traffic assignment: Basic theory and elementary algorithms,” Transportation Science 30, 93-11

[55] Leurent F. (1996) “The theory and practice of a dual criteria assignment model with continuously distributed values-of-times,” in J. B. Lesort, editor, Transportation and Traffic Theory, pp. 455-477, Pergamon, Exeter, England.

- Single-Occupancy Vehicle (SOV)
- High-Occupancy Vehicle with Two Persons (HOV2)
- High-Occupancy Vehicle with Two Persons (HOV3+)
- Medium and Heavy Trucks
- Commercial Vehicles
- Airport Passengers

The model performs assignment in two steps. First, all classes except for HOV 3+ are assigned. Then HOV3+ vehicles are assigned where the resulting volumes from the previous assignment are used as base volumes for HOV3+ vehicles.

Recommendations

1. Given that all four modeling software packages reviewed (Cube Voyager, TransCAD, VISUM, and Emme) use some variant of path-based UE algorithm, they are considered advanced methods and will likely produce comparable, reasonable results. However, we recommend exploring the possibility of select-link functionality within the path-based UE assignment option. In our experience, this is an important functionality for model validation as well as for model applications.
2. Although the four modeling software packages have the advanced path-based UE algorithms, it is likely that speed of convergence and specific results on toll roads and managed lanes could be different based on actual implementation of the assignment algorithms, embedded data structures, and how parallel computing steps have been utilized. Therefore, it is possible to get different numerical results using the alternate software platforms, especially given the complexity of the COG region’s transportation network and presence of managed lanes. So, we recommend that COG consider comparing assignment algorithms in Cube Voyager against VISUM, TransCAD and Emme multi-class assignment algorithm, with the same convergence criteria. The tests would compare base-year assignment results against each other as well as differences between a 'baseline' and 'build' network change such as a capacity increase and/or decrease. The results can help COG staff understand the strengths and weaknesses of different STA methods for the COG region and the extent of variability among the key performance measures along key corridors. Also, the results can inform future software directions in terms of gaining more speed in model convergence, and potentially better cross-sectional and longitudinal validation results.
3. It would be useful to test the need for running a separate assignment for the HOV3+ vehicle class under different assignment algorithms with comparisons to count data in different corridors. It is possible that the two-stage assignment approach would make it more challenging to interpret model results since the paths resulting from assignment of non-HOV3+ demand do not take into account the congestion effects of HOV3+ demand.

4. Engage with travel demand modeling software vendors to learn more about their internal testing and future plan in providing origin-based UE assignment algorithms as part of their software.
5. Overall, we think that it makes sense to continue to use the Cube Voyager's assignment methods for the first round of Gen3 Model validation purposes. Based on the first round of Gen3 validation results and parallel STA control tests, it may be desirable to revisit the need for improving on STA methods and related convergence criteria prior to the second round of Gen3 Model validation.

7.2 | DISCUSSION OF DYNAMIC TRAFFIC ASSIGNMENT

Dynamic traffic assignment (DTA) models allocate traffic among competing highway routes with explicit consideration of time. This is sometimes referred to as "time-dependent paths." The word "dynamic" means that these DTA models capture the changes in network traffic conditions as frequently as 10 minutes in choosing routes between origins and destinations. Consequently, the DTA results are disaggregate and probabilistic in nature. DTA models are used primarily to estimate dynamic traffic flow patterns over the network.

Because queuing and signal delay are dependent on arrival times to nodes, DTA models are more suited to simulating traffic flow and traffic control systems (e.g., intersection traffic signals, freeway ramp metering signals, freeway lane control signals, and dynamic message signs). Dynamic UE (DUE) methods are used to load and track movement of vehicles (or platoons of vehicles) throughout the transportation network and achieve convergence.

DTA models are typically implemented in three main sequential steps: 1) Network loading, either through analytical or simulation approaches, 2) Path set update, based on congestion patterns and travel times and using a time-dependent shortest path (TDSP) algorithm, and 3) Path assignment adjustment, to decide on the extent and location of the demand shifts. These three steps are repeated until a convergence is reached, measured by a Relative Gap measure.

DTA models have only begun to be used in practice in recent years, particularly at the regional level. Most of these DTA models have been applied for freeway management, integrated corridor management, managed lane operations, traffic management centers, incident management, congestion pricing, and emergency *management. For example,* the open-source TRANSIMS[56] DTA model is currently being used by the Northern Virginia Transportation Authority (NVTA) for corridor planning, operations, and project prioritization.

Most DTA models focus on route choice and relatively few are implemented for departure time or arrival time choice. Also, most DTA models are vehicle-based instead of person-based, and

---

[56] U.S. Department of Transportation, Federal Highway Administration. "TRANSIMS - Resources - Transportation Model Improvement Program (TMIP)," June 28, 2017. https://www.fhwa.dot.gov/planning/tmip/resources/transims/index.cfm.

were developed as standalone modeling tools. Also, to our knowledge, only a few activity-based models (ABMs) hadbeen linked with DTA models before 2010.

In the 2010-2014 timeframe, the second Strategic Highway Research Program (SHRP2) launched the Dynamic, Integrated Model System research program (C10)[57] that led to several model development efforts to link ABMs with DTA models.

SHRP 2 Project C10A involved integrating the open-source DaySim[58] activity-based model with the open-source TRANSIMS DTA model through a feedback loop. The integrated model was tested in Jacksonville, Florida for a large-scale regional test with 525,000 base year households and in Burlington, Vermont for a small-scale regional test with 55,000 base year households. The research team needed to put in substantive effort in TRANSIMS network model data development and debugging and coding intersection controls. The research team also noted that substantive efforts were necessary to test and understand the results of traffic operational improvement strategies when compared to testing demand management or pricing strategies.

Similarly, SHRP 2 Project C10B involved integrating the Sacramento, California region’s regional activity-based travel demand model, namely SACSIM (which is a DaySim family of ABM), with a mesoscopic DTA model originally developed by the University of Arizona as an open-source model, namely DynusT[59], but later made it a proprietary software in 2017. The C10B research team also noted about the intensive resource needs for DynusT application as well as some unexpected results from the integrated model for a freeway widening scenario test. The deployment of the DynusT required training on the model, significant network coding efforts, and substantial amount of time to examine and interpret model outputs. Due to time and resource constraints, the model developed in the C10B project was primarily used at a “proof of concept” level and was deemed not ready for real-world applications.

In 2014, the SHRP2 C10 program funded the following four pilot ABM-DTA integration projects and lead adopter projects across the country, which had been documented in a 2018 report[60]:

- Atlanta Regional Commission (ARC): CT-RAMP ABM with the DynusT DTA in a highway setting
- Ohio State Department of Transportation (ODOT): CT-RAMP ABM with the DynusT DTA in a highway setting
- Maryland Department of Transportation - State Highway Administration (MDOT SHA) and Baltimore Metropolitan Council (BMC): University of Maryland’s Agent-based microsimulation travel demand model, named SILK AgBM (for its emphasis on Search Information, Learning, and Knowledge in the travel decision-making process), with DTALite, an opensource mesoscopic dynamic traffic assignment

---

[57] http://www.trb.org/Main/Blurbs/169685.aspx

[58] https://GitHub.com/rsginc/daysim

[59] https://dynust.net/

[60] Scott Smith, et. al, “TravelWorks Integrated Model: Final Report,” June 2018, Prepared for the Federal Highway Administration, Office of Planning, Washington, DC

model that covered the State of Maryland; and BMC's InSITE ABM[61] with dynamic traffic assignment of DTALite that covered most of the urbanized areas in Maryland.
- Metropolitan Transportation Commission (MTC), San Francisco County Transportation Authority (SFCTA), and Puget Sound Regional Council (PSRC): Fast-Trips[62] dynamic transit passenger assignment model (an open source model) to address transit crowding and reliability. Fast-Trips is a schedule-based model but considers transit dwell time dynamically.
- San Diego Association of Governments (SANDAG): Lead adopter project to provide pricing and travel time reliability enhancements to their existing ABM

These pilot projects revealed that the networks for a DTA model require high degree of realism and consequently significant efforts are necessary to examine aerial photographs, street-level photography, signal phasing and timing files, and field data. To make the network development task manageable, some jurisdictions chose to work with multi-resolution networks. Some jurisdictions developed a routable network in-house by maintaining compatibility with INRIX data while others adjusted their DTA model to be able to read traffic signal data from other sources. These pilot projects also addressed ABM-DTA integration issues to maintain activity schedule consistency and exchange level-of-service data. These issues were addressed in the ARC and ODOT pilot projects by building additional computational steps such as an Individual Schedule Adjustment Module (ISAM), and the Accumulated Database of Individual Trajectories (ADIT) module.

Overall, there is a consensus in the modeling practice that travel demand models should consider the influences of travel time and cost in predicting travel behavior, and the ABM models should be integrated with DTA models such that these travel time and cost sensitivities are appropriately reflected in all aspects of the model chain. However, the agencies will need to make investments in staffing, training, and test applications before full-scale deployment.

Recommendations

1. We recommend that a DTA model be considered for the Gen4 Travel Demand Model to implement model enhancements in an incremental approach. This will allow adequate time for MWCOG to settle on the temporal resolution on the ABM side and develop a transportation network that includes traffic control and other operational attributes in addition to capacity-related features. This testing/screening process can start by using a subarea network from the COG regional model. COG should carefully select a subarea network where they have easy access to additional network data. These additional network data would be necessary to transition into a DTA model network. The network transformation process will entail enhancing the network with additional geometric details such as acceleration lanes, intersection turn lanes and turn restrictions; coding additional details on centroid connectors; coding additional traffic attributes such as

---

[61] https://www.baltometro.org/transportation/data-maps/travel-demand-forecasting
[62] http://bayareametro.GitHub.io/fast-trips/index.html#

saturation flow rates and jam density values for freeways and arterials; and adding traffic signal control data.

2. We recommend that a DTA model be thoughtfully integrated in Gen4 such that basic typological, spatial and temporal features are consistent with final model architecture on the demand side in Gen3. We also recommend special emphasis in model integration between the ABM and DTA model to ensure that iterative feedback loops in the integrated model will result in convergent, stable solutions.
3. COG should initiate testing and screening of available DTA models. We recommend that MWCOG focus on commercial software solutions (e.g., Cube Avenue,[63] TransModeler,[64] Dynameq,[65] Vissim,[66] and Aimsun[67]) as well as DTALite since it is used by BMC and already covers a significant portion of Maryland.
4. COG should initiate collecting additional traffic volume and speed data by time of day that will likely be necessary to calibrate and validate a DTA model. COG should explore using INRIX monthly average speed data in 30-minute increments for the selected subarea network for potential DTA model calibration and validation.

7.3 | DISCUSSION OF TOLL REPRESENTATION

The current MWCOG model uses an average value of time in assignment. Tolls are considered in the generalized cost path and represented on the network by time period. MWCOG has developed an automated toll setting/adjustment procedure that iteratively adjusts the toll cost based on a minimum level of service (defined as a volume/capacity ratio of between 0.95 and 1.00). There is no toll/non-toll choice option in the current MWCOG travel demand model. In the MWCOG Ver. 2.5 Model, value-of-time segmentation was added to auto assignment. However, this resulted in an increase in assignment runtime and unclear benefits and was not retained. We recommend that MWCOG re-evaluate the possibility of segmenting trip tables by value-of-time in the Gen3 Model, as we have found that such segmentation improved assignment results in the SANDAG model. MWCOG staff can pursue this experimentation with some oversight from the RSG team. We believe the automated toll adjustment script represents state of the art with respect to toll cost estimation. We suggest MWCOG explore the possibility of reducing the minimum level-of-service to a LOS C/D threshold (0.8 to 0.85) as this may results in a closer match to actual toll values. We also recommend integrating the toll optimization procedure within the main model flow with a user option to run the procedure within the model so that the script can be maintained and be available to all users.

---

[63] https://www.citilabs.com/software/cube/cube-avenue/
[64] https://www.caliper.com/transmodeler/default.htm
[65] https://www.inrosoftware.com/en/products/dynameq/
[66] https://www.ptvgroup.com/en/solutions/products/ptv-vissim/
[67] https://www.aimsun.com/

7.4 | TRANSIT ASSIGNMENT ALGORITHMS AND TRANSIT CAPACITY RESTRAINT

There are several types of static transit assignment methods including all-or-nothing, stochastic, and user equilibrium.[68] In an all-or-nothing transit assignment, all the passengers of an O-D pair are assumed to choose the least cost path. One key issue with transit shortest path, unlike auto shortest path, is that if there are multiple competitive lines between a given pair of boarding and alighting stops, passengers could choose between boarding the first vehicle to arrive or waiting for a later vehicle which results in a lower total travel time. Therefore, the shortest transit path could include multiple routes/lines, each with a certain probability of being used by passengers as a function of headway and in-vehicle time. Dial's algorithm[69] was one of the earliest shortest transit path-builders. It assumes exponentially distributed headways and random arrivals such that the wait time is equal to one-half of the headway. The expected wait time at stops with competitive routes is equal to one-half of the inverse of the sum of the routes' frequencies and assignment to those routes is proportional to their relative frequencies. TRNBUILD uses a similar algorithm.

PT uses an assignment algorithm more similar to Spiess,[70] in which a "strategy" is considered for each boarding decision point. A strategy is a choice of all of the competitive transit lines, and an optimal strategy is one in which the weighted travel time is minimized. This is also referred to as a "hyperpath." The PT module uses a series of logit choice models to determine the probability, and therefore the weight, of each hyperpath.

Because TRNBUILD is no longer being maintained or enhanced by Bentley (the developers of Cube software), we recommend moving to PT for transit skimming and assignment in the Gen3 Model. PT is also required if MWCOG wishes to implement transit capacity restraint in Gen3 models. We suggest experimenting with transit capacity restraint functionality when testing assignment of on-board survey trip tables to the transit network. This will allow the team to gauge the runtime implications and goodness of fit of enabling this functionality. We also may want to test future-year assignments to ensure that transit capacity restraint provides the desired response in future years when Metrorail demand may exceed capacity more significantly than current conditions.

---

[68] Liu, Yulin and Bunker, Jonathan M. and Ferreira, Luis Transit Users' Route-Choice Modelling in Transit Assignment : A Review. Transport Reviews, 30(6). pp. 753-769. 2010.
[69] Dial, R. B. Transit pathfinder algorithm. Highway Research Board, 205, 67-85. 1967.
[70] Spiess, H., and Florian, M. Optimal Strategies: A New Assignment Model for Transit Networks. Transportation Research, 23B (2), 83-102, 1989.(19)].

8.0 SOFTWARE APPROACH

8.1 | OVERALL SOFTWARE FRAMEWORK

The overall approach to the Gen3 Model software framework begins with an understanding of MWCOG's needs:

- **Powerful models –** Best-practice methods and models to create meaningful travel forecasts for transportation and land use planning
- **Up-to-date tools –** Continual improvements to take advantage of the latest advanced in transportation options, modeling methods, and software capabilities
- **Cost-effective ownership –** Sustainable modeling tools that avoid the high costs of custom efforts by leveraging the work of others
- **Proven roadmaps –** Comprehensive and well documented use cases and examples for increased reliability and reduced ramp-up and turnaround time
- **Partner community –** A network of agencies who have implemented similar models to draw on for advice

Based on these needs, RSG recommends building the Gen3 Model upon two software frameworks: the open source ActivitySim software and the commercial Cube software. ActivitySim satisfies these needs because:

- **Models are proven and state-of-the-practice –** Built on, and continually informed by, industry best practices
- **Pooled funding reduces costs –** The collective effort makes it possible to share development and maintenance costs of complex software tools
- **Economies of scale –** Collaboration makes it possible to develop, maintain, and document fast tools that are easy to use
- **User community –** The agency Consortium team members are a valuable resource for modeling support and program management

We recommend continuing to use the existing Cube software for network modeling because it satisfies existing and near-term network modeling needs while avoiding the significant startup costs associated with switching network modeling platforms. Given these two tools, the following model software framework is recommended.

Model Software Framework

Like the existing TPB Ver. 2.3 Model, we recommend running the overall model from the Windows Command Window or Python and calling command line programs such as ActivitySim/Python or Cube procedures. Each of these subprocesses works with their existing efficient binary I/O formats and converts inputs and outputs to open data formats for data interchange as needed.

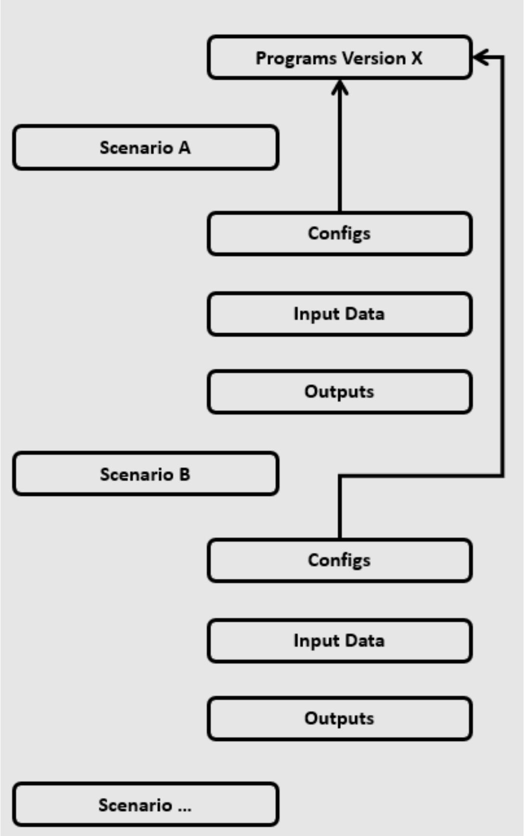


FIGURE 15: OVERALL SOFTWARE FRAMEWORK

Second, and like the existing TPB Ver. 2.3 Model, there is a separation of model components via folders. A set of programs define the current version of the model and each scenario is in a separate folder. The basic design matches the existing model design and the ActivitySim design. Figure 14 illustrates this design and an explanation of the key points is below.

1) The Gen3 Model implementation is managed in GitHub. This includes the programs and a template scenario setup (with settings but minus inputs).
2) Each release of the Gen3 Model is versioned and tagged as a release on GitHub for managing versioning and resolving versioning issues.
3) Each model scenario is in a separate folder and includes three key folders – configuration/settings, inputs and outputs. The version of the code used for each scenario is specified in the scenario settings.
4) A model scenario is run via a master Windows batch file or Python script, with the choice to be made later in consultation with the project team. Python is preferred because it offers a wealth of functionality.
5) Each model component works with its existing efficient binary I/O formats and converts inputs and outputs to open data formats for data interchange as needed. For example, skim matrices in binary open matrix format and data tables in text-based CSV files.
6) There are model components that simply translate data between formats as needed. For example, a model step near the beginning may convert Cube binary network format data into CSV format for use by downstream model components.

This approach to the overall model system framework is proven – having been used by several advanced travel models, reduces cost by using tools (i.e. Windows Command Windows and Python) with a large and active user base, takes advantages of economies of scale by again using Windows batch files and/or the wealth of Python data science libraries, and has a sizable community of users to ensure future success.

8.2 | COMMERCIAL MODELING SOFTWARE, CURRENT FEATURES AND DESIRED FEATURES

The Gen3 Model will continue to rely upon commercial modeling software for steps such as network management, traffic and transit skimming and assignment, and aggregate travel models such as freight models. Following are a list of features important for model integration. We note that such functionality is not limited to activity-based model integration. MWCOG staff conducted a survey of four software vendors in April 2020, including Bentley (Cube), Caliper (TransCAD), INRO (EMME), and PTV (VISUM). We reference vendor responses to the survey where appropriate below. However, we do not provide details as COG will be releasing a summary report of survey responses separately from this document. We note however that vendors claim very similar functionality in their responses, and, therefore, we suggest that COG acquire demonstration versions of the software so that features can be evaluated by COG staff directly.

- Data input\output: Most software has capabilities to import and export networks from shapefile format, comma-separated value (CSV) format, database format (DBF), Excel (XLSX) Open Street Map (OSM), General Transit Feed Specification (GTFS), and/or other transportation software programs. Tabular and matrix data can often be imported and exported from/to CSV, DBF, XLSX, and other formats. Background layers (OSM, Google Maps, satellite images) can be loaded from the web or disk. Most vendors report similar functionality here.
- Network editing: A graphical tool for displaying and editing network features. This tool is used for adding and removing links, nodes, and transit routes, adding to or changing the attributes of links, nodes or routes, etc. We note that the VISUM stands out from the other software packages in terms of its object-oriented network model.
- Network calculations: A tool for calculating network attributes, such as link capacity, from other attributes, such as number of lanes and other attributes. Network calculations are often used to automate modeling procedures, such that the user can rely on programmed calculations rather than manually change multiple data items. All software vendors report similar functionality, though Cube requires scripting for most network calculations, whereas other packages provide calculations through the GUI.
- Matrix calculations: A tool for creating and calculating matrix data. For example, performing mathematical operations on matrices (adding, subtracting, multiplying matrices), and transposing matrices.[71] Matrix calculations can be used to apply travel demand models, such as gravity models and logit mode choice models. Iterative proportional fitting (IPF), known as Fratar adjustment for two-dimensional matrices, is a special type of matrix calculation required for doubly-constraining gravity models but has other applications in travel demand modeling as well. Cube requires scripting for matrix calculations. EMME does not allow the user to edit matrix values in the GUI. TransCAD provides full matrix calculations in its GUI as well as copy-and-paste functionality.
- Traffic path-building and assignment: Tools for finding paths between nodes and/or TAZs through road, bike, or pedestrian networks, and assigning demand to those networks. Inputs include networks, demand matrices, and volume-delay functions. Outputs may include congested travel times on links, network skims, and link volumes. All software vendors report similar functionality. Runtime and stability of convergence

---

[71] A transpose matrix is one in which the rows and columns of the original matrix are flipped. A common application of matrix transposition is when changing a matrix format from production-attraction to origin-destination format.

properties between packages can only be determined via carefully controlled experiments.

- Transit assignment: Tools for finding paths between nodes and/or TAZs through transit networks, and assigning demand to those networks. Inputs include road and transit networks, transit demand matrices, and transit delay functions. Outputs may include transit skims, route level and stop level boardings and alightings, and transfer matrices. All vendors report similar functionality. Again, testing is required to compare runtime and functionality.
- Select link/route tracing: The analyst may be interested in identifying the demand associated with one or more selected links or transit routes. Or they may be interested in tracing the routes used by one or more zone pairs through the assignment process. Such analysis is typically required in the case of traffic impact studies or project-specific analysis. In either case, all software packages provide this capability, though it may be unavailable for certain assignment algorithms.
- Visualizing and reporting results: A travel demand model is a tool that, if successful, provides useful insights into the alternatives and policies being modeled. To that end, the visualization and reporting of results is a valuable feature of modern transportation modeling software packages. All software packages report similar base functionality. Some vendors provided compelling examples of visualizations, such as EMME animations of transit boardings and alightings and transit congestion. EMME and VISUM demonstrate visualization functionality specific to tour and activity-based model results. Other software packages may also offer such functionality.
- Automation: Travel models can take hours if not days to run from start to finish. Therefore automation is a required core feature to ensure replicability, accuracy, and efficiency and reduce labor costs and fatigue. We give high marks to PTV and INRO for relying on Python for scripting instead of relying on proprietary scripting tools or programming languates. We believe it is essential that vendors provide software functionality through a well-documented API that not only allows the user to call and execute software steps from within Python, but that actually returns Python data structures from those steps that can be manipulated in Python. This is particularly useful in the context of activity-based modeling and integration with ActivitySim, which is developed in Python. We can imagine future development activities, such as DTA, where accessing data structures provided by the commercial software will be particularly useful for tight software integration.

There are a number of other software features that are useful for modeling but not necessarily essential. For example, some tools offer Geographic Information System (GIS) integration and robust scenario management. COG uses ArcGIS and COGTools for scenario management, so built-in scenario management functionality may not be as important as it might be for other agencies. Subarea extraction functionality, origin-destination matrix estimation, built-in tools for choice modeling, and other functionality similarly may be very helpful for some agencies, but not necessarily as relevant for activity-based model integration. There are however some additional features that we wish to highlight for COG's consideration as it reviews responses to vendor surveys:

1) Transit capacity restraint: Although the DC area has seen some reductions in transit useage over the last ten years, and transit patronage during the COVID-19 shutdowns is at about 10% of the pre-COVID-19 levels, it is still believed that, over the long term, demand for transit will rebound and, thus, there will be a need and interest to have a model that can represent transit crowding (which, historically, has been an issue on both Metrorail and some bus routes). RSG performed carefully controlled tests of transit capacity restraint using Cube PT for the San Francicso Bay Area's Metropolitan Transportation Commission (MTC). We found that link-level crowding resulted in reasonable demand response and converged assignment, but stop-level capacity restraint resulted in unreasonable demand response and failure to converge. All software packages report some level of transit capacity restraint functionality. We strongly recommend that COG staff evaluate such capabilities using carefully controlled experiments in the coming months.
2) Cloud computing: Cloud computing offers huge economies of scale for intensive computing such as travel demand forecasting and we believe it is essential for software vendors to offer maximum licensing flexibility for deployment of models in the cloud. We found some of the survey responses on cloud computing to be somewhat vague and suggest COG staff follow up with vendors to clarify exactly what is permitted given current license restrictions and whether special pricing is required to enable cloud solutions. In particular, whether there are any constraints in terms of specific vendor-supplied cloud solutions versus publicly-available solutions. For example, we have found that most vendors allow their software to be used in Azure cloud services since such services are essentially an extension of the user's work network, but other cloud solutions may be more difficult to utilize.
3) Tour and activity-based visualization. As noted above, all software packages provide similar functionality with respect to aggregate model result visualizations. However, some packages are now offering visualization solutions specifically for disaggregate models, such as individual trip tracing. We note that CityPhi (INRO) offers compelling disaggregate model output animations, though this package is not included in the core EMME product and must be purchased separately. PTV also demonstrates disaggregate model output visualizations, though it is unclear whether such functionality is available in the core package.

8.3 | DEMAND SOFTWARE, CURRENT FEATURES AND DESIRED FEATURES

The current trip-based model is implemented in the Cube travel modeling toolkit, which was designed for aggregate trip-based modeling. As a result, the Cube toolkit supports the following functionalities:

1) Matrix management and matrix calculations
2) Network management and network calculations (such as for nodes and links and to a limited extent for transit networks)

3) Limited record processing functionality
4) Application of choice models such as multinomial and nested logit models
5) Scripting/customization
6) Submodel integration (e.g. Cube Application Manager)
7) Limited design to take advantage of advances in open source data science software due to a proprietary, and outdated, approach to scripting/customization

This set of building blocks represents the common tools used in aggregate trip-base models. However, with the development and maturation of more powerful and comprehensive disaggregate activity-based models, a new set of software requirements will soon be needed. By focusing on individual households, persons, tours, and trips, activity-based models need a separate and well-developed set of software building blocks. These activity-based model building blocks are:

1) Comprehensive capabilities for managing and manipulating data tables (such as the functionality in Python pandas)
2) Comprehensive capabilities for managing and manipulating matrices (such as the functionality in Python numpy)
3) Comprehensive capabilities for managing and manipulating user defined expressions across matrices and tables (such as the functionality in ActivitySim)
4) Comprehensive capabilities for managing and manipulating both in-memory and persistent data pipelining (such as the functionality in ActivitySim)
5) Comprehensive capabilities for managing and manipulating logging and tracing of calculations (such as the functionality in ActivitySim)
6) Comprehensive capabilities for managing and manipulating multiprocessed/threaded applications (such as the functionality in ActivitySim, including the ability to restart and maintain stable random numbers at any submodel step and also across processes/threads)
7) Comprehensive capabilities for managing and manipulating individual person time/space windows and availability (such as the functionality in ActivitySim)
8) Comprehensive capabilities for extending software with modern open source data science libraries (such as those found in the Python community)

And with any toolkit developed, a user community of sufficient size to manage and improve the toolkit overtime, plus a user community of sufficient financial health to own and fund the development of a quality software solution, including testing, documentation, and improvements is required.

The recommended ActivitySim activity-based modeling software framework includes several of the desired characteristics of demand software. As described earlier, it improves upon the existing Java-based CT-RAMP and C#-based DaySim implementations by building upon

Python, which has the right mix of software engineering and data science functionality for implementing travel modeling software and individual travel models.

With that being said, the ActivitySim framework continues to evolve and is expected to be improved over the years to come. Functionality that is expected to be improved is listed below:

1) More error handling and input checking to improve the user experience under unexpected situations
2) Additional documentation as the user community grows
3) Additional software test coverage for features added for specific regions in order to better future proof everyone's collective core software platform
4) Advances in model design motivated primarily by data-driven analysis as opposed to speculative / research-oriented ideas and interests
5) Continued emphasis on performance improvements and tuning

8.4 | NETWORK MANAGEMENT SOFTWARE, CURRENT FEATURES AND DESIRED FEATURES

Synopsis of Current Practices

MWCOG currently uses a combination of ArcGIS[72] and COGTools[73] for editing and managing transportation networks and scenarios for their Gen2 family of trip-based travel demand models (e.g., Ver. 2.3.70 or Ver. 2.3.75). MWCOG staff also uses Cube Base for some network editing tasks, such as sub-regional or project-planning work. COGTools is an ArcGIS application extension, developed by Daniel Consultants, Inc. (DCI) in 2007, and maintained by DCI and MWCOG staff. COGTools provides the necessary tools for creating, editing, and managing multiple highway and transit networks in an ArcGIS personal geodatabase. In theory, it can also be used with enterprise geodatabase formats, but that approach has not been used in production by MWCOG staff. COGTools is updated on a regular basis, but its last formal documentation is from 2013.

COGTools allows one to edit a master highway network, add future-year network layers and features, build transit routes and transit stops layers and related attribute features, and export networks for each modeled year in Cube and geodatabase formats.

COGTools provides the functionality of maintaining topological integrity to make consistent edits across highway and transit layers when modifying an existing highway network. COGTools allows users to work on enterprise geodatabases (accessed via ArcSDE) of highway and transit

---

[72] ArcGIS Desktop (10.1 or 10.6) or ArcGIS Engine (10.3.1 or 10.6.1, distributed with Cube software)

[73] Qiang Li, Jim Yin, and Daniel Consultants, Inc. COGTOOLS User Guide, Revision 3.0, Washington, D.C.: Metropolitan Washington Council of Governments, National Capital Region Transportation Planning Board, October 17, 2013

networks to manage different versions of network scenarios and to keep track of edits and changes by multiple users. COGTools is used primarily by internal COG staff, though it has been shared with the staff of one local jurisdiction, who is developing its own software, but with similar functionality.

COG staff recently conducted a survey[74] of twelve peer Metropolitan Planning Organizations (MPOs) across the country to understand the current state of the practice in managing transportation networks for travel demand modeling. This survey revealed that out of four peer MPOs who are using the Cube Voyager software[75] for travel demand modeling, two MPOs (namely Baltimore Regional Transportation Board and East-West Gateway Council of Government) are using ArcGIS as the companion software platform for multi-year highway and transit network development and management, similar to the MWCOG's network management approach of pairing ArcGIS and COGTools with Cube Voyage.

This pairing of a companion network management tool is necessary to efficiently and conveniently manage multi-year multimodal networks in a master network geodatabase because Cube Voyager does not currently have a multimodal network manager. The ArcGIS Engine provided as part of the Cube software bundle enables the development of custom scenario management tools such as COGTools by extending the data storage formats beyond traditional highway and zone network (NET) files and a series of text-based transit line (LIN)[76] files. The San Francisco County Transportation Authority (SFCTA) and Metropolitan Transportation Commission (MTC) in the Bay Area, California also uses Cube Voyager for travel demand modeling paired with NetworkWrangler,[77] an open-source Python-based system developed for multimodal network management using a Git repository snapshot to define a scenario.

In contrast, five peer MPOs who are using TransCAD travel demand modeling software, the multiyear network and scenario management task is typically performed using some version of a "Net Manager" type "add-in" script within the TransCAD GIS Developer's Kit (GISDK) environment. The latest version of TransCAD (version 8.0)[78] offers a "Model Manager" for managing scenarios and running individual model steps using an intuitive flowchart-based user interface.

The other major network management software packages, EMME and VISUM, manage multimodal network scenarios through the concept of databank (database) scenarios and separate version files, respectively. In either case, scenarios and version files are independent

---

[74] Jim Yin to Mark S. Moran et al., "MPO Survey of Network Management Practices," Memorandum, January 6, 2020

[75] https://www.bentley.com/en/products/product-line/mobility-simulation-and-analytics/cube

[76] At COG, transit line files are given the file extension of "TB" for TRNBUILD-format files and "PT" for PT-format files.

[77] https://github.com/sfcta/NetworkWrangler

[78] https://www.caliper.com/pdfs/TravelDemandModelingBrochure.pdf

of one another, which clarifies the definition of input files while requiring additional procedures/scripts for building composite scenario files when needed.

In summary, the current state-of-the practice in network management for travel demand modeling appears to involve pairing an outside software toolkit/custom application extension with one of the commonly used travel demand modeling software such as Cube Voyager, TransCAD, VISUM, or Emme.[79]

The other emerging approach of using open-source Python based network and database management scripts seem to work well for staff with computer programming training but could be a source of confusion and errors for new, non-technical, or external agency model users.

Recommendations

1. Given that the ArcGIS-COGTools generally works well for COG's internal and external staff, we propose that we maintain and preferably enhance or refresh the tool for development of highway and transit networks for the Gen3 Model. In our view, there are no strong incentives to replace or change this auxiliary COGTools until such time when COG decides to change the main modeling platform from Cube Voyager to some other software package, especially potentially one with a stronger application progrmaming interface (API). This approach makes sense within the underlying assumption that we are developing the Gen3 Model in Cube Voyager due to schedule and cost constraints. In essence, we do not recommend changing COGTools for the Gen3 Model to avoid the proverbial situation of putting the "cart before the horse." Instead, it is desirable to update/refresh the COGTools so that it has adequate batch processing and feature editing capabilities to address Gen3 Model network development needs. It is very likely that we will need to add new attributes to the transportation network based on the selected Gen3 Model design elements.
2. For Gen4 model development or future phases of Gen3 model updates, we recommend that COG first decides if any change is necessary on the modeling software platform based on tradeoffs between costs and desired functionalities, including the need to have embedded scenario and network management within the same software platform.

8.5 | RECOMMENDED DATA MANAGEMENT SOLUTION - MAINTENANCE AND MANAGEMENT OF INPUT DATA, OUTPUT DATA, VISUALIZATION TOOLS

RSG's recommended data management solution depends on the requirements of the model system and the capabilities and interest of the model users. We understand the MWCOG model needs to satisfy the following requirements and user needs:

---

[79] https://www.inrosoftware.com/en/products/emme/

1) Work with Cube and ActivitySim
    a) Cube manages data in proprietary formats such as binary network (NET), ArcGIS, and Cube binary matrix (MAT or MTX) formats. Routines exist to import and export to open formats.
    b) ActivitySim manages data in Python pandas HDF5 format and OMX format (which is also based on HDF5).
2) Work with COGTools, which manages multimodal networks via ArcGIS geodatabase format and integrates with Cube, via import and export functions.
3) Work with third-party tools used by MWCOG model users such as Cube, Python, R, ArcGIS, and Excel. Cube reads and writes ArcGIS, DBF, CSV, and OMX files. Python reads and writes DBF, CSV, and OMX files. R reads and writes DBF, CSV, and OMX files. ArcGIS reads ArcGIS, DBF, and CSV files. Excel reads DBF and CSV files, but not ArcGIS and OMX files, and no longer writes DBF files. File formats common to all tools are CSV files.
4) Work with binary-based storage options (some proprietary, some open source) for efficient input/output.
5) Tools used by the MWCOG model should be actively maintained, used by a community of significant size in order to ensure the likelihood of maintaining and improving the tool over time, include developing and maintaining high-quality documentation, and be backed by sufficient software functionality test coverage to ensure product stability, overall quality, and user satisfaction.

Based on the requirements and the familiarities of model users, RSG recommends managing the Gen3 Model system as follows:

1) Cube-related modules should work with native, proprietary binary formats for efficient I/O. For example, Cube NET files and/or geodatabase formats.
2) ActivitySim related modules should work with native, open binary formats for efficient I/O. For example, HDF5 and OMX.
3) Both toolkits – Cube and ActivitySim – support import/export from efficient binary formats to open (published) formats for data exchange (i.e. CSV and OMX) between tools.

Maintenance and management of the Gen3 Model implementation should focus on developing and maintaining tools that work with agreed-upon data management technology while making the software easier to use, more capable, and more stable.

Visualization of model results is a “hot topic” in travel modeling, especially with the larger movement to a more data science-oriented profession. Broadly speaking, there are two types of visualization:

1) Static reporting / visualization in the form of model reports such as Cube reports, html pages, R markdown pages,[80] and Python Jupyter notebooks.[81] These visualization tools often provide limited interactive capability while also running in the browser for easy distribution. Reports are typically focused on model users. Cube, ActivitySim, R, and Python provide tools for outputting such reports.
2) Interactive online (cloud-based) solutions for model visualization. Examples of these solutions include MTC's use of Tableau[82] for public involvement and ActivityViz[83] – RSG's open-source travel and activity data visualization toolkit for visualizing travel models, travel surveys, and passive data analysis. These solutions are hosted, provide significant capabilities for querying, filtering, and displaying data, and are typically tailored for an audience beyond the core modeling community.

At a minimum, the RSG team recommends improved static reporting functionality for the Gen3 Model components. A model built with the ActivitySim framework would logically include Python Jupyter Notebook-based reports and, thus, this visualization/reporting method is therefore recommended. Cube based modules, with their limited data science capabilities, would likely export data to open data formats for eventual reporting and visualization with the more capable ActivitySim-based technologies.

A more comprehensive, interactive online (cloud-based) solution for model visualization is recommended once the model system achieves a level of stability and maturity to warrant investment. Developing a professional interactive online solution draws upon a separate software development skillset, and therefore requires the additional cost of translating requirements that is typically only cost effective once the economies of scale are in place from having several sustained users.

---

[80] https://rmarkdown.rstudio.com/
[81] https://jupyter.org/
[82] https://www.tableau.com/
[83] http://rsginc.github.io/ActivityViz/

9.0 QUALITY CONTROL AND QUALITY ASSURANCE (QC/QA)

9.1 |CURRENT NETWORK CODING AND QA/QC PROCEDURES AND RECOMMENDATIONS

MWCOG currently uses COGTools for network coding and Git for version control, though at this point, only some of the COG modelers are using Git. As described above, we see no reason to move away from COGTools until such a time as either COG discovers significant limitations with the software (for example, it may be incompatible with data required for DTA development) and until COG makes a determination on commercial transport software for Gen4 Model development.

MWCOG uses a public-facing database software system (iTIP) to collect information on projects to include in the Transportation Improvement Program (TIP). We believe this represents a best practice solution for management of the TIP process and recommend that MWCOG maintain this software and enhance it as necessary to meet COG needs.[84]

As noted above, some MWCOG model development staff use Git and GitHub for version control, but not all modeling staff have adopted these tools. As a result, there may be some modifications made to existing scripts or new scripts and procedures that may not be versioned (until they get versioned by staff who use Git). We suggest that COG staff adopt Git more uniformly throughout the travel modeling group. We recommend that the current model code be versioned. We recommend a branching structure such that the release branch contain the current release version of the code. Under the release branch there would be a develop_tbm for development of trip-based model code and develop_abm for development of the activity-based model. Once the activity-based model becomes the new release version, the current trip-based model code would be archived and split off from the main release branch of the code. We also recommend that properties, files and other common model inputs be maintained in GitHub under whatever branches those inputs are applicable for. Note that we do not recommend using GitHub for maintenance of large files. These are better shared using a common server location or cloud-based file-sharing with version control.

We recommend that anyone making modifications to the code actively maintain and contribute to the GitHub repository. This may require some internal training in GitHub, but there are many online resources for such training. We also recommend the use of TortoiseGit for the Windows Git plug-in, which makes it very easy to commit changes, pull from and push to remote repositories. We recommend that users commit their changes as soon as they verify that their changes are bug-free so that the repository is up to date.

---

[84] COG staff is currently transitioning to a new TIP database software package, called the Project Info Tracker, which is being developed by EcoInteractive.

We also recommend that COG travel modeling staff use issue tracking software. Such software can be integrated with Git if desired, which can be useful to link commits back to specific issues, but we leave this to the discretion of COG staff. Regardless, issue tracking software such as GitHub Issues can be a useful tool for understanding limitations of current software, potential future enhancements, and which members of the team are responsible for different aspects of the code. We currently use GitHub Issues for development of ActivitySim, and the RSG team will use GitHub issue tracking for development and maintenance of the Gen3 Model code as well.