CMAP Strategic Plan for Advanced Model Development

Final Report of the CMAP Advanced Travel Model Cadre

September 2009 – June 2010

- Thomas Rossi (Cambridge Systematics)
- John Bowman (Bowman Research and Consulting)
- Peter Vovsha (Parsons Brinckerhoff)
- Kostas Goulias (University of California Santa Barbara)
- Ram Pendyala (Arizona State University)

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Strategic Plan for Advanced Travel Modeling at CMAP

As part of our Strategic Vision submitted to the State Legislature in 2006, CMAP seeks to be the authoritative source for regional data collection, exchange, dissemination, analysis, evaluation and modeling for the Chicago metropolitan region.

This is the final report of the CMAP Advanced Travel Model Cadre. The Cadre convened, under contract with CMAP, for the period September 2009 through June 2010 to investigate and recommend a program for improved travel modeling practice at CMAP.

Advanced travel modeling remains at the cutting edge of the transportation planning industry. Comparatively few travel modeling professionals have dedicated themselves to extending the state of the practice by embracing an entirely new theoretical paradigm for mathematically representing travel activity.

The Cadre was composed of national leaders in advanced travel model practice:

- Thomas Rossi (Cambridge Systematics)
- John Bowman (Bowman Research and Consulting)
- Peter Vovsha (Parsons Brinckerhoff)
- Kostas Goulias (University of California Santa Barbara)
- Ram Pendyala (Arizona State University)

The Cadre was led by Kermit Wies, CMAP Deputy Executive Director for Research and Analysis.

Executive Summary

Advanced travel modeling and forecasting is generally composed of three frontiers in transportation research. 1) Microsimulating **individual daily activities** rather than aggregate trip-making behavior, 2) representing **transportation network use** in response to real-time conditions rather than under average equilibrium conditions and 3) estimating, calibrating and validating transportation supply and travel models using **richly detailed empirical data**.

Following completion of our 2007 Travel Tracker Survey, we are now prepared to advance our travel forecasting capabilities at CMAP. In August 2008, we hosted an Activity-based Travel Demand Model Symposium to identify the agency's needs and potential directions in travel model practice. Based on the symposium proceedings, CMAP assembled a Cadre of nationally recognized professionals to develop a strategic plan for advanced modeling at CMAP.

The Cadre's product is this multi-year strategic work program to replace the current sequential (a.k.a. 4-step) trip-based models at CMAP with a set of advanced activity-based demand estimation procedures for use in regional transportation and land use planning.

The Cadre was directed to address the topic in terms of CMAP's planning and policy priorities, then to translate these into advanced modeling features and finally to specify the computing requirements for implementing those features. In the course of doing this, salient management principles for CMAP's advanced travel model work plan emerged:

- **Microsimulation** techniques will form the basis for all new travel forecasting tools. These are refined applications of agent-based mathematical and statistical procedures that individually enumerate and explain decisions and actions in comparatively microscopic detail.
- Successfully implementing advanced travel models hinges on CMAP's commitment to an **ongoing survey and data acquisition** program that ensures a full array of empirical information to properly estimate, calibrate and validate each model component.
- CMAP's **planning and policy development priorities** will establish the order in which advanced modeling components are developed and incorporated into the agency's research and analysis stream. For example, Freight Planning and Transportation Pricing are current policy priorities at CMAP. Advanced modeling techniques will be pursued with special attention to incremental implementation in order to prove their worth in immediate application.
- Persistent attention to **model integration** provides CMAP with the assurance that activity-based model techniques provide a compounding return on agency investment; i.e. they continually increase CMAP's capacity to provide unified analyses of the urban system.

Management

A fundamental management principle is being recommended in this plan based on CMAP's long-standing commitment to actively linking technical evaluations to its policy actions. CMAP will continue to pursue advanced model development guided by its policy priorities and will expect the new modeling tools to satisfactorily perform in the time frame dictated by the policy action at hand. There are advantages and risks associated with this approach.

The advantages of this approach are found in the immediate policy-relevance of the modeling tools. Policy makers, when confronted with complex technical decisions are typically unsympathetic to the idea of delaying their choices until adequate analysis tools can be developed. Consistently orienting model development priorities with policy priorities sends a satisfying signal to policy-makers that their technical investment is responding to customer needs. Maintaining this congruency is an attribute of CMAP's organizational culture and this plan seeks to reinforce that.

In terms of risks, we observe that some of our peers have embarked on model improvement programs that separate budgeting, staffing and schedule for model improvements from the agency's primary plan evaluation activities. This is done to minimize the risk of adopting technical innovations upon which the agency's policy commitments depend; output might be opaque or counter-intuitive, not consistent with earlier methods, or require unavailable data and computing resources.

At first, this isolation from the policy mainstream might be preferred. Model development is, in effect, under quarantine until deemed safe. But it also places the value of advanced modeling tools, i.e. improved policy responsiveness, off the critical path of the agency's primary public policy mission. Over time, this segregation weakens our analytical responsiveness. Instead, we accept the risk of coupling our model development work program with policy priorities to ensure that robust model-based forecasting remains an agency asset.

Resources

CMAP' modeling legacy permits us to consider our modeling and forecasting tools as long-term institutional assets. Tracking model development with the agency's policy priorities will extend the time frame for full-scale conversion to the activity-based paradigm across all aspects of travel forecasting, but will also establish a comfortable framework for CMAP planners and policy analysts to consider, test and adopt more transparent methods, robust data and advanced computing resources.

CMAP **staff** intends to develop its suite of advanced models in-house with consultant support. To accomplish this, we will foster a commitment among professional staff by maintaining advanced model specialization as a career track with a strong emphasis on the need to integrate modeling skill with planning knowledge. CMAP's entry-level professionals, regardless of their previous educational background or training, are presently encouraged to immerse themselves in planning and policy development activities. Success in evaluating and analyzing planning strategies provides a good indicator to managers of the potential to thrive in a more rigorous quantitative environment. Once basic modeling cognition is identified at the entry level, advancement can be benchmarked against tangible and consistent contributions to CMAP's planning products. This approach to staff development parallels the plan for continuous model integration in that CMAP modelers will be rewarded based on their contributions to the agency's mainstream policy agenda.

CMAP presently employs four full-time-equivalent (FTE) staff who can be fairly labeled "modelers" ranging from two to 25 years experience. Another six FTE provide direct data development support to modeling activities, though their efforts also support general planning evaluations. An additional four FTE are similarly devoted to data development oriented toward Web services. Information technology at CMAP is supported by four FTE with additional support under contract. This total is roughly equal to the number of CMAP professional staff devoted to preparing technical evaluations under CMAP's general planning work program. With an overall agency headcount of nearly 100, this seems an acceptable and productive balance of skills and effort. Nevertheless, because modeling is a relatively arcane element of urban planning, recruitment and retention of staff modelers should not be neglected.

CMAP **funds** model development through our annual Unified Work Program in which agency priorities are established in response to policy objectives and regulatory requirements. CMAP has clearly defined policy products such as the Long-Range Comprehensive Plan and Transportation Improvement Program that must be updated at prescribed multi-year intervals. Supporting activities such as planning evaluations, data development, public involvement and community assistance are scheduled within the multi-year cycle with some predictable regularity. Fundamental research and analysis is most defensible when responding to real, rather than perceived, agency needs. For this reason, while advanced modeling techniques certainly existed during preparation of the most recent long-range comprehensive plan, their value was not especially apparent until we began enumerating the policy questions that *could not* be addressed using current techniques.

During the four years leading to the adoption of *Go To 2040*, CMAP expended approximately $750k on basic modeling research and nearly $3M on data collection and acquisition (including a large household survey). While these data development activities are essential to a wide variety of CMAP programs beyond modeling, this total (approximately $4M) is an acceptable and productive balance of expenditures. Preliminary budget planning for two years beyond adoption of the *Go To 2040* indicate a similar overall level of financial commitment with greater emphasis on advanced model development now that the basis for an ongoing household survey program has been established.

CMAP's Goal for Advanced Modeling Capabilities

The Cadre recommends that CMAP share the same principal model improvement goal as its peer MPOs nationwide: To substantially convert from trip-based to activity-based modeling procedures over the next decade.

In preparing this report, the Cadre observed that practical implementation of activity-based models by MPOs is a relatively young endeavor; beginning in the mid 1990s in the U.S. Early adopters included New York, San Francisco, Sacramento, Atlanta, Denver, Portland OR and Columbus, OH. Each of these regions embraced the new paradigm for their own unique set of reasons. But one attribute they all shared was that there was no previous functioning application from which to learn or borrow.

In reviewing available documentation, it appears that these pioneering model development teams approached their task deductively. That is, they began with fundamental premises about how to mathematically represent personal activity and aimed to assemble a closed sequence of computational expressions. This closed sequence of expressions is what most practitioners label "the model". And, often as not, "the model" is portrayed to non-practitioners as a comprehensive explanatory tool. Experienced model managers are certainly aware that, in reality, the complex web of information interpreted by millions of individuals when making urban travel choices defies intellectual comprehension much less explicit mathematical representation. Nonetheless, the fundamental relationships between policy objectives (e.g. congestion reduction, air quality conformity, land use management) and existing analysis tools (e.g. regression, gravity, logit and network path algorithms) are so primitive that the promise of a new method that explains everything at once is irresistible.

CMAP's present advantage is that we can study these pioneering examples and gain significant insights into which activity-based model features are fundamentally robust and where the vulnerabilities lay, particularly in demonstrating the efficacy of proposed planning strategies. We can then proceed inductively and implement those model features most likely to bear immediate fruit. Salient strengths and weakness evident from our review are:

- Advances in **GIS technology** significantly increase the availability of spatially derived data as model inputs.
- Digital availability of **parcel-based** property assessments is an important addition to the set of modeled land use attributes.
- Portable **GPS tracking devices** significantly increase the quality and accuracy of travel surveys.

- Data resulting from traditional travel diary **surveys** remain insufficient in volume, variety and quality to adequately inform models of the full range and significance of urban activity.
- Current activity-based model results remain beholden to **transportation level-of-service measures** derived from traditional (averaged) network path-building techniques. This results in a dilution of the advantages that microsimulation brings by the inferior, but readily available, data resources produced by traditional methods.

CMAP's guiding principle in prioritizing model development tasks, therefore, is to adopt activity-based microsimulation components that yield immediate policy explanations to CMAP planners while focusing our development resources on overcoming the disadvantages that make these implementations vulnerable.

Cadre Proceedings

The specific technical recommendations of this strategic plan are presented as three independently authored reports by the Cadre team leaders with emphasis on distinct areas of advanced model development. The three teams were:

- Policy Response
- Data Development
- Computing Environment

In authoring their reports, team leaders began by refining their topic's overall construct. All cadre members were encouraged to communicate with each other during this step to clarify roles. Each team leader assembled his team from the remaining four cadre members. Once each team was assembled, team leaders produced a work plan for their topic based on objectives provided by CMAP. The work plan included individual team assignments and a schedule for product delivery.

Each team leader independently spent 30 hours in residency at CMAP. The purpose of the residency was to examine CMAP's day-to-day operation and meet with staff from all areas of the agency. In this way, CMAP was able to convey its research and analysis culture, particularly with regard to regional planning research. In addition, cadre members used this opportunity to investigate CMAP's data holdings and information technology resources as they relate to this project.

All project communications and documentation were managed in a wiki environment using Google Sites. This permitted active collaboration in the preparation of project documentation. The site's organization was driven by its content in order to permit exploration of new directions, but also to host the final deliverables. In addition to original postings, the site permitted active feedback and commentary. The site included personal blog space permitting cadre members to articulate their individual perspectives. The site has been archived and is available for review upon request.
http://sites.google.com/site/cmapadvancedurbanmodel.

In addition, all cadre members participated in monthly on-line 60-minute teleconferences. During these teleconferences, project progress and wiki content was reviewed and discussed. For these teleconferences, cadre members provided up-to-date posted content; actively defended their contributions and discussed project objectives and concerns.

Upon completion of the primary reports, each cadre member was invited to compose a short essay briefly outlining any outstanding issues or concerns they felt were not sufficiently addressed in the main body of the report. These are included as appendices.

Chapter 1: Policy Response

Principal Author

Thomas Rossi
Cambridge Systematics, Inc.
100 CambridgePark Drive, Suite 400

Cambridge, MA 02140

tel 617 354 0167

fax 617 354 1542

e-mail trossi@camsys.com

www.camsys.com

Other contributors

Kermit Wies, CMAP

John L. Bowman, Bowman Research and Consulting

Peter Vovsha, Parsons Brinckerhoff

Konstadinos Goulias, University of California, Santa Barbara

Ram Pendyala, Arizona State University

Summary

The Chicago Metropolitan Agency for Planning (CMAP) formed a cadre to develop a strategic plan for advanced travel demand modeling at CMAP. The objective of the cadre was a multi-year strategic work program to replace the current sequential ("four-step") travel demand models at CMAP with advanced activity-based demand and microsimulation procedures for use in transportation and land use planning.

This document represents a plan for implementing advanced travel demand models to inform policy decisions in the Chicago region.

This document recognizes that CMAP, in its role as the regional planning agency for the Chicago area and the designated metropolitan planning organization (MPO), has a wide variety of policy and planning analysis needs that require a complex set of analytical procedures. These analysis needs include:

- Transportation plans programs and air quality conformity
- Comprehensive regional plans and policies;
- Freight and commercial vehicle movements;
- Land use and economic impact assessments.

To adequately address these analysis needs, it is necessary to have improved modeling tools that go beyond the state of the practice and may require data that are not yet available and may take years to assemble. However, it is also desirable to have improved model components in place quickly to meet immediate planning needs.

To address this paradox, the document presents the components of a "**first generation**" model that can be completed within three years of the initiation of model development, and a "**second generation**" model that can address issues that require additional data and model components development before they can be confidently implemented. The first generation model will be a complete model system using the best available relevant procedures, which will be capable of producing much of the necessary policy and planning analysis information required by CMAP. The second generation model will use data to be collected over the next few years as well as newer analytical procedures, including those not yet tested in a "real" urban area setting. This model will allow CMAP to address problems that cannot be examined completely using the procedures currently available. The second generation model will likely be implemented in phases over several years, with some features and capabilities introduced earlier as data and other resources become available.

The proposed **first generation** model is characterized by the following features:

- A population synthesis procedure based on a proven existing method.
- A spatial level of detail similar to the existing zones and subzones used in CMAP's existing model, with added detail in specific locations where needed (such as for transit access analysis).

- Transportation networks that are based on existing model networks – static, with some added detail.
- Models that estimate long term choices such as regular workplace and school locations, vehicle availability, transit pass holding, and workplace parking subsidy.
- Activity based models of person travel by the region's residents, including daily activity pattern, tour formation and composition, tour and stop level destination, mode, and time of day choices, with intra-household interactions modeled explicitly.
- Tour based models of external travel (residents leaving region, non-residents entering region, through travel).
- Simple models for other person travel, including airport access and "special generators".
- Tour-based and supply-chain models of freight and commercial vehicles activity.
- A static multi-class highway assignment procedure, including trucks, tolls, etc.
- A static transit assignment process including capacity constraints and parking lot choice.

The proposed **second generation** model is characterized by the following features:

- A dynamic population forecasting procedure that "ages and migrates" individuals to represent changes in household's life-cycle
- A disaggregate spatial level of detail such as parcels or uniform grid
- Dynamic networks suitable for traffic/transit microsimulation
- Expansion of the vehicle availability model to include vehicle type
- New activity and tour-based models to estimate external travel, visitor travel, airport use, transit tours and commercial vehicles tours.
- Economic and land use prediction tools using the added capacity provided by activity-based models.

The general order for the development of the proposed first generation model is to assemble and prepare the model estimation data sets and being model estimation, then to perform the bulk of model estimation and prepare validation data sets, and finally to complete model implementation/programming and perform model validation. Data collection for the second generation model, including freight, visitor, and airport surveys and the development of a parcel database, can begin immediately and should continue into future years. The second generation model development may include phased component introduction, with components such as freight, airport, external, economic/land use, and traffic/transit simulation as well as an enhanced activity based model phased in and continuing over several years.

Introduction

The Policy Response Chapter of the Strategic Plan represents a plan for implementing advanced travel demand models to inform policy decisions in the Chicago region. While the cadre did not unanimously agree on all aspects of the various details of the modeling, data, and computing approaches, a general consensus was reached on the approaches as documented.

The recommendations are more conceptual than specific in many places, particularly in describing the second generation model. It is not intended to provide all technical details of the modeling approach, for three reasons. First, since the model development process, especially for the second generation model, will not begin immediately, there is no need to make decisions that may change before that process starts, due to changes in available data, institutional issues, or resource constraints. Second, it makes sense to leave some of the details to those who will develop the model, who can make decisions along with CMAP after thoroughly examining the data and any additional information that may become available. Finally, the document will be more readable to a wider audience if it is not bogged down in such details.

The Cadre recognizes that CMAP, in its role as the regional planning agency for the Chicago area and the designated metropolitan planning organization (MPO), has a wide variety of policy and planning analysis needs that require a complex set of analytical procedures, as described in the next chapter. It is necessary for improved modeling tools to be in place quickly to meet immediate planning needs. To address this objective, the document presents the components of a "**first generation**" model that can be completed within three years of the initiation of model development, and a "**second generation**" model that can address issues that require additional data beyond what are available now or can be collected quickly, or model components that may require additional research and development before they can be confidently implemented. Basically, the first generation model will be a complete model system using the best available relevant procedures, which will be capable of producing much of the necessary policy and planning analysis information required by CMAP for its current work program. The second generation model will use data to be collected over the next few years as well as newer analytical procedures, including those not yet tested in a "real" urban area setting. This will allow CMAP to address issues that cannot be examined completely using the procedures available for the first generation model. The second generation model will likely be implemented in phases over several years, with some features and capabilities introduced earlier as data and other resources become available.

Policy Analysis Requirements for CMAP

As the federally designated MPO for the Chicago metropolitan area, CMAP is responsible for fulfilling all federal transportation planning requirements, including the development of the long range Regional Transportation Plan and the Transportation Improvement Program and the performance of air quality conformity analysis. CMAP also performs many other transportation planning functions in the region, including freight planning, coordination of project construction, congestion management, analysis of proposed transportation investments, and bicycle and pedestrian planning. CMAP is currently developing GO TO 2040, the comprehensive plan for the region.

CMAP views travel models as hypothesis testing tools, not simply formulas that are given a set of fixed inputs and spit out "the answer." The models can therefore help shape transportation policy and planning and are not just a means for analyzing expected results of plans developed exogenously.

Analysis of Transportation Investments and Strategies

The existing model is frequently used to analyze major capital projects, including both highway and transit projects. About 60 major capital projects and numerous operational strategies needed evaluation in the course of preparing the most recent regional plan. CMAP also needs to examine individual proposals and strategies in the context of project and environmental studies.

As is the case in most large urban areas in the U.S., projects proposed to be funded under the FTA New Starts program also need to be analyzed. There are specific requirements for ridership forecasting associated with New Starts projects, and the new model will need to be capable of fulfilling these requirements.

Air Quality Conformity

Federally required air quality conformity analysis uses input data that are produced by the travel demand model. The Motor Vehicle Emission Simulator (MOVES) has been released by the U.S. Environmental Protection Agency (EPA) for estimating emissions from highway vehicles (U.S. Environmental Protection Agency 2010). MOVES2010 will be officially approved and will serve as a single comprehensive emissions modeling system to replace MOBILE for state implementation plans (SIPs) and regional or project-level transportation conformity analyses. MOVES is designed to estimate emissions at scales ranging from individual roads and intersections to regional and national scales. This provides an opportunity to use more disaggregate outputs form travel demand models in emissions analysis.

Policy Analysis

Policy analysis is an important component of regional planning at CMAP. There are various transportation and planning policies where travel models can evaluate the effectiveness of proposed policies and strategies. These mightinclude:

- Transportation pricing policies (tolls, congestion pricing, parking, transit fares, fuel prices, etc.);
- Land use policies (development regulations or subsidies, jobs-housing balance, open space preservation, etc.);
- Traveler information/incident management (generally, "Intelligent Transportation Systems");
- System management, including signal priority, fare and service coordination, managed expressway lanes, and

- Policies related to mode choice incentives including pedestrian and bicycle travel.

Some policies of interest in the region have been difficult to analyze using the existing trip-based models. In some cases, the aggregate, static nature of the conventional travel modeling process used by CMAP does not provide sufficient detail to properly understand the impacts of the proposed policies.

Freight and Commercial Vehicle Analysis

The Chicago region has one of the largest and most complex freight systems in the world. Contributing to its size and complexity are the region's location in the middle of the U.S., the locations of historic termini of railroads in the area, the substantial agricultural production within a relatively short distance, and of course the region's large population. A substantial number of warehouses and freight terminals are located in the region. For this reason, CMAP needs to understand the region's position in the national and international freight economy.

A large part of the transportation activity in the Chicago region is related to freight and commercial vehicle movements. Much freight is moved into, out of, and through the Chicago area, and goods from elsewhere must be distributed through the region. There are also many commercial vehicles in the region that do not carry goods; these include vehicles used by commercial service providers, vehicles used for public services such as mail delivery and waste collection, and fleet vehicles used in certain businesses. The freight system in the region is multimodal. There is a robust rail system, and how goods travel on the region's rail system, yard operations, getting goods from yards to destinations and to other cities, and transfers to other railroads must be considered.

Non-Motorized Travel

There is a significant amount of non-motorized travel (walking and bicycling) in the Chicago region, especially in denser areas served by transit. Since non-motorized trips, especially walking trips, are relatively short, the relatively aggregate level of spatial detail in the existing model makes accurate estimation of non-motorized travel demand difficult. The same can be said for pedestrian access to transit.

Several issues related to non-motorized travel are important to consider, including the following:

- The effects of sidewalk availability including its relationship to transit use (walk access/egress).
- Disability status, including the ability to use the transit system and to drive.
- A large portion of school related activity uses non-motorized travel modes. School travel has many unique characteristics, including the youth of most travelers, a high percentage of escorted trips, provision of school bus service in some areas and use of public transit in others, and the concern over safety for student travel.

Land Use and Economic Analysis

Like many travel demand models, the current CMAP model considers the effects of land use and economic effects through fixed socioeconomic inputs. The effects of the transportation system and changes in accessibility on land use are not as systematically considered. Forecasting land use and demographics, particularly locations of future development, is affected by a number of factors beyond transportation accessibility, including political considerations, physical constraints on land development, governmental regulations such as zoning, and the relatively small number of decision makers involved in development decisions for large tracts.

The impacts of independent jurisdictional land use development decisions on transportation system performance are of concern, as are the effects of regional transportation operations and capital investments on future land use patterns.

Proposed Plan for Advanced Models of Policy Response

The proposed model plan described in this section is designed to meet the policy and analytical needs discussed in the previous section while considering the desire to have usable advanced modeling tools as quickly as possible. These competing objectives resulted in the concept of **first generation** and **second generation** models.

The **first generation model** is intended to be able to be completed within only a few years of the initiation of model development. It therefore is proposed to be estimated using data that are available now or can be collected quickly and uses the best available relevant procedures. This model will include an activity based model of travel by the region's residents where the activities and travel of each individual in a synthetic population representing the entire region will be modeled individually, with the ability to sum results at various levels of aggregation as needed for particular analyses. The first generation model also includes more conventional procedures for modeling trucks and freight, travel by visitors, and special markets such as airport access. Also included are enhanced static traffic and transit assignment procedures, which, while similar to the processes used by CMAP now, will include some advanced features such as the use of capacity constraints in transit vehicle loading. While there are some policy issues whose analysis would require more advanced procedures that will take more than a few years to implement, the first generation model can function as a complete stand-alone model system that will provide CMAP with the means to address current planning analysis functions more efficiently and to address some deficiencies where current procedures are lacking.

The **second generation** model will address new policy issues beyond what the first generation model is capable of and will provide more robust ways of performing existing analyses. The second generation model will likely be implemented in phases over several years, with features and capabilities introduced as data and other resources become available.

The second generation model will also use new data that will be collected over the first few years of the model development process. It will also feature newer modeling procedures that have not yet been field tested in and may require additional research and development before they can be confidently implemented. These features will allow CMAP to address issues that cannot be examined completely with the procedures available for the first generation model. It is planned that the second generation model will be more detailed spatially, perhaps at the parcel or small-grid cell level. More advanced procedures for special markets will also be implemented. The second generation model will also feature the integration of the activity based personal travel model with microsimulation of passenger, commercial vehicle, and transit vehicles as well as transit passengers.

Table 3.1 summarizes the features of the first and second generation models. The remainder of this chapter describes these features for both models. The first generation procedures are presented in greater detail, not only because they are based more on existing procedures in use elsewhere, but also because the second generation procedures need to be flexible to consider future advances in modeling research and data collection.

Table 3.1. First and Second Generation Model Summary

	First generation model	**Second generation model**
Population synthesis	Based on a proven existing method (CMAP, ARC, PopGen, etc.)	Dynamic process considering changes in households over time (based on longitudinal data) and responsive to scenario conditions.
Unit of geographic detail	Zones/subzones, with point locations simulated for choices where geographic specificity is critical	Parcels or microzones
Transportation networks	Based on existing model networks–static, with added detail	Time-dependent networks for dynamic traffic/ transit microsimulation
Long term choices	Regular workplace and school locations, vehicle availability, transit pass holding, workplace parking subsidy	All first generation choice models (possibly modeled as part of a dynamic synthetic population generator or land use model), plus vehicle type model
Person travel by region's residents	Estimated from existing household survey: Daily activity pattern, tour formation and composition, tour and stop level destination, mode, and time of day choices, intra-household interactions modeled explicitly. Based on best existing activity modeling practices.	Take advantage of currently ongoing research and development in activity based models to make models more robust and efficient. Use additional survey data collected over next few years.
External travel (residents leaving region, non-residents entering region, through travel)	Residents: Simple tour based model (home-external-home) estimated from household survey. Non-residents: Trip based model based on external station survey (if available in time) No explicit visitor travel model	Residents: Updated as necessary to be consistent with internal person models Use all external station survey data Visitor travel model – Tour based model based on visitor survey data

Table 3.1. First and Second Generation Model Summary (continued)

	First generation model	**Second generation model**
Other person travel	Generally use existing procedures for airports, “special generators,” etc. Analyze existing household survey and other available data to improve these procedures in the short term (e.g. group quarters).	Develop specific models for airport, university, and other markets using data collected over the next few years, including an airport survey.
Freight and commercial vehicles	Enhancement of existing procedures, making process consistent with the new activity based model.	New tour based models of freight and commercial vehicles, using data from new commercial vehicle surveys and other data. New national (or greater) scale models of commodity flows/economic activity
Trip assignment	Static multi-class highway assignment procedure, including trucks, tolls, etc. Static transit assignment process including capacity constraints and parking lot choice	Dynamic traffic simulation a la SHRP C10, TRANSIMS, or other process tied directly to disaggregate model application. Transit rider and vehicle simulation a la SHRP C10 or other process tied directly to disaggregate model application.
Economic and land use modeling	Continue using existing procedures for demographic data development and forecasts	Integrated land use-transportation model, perhaps using an existing model (e.g. PECAS, TRANUS, UrbanSim) or new model developed specifically for Chicago area Economic modeling may be coordinated with state agencies

Population Synthesis

Both the first and second generation models will generate a synthetic population and stochastically microsimulate the movements of each member of the population. The synthetic population will represent all persons living in the CMAP modeling region for the time period (typically defined by an analysis/forecast year and land use scenario).

First generation

Population synthesizers are being used successfully in a number of existing activity based models. Many of these programs are publicly available through host agencies or academic institutions. While there are some algorithmic and programming differences among these programs, the outputs are similar—essentially they are lists of households and the persons living in them with their locations, household characteristics (such as income level), and person characteristics (such as age, gender, worker status, and student status). It is recommended to examine the range of options available for population synthesis at the time that the first generation model begins and determine the most advantageous program for CMAP. The decision process will involve discussions with the agencies/organizations who maintain the programs. At this time the following population synthesis programs are known to be available:

- PopSynWin (developed by UIC and currently in use at CMAP)
- TRANSIMS (Federal Highway Administration)
- PopGen (Arizona State University)
- Atlanta Regional Commission
- Denver Regional Council of Governments
- Sacramento Area Council of Governments

Second generation

There are two main reasons to consider, for the second generation population synthesizer, a dynamic procedure that incorporates household formation, aging, births and deaths, and migration as an evolutionary process in the region. First, it will produce a more accurate set of synthetic households and persons for use as inputs in forecast year scenario model runs since it is based directly on existing household information and data on how households and persons change over time, through the use of a longitudinal survey. Second, it will allow examination of past experience in explaining present day travel behavior. For example, questions such as the following could be answered: Do the temporarily or newly unemployed individuals travel more or less than the long term unemployed? Do persons who recently migrated to the Chicago area behave the same as persons of similar personal and household demographics who have lived in the region for a long time?

The incorporation of dynamically evolving synthetic populations may be one of the later improvements implemented in the second generation model, mainly because it will take time for sufficient longitudinal data to be collected, and also because no specific procedure to do this has been implemented in an existing urban area model.

Transportation Networks and Geographic Resolution

Unit of geographic aggregation

The basic unit of geography for model application should be the smallest possible at which relevant data can be gathered and processed. Most travel models use aggregate geographic zones that are sized based on available information and computational limits. Unfortunately, aggregation results in statistical error resulting in a loss of fidelity and policy sensitivity.

The recent availability of GIS layers used for property tax assessment, have made development of parcel-based modeling geography appealing. Advanced microsimulation techniques and computing resources are also oriented toward overcoming the computational limits imposed on the matrix handling found in trip-based models

Considerable effort is required to prepare parcel data for model estimation, validation, and application at the parcel level. Furthermore, practical accommodation of the required amount of computation for a parcel level model has not yet been achieved in a region as large as CMAP although computation is not expected to be a significant issue if matrix handling of parcel-to-parcel times and costs are not used in the model.

The recommendation, therefore, is for the first generation model to use a more aggregate level of geography, with a goal of a finely grained level for the second generation model. While aggregation to some level will be done for the first generation model, it is still desirable to use a somewhat finer level of spatial resolution than exists in the current CMAP model. This will allow the new model to address issues of “lumpy” loading of trips to the highway system, variations in transit walk access and egress times causing inaccurate average measures of transit time, and accuracy in estimating travel time for walk and bicycle trips. Perhaps the most efficient means of developing the zone system for the first generation model is to use the existing CMAP subzone system (used for trip generation) as a starting point, and aggregating according to desired fidelity criteria.

In the second generation model, as sufficient parcel level information becomes available throughout the CMAP region and computing advances improve the performance of a potential parcel level model. The recommendation is to have a complete parcel based approach. It is still likely that there will need to be some accommodations to the large number of parcels, such as sampling of destination choice model alternatives. A systematic method of disaggregating large rural parcels for use in evaluating future scenarios will also be needed.

Highway network

The type of data required for the highway network depends on the type of highway assignment procedure used in the model. As discussed in Section 3.8, a static equilibrium assignment process is recommended for the first generation model and a disaggregate, dynamic traffic simulation process is recommended for the second generation model. Since the existing CMAP model uses a static equilibrium highway assignment process, the existing highway network provides a good starting point for the first generation model network, adding detail as necessary. It will also be worthwhile to consider other network revisions to enable improved analytical techniques.

The dynamic traffic simulation envisioned for the second generation model will require a different type of highway network. Features such as turning lanes, acceleration/deceleration lanes, traffic signal timing, and other traffic control devices will be required. While this will require a substantial data collection and assembly effort, an additional process to synthesize information for some locations will be necessary as well. The data synthesis process could also be used to address any gaps in existing data, which will be time consuming and expensive to assemble.

Transit network

As is the case with the highway modeling process, the first generation model will use a similar transit assignment process to that used in the existing CMAP model. This means that the existing route based CMAP model transit network can be used as a starting point for the first generation transit network. Since the existing CMAP model already includes virtually all transit routes in the region, there is no additional detail to be added to the transit route system.

As discussed in Section 3.8, the second generation model will include dynamic transit rider and vehicle simulation. This requires the use of a transit network that is timetable (schedule) based, rather than route based. Rather than using individual routes with headway information, the second generation model will require information on each individual run made by a transit vehicle between its start and end points, along with the starting time of the run and the times the vehicle arrives/departs each stop. This information is, of course, available for the existing transit system but in many cases will have to be synthesized for forecast year and alternative scenario transit networks.

Long Term Choices

Long term choices are the result of singular personal attributes that affect activity and travel behavior; such as age, wealth, health, job and children. The attributes induce quite permanent choices including the location of the regular workplaces, school locations, the number and type of vehicles available to households, transit use and parking cost.

Location choices for most activities are dependent on traveling from home for activities organized along multi-activity tours. Changes in transportation accessibility or costs, which often result from the implementation of transportation policy decisions, can affect activity location choices. However, the effects on some types of activity locations, notably workplaces and schools, cannot as easily be affected by changes in accessibility or travel cost. To properly analyze the effects of policy decisions, location choices that can be characterized as longer term decisions need to be modeled separately from other activity location choices.

The subjects of the other proposed long term choice models, including vehicle availability, parking subsidy at work, and transit pass holding are more likely to be directly affected by policy decisions. Vehicle availability is affected by transportation investment decisions (greater transit accessibility in very dense areas, for example, is hypothesized to decrease auto ownership) as well as other policy decisions such as land use planning actions to limit parking. Parking subsidies at work can be subject to policy decisions about parking pricing and taxation. The decision of whether to use transit passes can be affected by transit investment decisions and by policies related to pricing and availability of the passes. Since these choices impact subsequent travel decisions and therefore the use of the transportation system, their effects need to be considered properly to accurately analyze the effects of the policy decisions, and these models therefore need to be part of the analysis process.

Data to estimate most of these models would come from the household activity/travel survey, supplemented by data from the transportation networks and information such as school enrollment and employment by area. Assuming the necessary data from the survey are available, these sub-models are all included in the first generation model, with one exception. While a model of the number of vehicles owned by each household can be implemented in the first generation model, implementation of a vehicle type model is proposed for the second generation model.

Regular workplace location

In the first generation model, the household locations will be determined by the population synthesizer, based on control totals at a specified geographic level such as zones. Workplace locations, however, will not be determined by the population synthesizer in the first generation model although there will likely be exogenously determined zonal totals for employment based on existing employment data or forecasts.

The purpose of the regular workplace location model will be to determine the workplace location for each (synthesized) worker where he or she most commonly reports to work. The model will use information on regular workplaces reported in the household survey data set to determine the probabilities that the regular workplace is in specific locations (including, possibly, the individual's home). It is likely that this will be a choice model with zones as the alternatives, with a specific workplace location simulated among parcels in the chosen zone based on employment data.

The second generation model will consider the use of dynamic models that determine changes in workplace locations over time. This would be part of the dynamic model discussed above under the second generation population synthesizer. Another possibility for the second generation model is to use a land use model (see Section 3.9) to tie workplace locations (or specific jobs) to workers using information on such characteristics as industry type, person occupation, income group, full-time/part time worker status, and other information that might be generated by the land use model.

Child's school location

Similar to the workplace location model, the child's school location model will determine the school location for each (synthesized) student. The model will use information on school locations reported in the household survey data set to determine the probabilities that the school is in specific locations. It is likely that this will be choice model with zones (that contain schools) as the alternatives. In most cases, there will be only one school per zone.

As with the regular workplace model, the second generation model will consider the use of dynamic models that determine changes in student status and school locations over time. This would be part of the dynamic model discussed above under the second generation population synthesizer. If there is a land use model (see Section 3.9) that includes information on school locations, this information could be incorporated into the second generation model as well.

Vehicle availability

The first generation vehicle availability model will estimate probabilities that each synthesized household owns zero vehicles, one vehicle, two vehicles, etc. The choice will be based on household/person characteristics (for example, number of persons, number of workers, income level, and regular workplace locations of workers), land use characteristics of the residence area, and accessibility measures of the residence area by auto and non-auto modes. This model will be estimated using household survey data.

The second generation model will consider not only the number of vehicles owned by households but also the vehicle types, which would be useful information for emissions or energy analysis. This type of model is not commonly used in urban area travel demand forecasting, but it is not uncommon in the literature. It will also be estimated using household survey data although additional information beyond what is available in the existing household survey data set on vehicle types may be needed.

Parking subsidy at work

The choice of mode to work may depend on the cost of parking at the workplace and whether and how much the employer subsidizes the employee's parking cost. It may be possible to develop a model that predicts whether a worker pays to park at his (non-home) regular workplace, the cost of parking, and how much of the cost is subsidized by the employer. Input variables might include characteristics of the worker (industry, occupation, income level, etc.) and of the workplace location. Model estimation data would include the household activity/travel survey and parking cost data for different parts of the region.

Transit pass holding

Whether a person holds some kind of transit pass is typically determined by frequency of use. It is proposed to develop a model that determines the probability that an individual chooses to hold a transit pass and if so which type. The model would be based on characteristics of the person and his or her household (for example, vehicle availability, income level, work location, etc.), the pricing of passes relative to pay-as-you-ride pricing, and the type of transit service available in the residence area, especially to the regular workplace location.

Activity Pattern

It is the consensus of the cadre that to meet CMAP's policy analysis objectives, the complete daily activity patterns of individuals living in the CMAP model region must be modeled, including the list of activities performed and their sequence, the activity locations and timing (start/end times), the tours that comprise the travel related to the activities, and the modes used to travel between activity locations and the home. Because activity patterns of other household members have been shown to significantly impact individual activities, interactions among household members will be explicitly modeled. These interactions include activities performed jointly among household members and escort activities where the main activity of one household member is to accompany another.

Generally, the process for modeling activity patterns is similar among existing models. The main differences among the various approaches are:

- There are different ways in which intra-household interactions are considered in the models. Some models do not consider any household interactions. The consensus of the cadre is that including intra-household interactions is important for CMAP and feasible for the first generation model.

- There are several ways in which activity-based models consider temporal and spatial constraints in activity pattern, location, and time of day choice. Some models consider these constraints explicitly only in the time of day choice model, where time windows required by previously scheduled activities and related travel are blocked out in scheduling other activities. In other models, the numbers and locations of activities may be constrained based on duration requirements for activities and the times required to travel between their locations.

- The amount of decision making that is modeled sequentially and the specific sequence in which choices are modeled varies. While sequentially applied choice models are recognized as oversimplifications, they are often used to make models computationally practical. It is common to model activity pattern, location, mode, and time of day choice sequentially, as well as the latter three choices separately for main activities (tour level) and intermediate tour stops (trip or stop level). One of the areas with the most variation among models is the point at which time of day choice for the main activity is modeled, which may be before location and mode choice, between destination and mode choice, or after destination and mode choice. There is no consensus to suggest that one particular placement of time-of-day choice is more beneficial than another.

In the first and second generation models, for this series of sub-models an activity pattern for each person will be simulated. Consideration of intra-household interactions will require that household members' activity patterns be simulated together. The general process will be as follows:

- The activities to be performed will be determined.

- The activities will be organized into tours. Each tour represents a time when the individual leaves home, and the specific activities performed on each tour and their sequence are determined.
- The locations of activities will be determined.
- The timing of each activity (start and end times) will be determined, using information on travel times to and from activity locations and considering times needed to perform other activities in the daily pattern.
- The mode choices for trips between activity locations will be determined.

Note that the sequence in which these steps are performed—indeed they generally represent simultaneous choices and may not be modeled separately—is not specified in this recommendation. Existing activity based models have implemented choice models in different orders, and no consensus on the best sequencing has emerged. The details of sequencing will be determined by CMAP and the model development team prior to commencement of the model development process, and the approach adopted in the first generation model may be changed for the second generation model, based on continuing research on activity based models.

Table 3.2 summarizes CMAP's policy analysis requirements and discusses how the activity based modeling approach, in connection with the other model components proposed here, will provide the best analysis tool for these policies.

To meet the policy and planning objectives, it will be necessary to include variables throughout the model system to consider specific objectives. For example, variables related to land use type and density must be included to allow the model to be sensitive to the effects of alternative land use patterns. The pedestrian and bicycling environment affects how people travel by non-motorized modes, and so variables representing this environment must be included in models of mode choice, and possibly activity generation and destination choice. To correctly analyze policies such as road pricing, how travelers consider tradeoffs between cost, travel time, and reliability of travel time must be explicitly included in the model, and so individually simulated values of time as well as measures of system reliability must be represented in the model's variables.

Table 3.2. Relation of Activity Based Model Features to CMAP's Policy Analysis Requirements

Policy Analysis Requirement	Relation to Activity Based Model Features
Analysis of Transportation Investments	Generally, transportation investments improve the transportation level of service for those trips that encounter the benefits. These level-of-service changes result in some travelers changing their activity locations and timings, as well as mode and route choices. Activity- based modeling is the preferred method of capturing these effects because it integrates the various choices, considers tours, and reduces aggregation error. This approach also provides the necessary disaggregate input to dynamic traffic simulations permitting a more accurate means of analyzing traffic impacts of projects.
Air Quality Analysis	Performing accurate air quality analysis requires accurate estimates of vehicular travel, vehicle speeds, and the temporal distribution of travel. The activity based approach is seen as being essential to accurate estimation of travelers' time of day decisions. The traffic simulation process that will rely on the activity based model for input will provide more accurate estimates of travel speeds than conventional models.
Regional Planning Policy	Transportation pricing, land use policy, traveler information/incident management, roadway management, and policies related to pedestrian and bicycle travel. Land use and non-motorized travel are discussed below. Pricing analysis using travel demand models has long suffered from the need to aggregate travelers into market segments and assume that all travelers within each segment made the same time-cost tradeoff decision regarding whether to use a priced facility. The disaggregate activity based approach eliminates the need for this assumption, and each traveler can have a unique simulated value of time. Analysis of traveler information/incident management is dependent on traffic simulation, which will be implemented using inputs from the activity based model as previously discussed. Analysis of roadway management policies combines the techniques used for traffic simulation and pricing.
Freight and Commercial Vehicle Analysis	Freight analysis is not directly related to the person travel considered in the activity based modeling approach and is discussed separately in Section 3.7. However, the more accurate modeling of person travel—and therefore auto travel—will result in better highway level of service estimates that affect truck and commercial vehicle travel.

Non-Motorized Travel	While the major factor in accurately analyzing non-motorized travel is a fine level of spatial detail, the tour based nature of the activity based model is also critical in identifying walk trips that are part of trip chains made by travelers.
Land Use and Economic Analysis	Activity based models by themselves are insufficient to perform land use and economic analysis, but they will work hand in hand with the economic and land use modeling procedures proposed for the second generation model. These processes are discussed in Section 3.9.

Activity categorization

The precise categorization of activities will require detailed study of the household survey data set. However, it is advantageous to have as detailed a categorization as possible as experience in other areas shows significant differences in travel behavior based on activity purpose. Depending on data availability, it may be possible to introduce more detailed activity categorization in the second generation model.

Activity categorization may be considered by activity type (e.g., work, school, shopping), companionship (with family members, friends, other known or unknown), and location (home, usual workplace, usual school location, other). In this way critical data gaps for future survey efforts might be identified.

Tour formation/stops

The tour will be the primary unit of travel in the first generation model, and quite possibly in the second generation model. Each tour made by a person will begin and end at home; work based subtours will also be modeled. An activity pattern will be simulated for each (synthetic) person, specifying the number of tours, the purposes of the primary and secondary activities, and the number of stops by purpose on each tour (on the way to and from the primary activity).

The first generation model will incorporate consideration of interactions among household members in the specification of individual activity patterns. Interactions among household members affect activity and travel decisions in a number of ways, including the following:

1. Explicit linkages among activity patterns for members of the same household.
2. Joint activity participation among household members.
3. Escort activities, where one household member drives or picks up another but does not participate in the activity himself or herself.
4. Allocation of activities, such as household maintenance activities, to specific household members.

5. Allocation of automobiles to individual household members.

From a policy analysis standpoint, the most important of these are the first three since they most directly affect the amount of travel, activity location, and mode choice (including carpooling), all of which might be affected by policies or transportation improvements. These areas should therefore be given priority in the development of the daily activity pattern models. The first two have been implemented in existing activity based models for U.S. urban areas (along with #4), and so they should be included in the model. The remaining items are optional for the first generation model, but they are expected to be included in the second generation model.

Tour level mode and destination choice

It is common practice in existing activity based models to designate for each tour a primary activity—the main reason for leaving home on the tour. For tours with more than one activity (stop), this activity maybe defined by duration, distance from home, based on a hierarchy of activity purposes, or in some other way.

If the primary activity is school, the activity location is assumed to be the school location simulated in the school location model. If the primary activity is work, then whether the work activity takes place at the regular workplace location must be simulated. One way of doing this is through a binary choice model. If the regular workplace is chosen, then no tour level destination choice is needed.

For all tours except school tours and work tours to the regular workplace, the location of the primary destination is simulated. For the first generation model, zones are likely to be the spatial alternatives as finer spatial resolution would likely result in a model that would be impractical to implement. Size variables for the zones would be similar to the input data used in trip attraction models of four-step model systems (e.g. retail employment for shopping activities). Specific activity locations could be simulated among parcels in the chosen zone based on appropriate data, depending on the activity purpose. Again, the appropriate data would be similar to the data used in trip attraction models.

The tour level mode choice model for the first generation model is likely to be similar to those used in existing activity based models (and conventional trip based models, for that matter). Modal alternatives for CMAP are likely to include single occupant, two-occupant, and multiple (more than two) occupant auto; bus (perhaps segmented by local versus express) with walk access and with auto access; commuter rail (i.e. Metra) with walk access and with auto access; el/subway with walk access and with auto access; walk; and bicycle.

The modal alternative chosen is sometimes referred to as the “main” mode for the tour. Since mixed mode tours are not uncommon, the main tour mode represents a judgment by the model developer. Often a hierarchical approach is used, where the modes are ranked, and the tour mode is determined as the highest ranked mode in the hierarchy. For example, if commuter rail with auto access were ranked highest, any tour with a commuter rail-auto access trip would be assigned that tour mode; if walk were ranked lowest, a walk tour would consist only of walk trips. The purpose of the tour level mode choice is to condition the mode choices for the trips between stops on the tour. On an auto tour, the auto must usually accompany the traveler to every stop (perhaps excepting some nearby stops within walking distance); on a non-auto tour, the single occupant auto mode is generally unavailable (since an auto was not brought along), and the other auto modes are less likely than they would be on auto tours.

No specific changes are assumed for the second generation model from the first generation model. CMAP should track future research in activity based modeling to assess whether superior modeling methods might emerge for mode and destination choice. One possibility might be joint mode and destination choice models although it might also be worth considering time of day choice jointly as well.

Stop level mode and destination choice

The stop level mode and destination choices depend on the tour level choices. For destination choice, the locations of activities for intermediate stops (i.e. not the primary activity) depend on both the home and primary activity locations, and the modal alternatives available (e.g., on a transit tour, one can only visit locations that are served by transit). Mode choices for trips to intermediate stops depend on the main mode choice (whether or not an auto was brought along on the tour).

The stop level destination choice model will be similar to the tour level model. The main differences would be the following:

- The level of service (cost/time) variables would be based on the locations of both the home and the primary activity location (or a previously simulated stop).
- The set of alternatives might be constrained based on the locations of the home, primary activity, and previously simulated stop, and available modes for the trip. This might allow for a finer resolution of location alternatives.

The stop level mode choice will consider the constraints implied by the choice of tour mode. In some cases the constraints might be substantial (one or two alternatives). Rules such as the need to bring the auto on an auto tour or the lack of availability of an auto if it was not brought or left at a park-and-ride facility would be imposed.

No specific changes are assumed for the second generation model from the first generation model. CMAP should track future research in activity based modeling to assess whether superior modeling methods might emerge for mode and destination choice.

Time of day choice

Time of day choice is among the most complex model components. Time of day choices are dependent on a number of factors, including desired times to perform activities, required start times (e.g. work hours, school schedules, start times for attended events); hours of operations at some types of activity locations (e.g. stores, libraries); the cost of travel, including levels of congestion at different times of day (and perhaps variable pricing by time of day as well); the need to schedule other activities, perhaps on separate tours; and the need to schedule sustenance activities such as eating or sleeping, which often do not require travel and are not part of a tour.

Both discrete choice and continuous models have been used to simulate time of day choices in activity based model research although models so far implemented in urban areas have used discrete choice formulations. The first generation model, therefore, is recommended to be discrete choice although the inherent correlation among temporal alternatives may not be explicitly addressed. At the tour level, the model must jointly simulate the start and end times of activities. Stop times for intermediate stops before or after the primary activity will be simulated based on the times of the primary and other activities already simulated, as well as the travel time needed to travel between consecutive activities. These requirements effectively constrain the choice sets for most activities.

The level of temporal resolution may be determined as part of the model development/design process, although it is recommended to be no greater than a half hour for the first generation model. A resolution of 30 minutes or finer allows for consideration of time-related planning issues such as peak spreading and congestion pricing while considering the availability and accuracy of the available data, including household survey data and traffic counts. Aggregation of overnight periods might be considered although this would introduce issues with having alternatives of different sizes (in this case numbers of minutes).

The second generation model might consider the use of continuous time. This has already been implemented within some models developed in academic settings and tested using real U.S. urban area data (e.g. CEMDAP, FAMOS). One possibility would be to develop a synthetic schedule generator that simulates activity generation first and then performs a scheduling function, as in the CEMDAP application for the Southern California Association of Governments called SimAGENT. Moreover it may also be possible to account for temporal and spatial constraints in scheduling activities as well as the coupling constraints that are requirements to schedule activities with other persons.

External Person Travel

"External travel" includes travel entering and leaving the Chicago region by land transportation (air travelers are discussed in Section 3.6), travel that begins and ends outside the region but passes through it, and travel made within the region by visitors to the region. Since the activity/travel models described in Section 3.4 pertain to residents of the region, the only component of external travel that will be considered in those models is travel made by residents from their homes to locations outside the region. The easiest way to model such travel would be as tours where the primary activity location is designated as the point at which the person leaves the region. While such tours may not be completed in the same day, or return to the region through the same route, they can be assumed to do so in the model without loss of generality, and the tours from the household survey data set "completed" for the purpose of model estimation.

The remaining components of external travel will have to be modeled using separate procedures from those described in Section 3.4. These procedures are described below.

Because the CMAP modeling region is so large and because of the travel barrier of Lake Michigan, external travel represents a small component of total travel in the region. It has therefore been decided that data collection for trips entering and leaving the region will not be a high priority, as discussed in the data development plan. This means that the remaining components of external travel will have to rely on alternative data sources, including older surveys conducted in the region, national surveys such as the National Household Travel Survey and American Community Survey (ACS), traffic and transit ridership counts, growth forecasts for places outside the Chicago region, and the existing CMAP model.

No changes to the modeling of external travel are envisioned in the second generation model.

Through person/auto trips

Through travel—that which passes through the region but originates and is destined outside it—will be modeled as highway vehicle trip tables with origins and destination at external stations—the points at which roadways in the highway network enter/leave the region. The through trip tables from the existing CMAP model will serve as the starting point for these tables, with updates based on traffic counts at the external stations and, perhaps, growth forecasts for places outside the Chicago region.

Travel to/from Chicago region

This component consists of modeling trips made by non-residents of the region which have one end inside the region and one outside. Because tour information will not be available for these trips, they will be modeled as individual trips, or perhaps as simple tours between external stations and internal zones. The home, or production, end would be considered to be the external station.

For work travel, it should be possible to estimate person trip tables of external travel to the region. The primary data source would be the ACS. Once such trip tables are developed, it may be possible to develop simple destination choice models based on distance, time, and cost for the trip and employment in the zones of work activities. This model would enable trip tables to be developed for forecast scenarios. There would likely be insufficient ACS data to develop mode choice models for external work travel, and so it is likely that the tour mode choice model would have to be used to model mode choice for these trips. Time of day could be determined using simple percentages based on ACS data although there are no data on return trips, and so time of day percentages for work to home travel would have to be determined from other data (perhaps the regional household survey).

For non-work travel, in the absence of a sufficient survey data source, there are two options: use of existing trip tables, perhaps updated using traffic count information, or transfer of a model from another region. The transferability of such a model has not been validated, but its use would be the only way to estimate changes in destination or mode based on policy or transportation system changes.

Internal travel by visitors

Data from a visitor survey will be required to develop the models of travel by visitors to the Chicago region. The CMAP visitor survey, part of the continuous survey program proposed in the data development plan, will provide such a data source. The timing of the availability of data from the visitor survey will determine whether the visitor model component is part of the first or second generation model.

The visitor model will use two main segments, based on whether the visitor is in the region for business or personal/leisure reasons. It will be an activity/tour based model similar to that used for resident travel. Activity pattern, destination, mode, and time of day choice components will be included.

Some socioeconomic/demographic variables used in the resident models will be eliminated or deemphasized in the visitor models due to lack of data from the shorter visitor surveys (for example, rather than the number of vehicles from a vehicle availability model, a simple model of whether the visitor has a vehicle available (through rental or bringing one's own car on the trip) may be used. The visitor models will also differ from the resident models in terms of additional variables (such as party size) and choice alternatives (such as taxi and hotel shuttle modes).

Other Person Travel

This section discusses special markets of travel that, while included in the resident and external travel modeling processes described in Sections 3.4 and 3.5, likely require special treatment to accurately analyze.

Airport ground access and egress

Data from an airport survey will be required to develop the models of travel to and from O'Hare and Midway airports (and possibly future proposed airports). The proposed airport traveler survey, part of the continuous survey program in the data development plan, will provide such a data source. Data from the airport survey is not expected to be available in time for use in the first generation model, and so air travelers will continue to be modeled as in the existing CMAP model (i.e. as a "special generator").

The second generation airport traveler model will be estimated from the new airport survey data. It will use two segments of travel type, business or personal/leisure reasons, and two segments of traveler residence, visitor to the region or resident, for a total of four segments. Generally, the model will be a trip based, with the round trip between the airport and a non-airport location (home, hotel, etc.) as the unit of travel. Destination, mode, and time of day choice components will be included.

Group quarters resident travel

No specific survey data have been collected for group quarters residents. For the first generation model, therefore, activity patterns for group quarters residents will be obtained using the same models estimated for all residents (which is essentially what happens in the existing CMAP models).

The data development plan includes specific plans to include a sample of non-institutionalized group quarters residents, especially college students and older residents living in group quarters. For the second generation models, therefore, these residents can comprise separate market segments for the resident activity based models.

Special events travel

There are "special" events that generate considerable travel but do not occur on most weekdays. They include sporting events, concerts and music festivals, large block party type events, and parades.

Because each event does not occur on the average day, the best way to consider the generated travel is to model the travel to and from each event separately, perhaps adding trip tables/tour lists for event travel to those from the "everyday" models (resident, external, airport, and visitor travel) prior to trip assignment to represent an "event day."

It may be possible to conduct surveys of attendees at various types of special event to obtain data to estimate special event models, as is being done in some other urban areas (Phoenix, Dallas). These models consist of location choice (for the non-event end of the trip) and mode choice. If these surveys are not conducted, consideration of transferring models from other locations (although transferability of such models has not yet been demonstrated) may be considered, or perhaps information from another model component that includes destination and mode choice could be used.

Freight and Commercial Vehicle Modeling

A survey of business enterprises conducting commercial movements in the region is proposed in the data development plan. While this survey is expected to provide valuable information about freight and commercial vehicle movements, data from it are not expected to be ready for use in the short term. The data development plan classifies the demand for freight and commercial vehicle movements as **intra-regional commerce** or **inter-regional goods shipments**.

Intra-regional commerce

The modeling of intra-regional commerce can be improved by using a tour based structure, as has been done in some urban areas outside the U.S. such as Calgary, Canada (Hunt and Stefan 2007). Tours within the region are conducted by commercial enterprises, including transportation and warehousing interests, and public sector service providers (e.g. postal services, garbage pick-up). The tour based modeling approach can be implemented by estimating models using data from the survey of business enterprises conducting commercial movements in the region, and this is recommended to be part of the second generation model.

Inter-regional goods movement

The modeling of inter-regional goods shipments is more complex, in large part because of the need to consider information from a far greater geographic area than the Chicago region. The Chicago region's location and size makes it a national hub for commercial goods movement and a significant international port, which results in substantial commodity flows into, out of, and through the region. Decisions about mode and route choices within the region depend on a variety of factors, including the type of commodity, the origin or destination of the shipment outside the region, the availability and suitability of freight modes for the particular shipment, the need for mode switching, and pricing considerations involving private transportation providers (trucking companies, railroads, marine transportation companies, etc.) and public organizations (agencies levying tolls or regulating providers), and economic conditions inside and outside the region. Further complicating the analysis is the movement of a significant amount of freight through the region (truck, rail, and intermodal), some of which uses intermodal facilities within the region.

The effect of regional conditions on inter-regional freight movements might be analyzed using an economic model (the Computable General Equilibrium (CGE) model described in Section 3.9 would be one example). Another option might be a national scale (or greater) commodity flow model although it might be difficult for such a large model to exhibit the proper sensitivity to regional conditions.

Tour based and logistics supply chain models have begun to be used in freight forecasting, and current research has continued to refine these types of models. This work includes the framework originally developed by Cambridge Systematics for the Los Angeles area (Fischer et al 2005), statewide models in Ohio and Oregon, and research conducted by the University of Illinois at Chicago (UIC) in the Chicago region and nationally. The logistics chain includes operations such as intermodal and same-mode transfers, shipment reconfigurations, and warehousing.

UIC has been developing an activity based freight modeling framework consisting of five modules: firm generation, supply chain replication, shipment forecasting, logistic decisions, and network analysis (Samimi et al 2010). This process is analogous to the passenger activity based modeling approaches such as the one described in Section 3.4. In the UIC framework, individual firms or groups of firms are the modeling agents.

The intra-regional commerce and inter-regional goods movement modeling processes must be linked. The beginning and end of the logistics chain within the Chicago region for an inter-regional commodity movement are part of the intra-regional commerce being modeled, and accurate analysis would require the tour based process described above. This includes goods that might enter the region via truck, rail, air, or water that might reach their final destinations by transferring to other trucks (or perhaps trains), as well as the reverse for goods leaving the region.

The outputs of the second generation freight and commercial vehicle modeling process will be lists of vehicle tours and trips that would be input into the traffic simulation process described in Section 3.8. External freight movements will be treated as individual trips to/from the appropriate gateway (external station), with a tour of movements within the region at the beginning or end of the trip modeled by the intra-regional commercial vehicle model. The tour lists should include the modes (by segment, so that intermodal transfers can be considered) and times of stops, preferably at the same level of resolution as the passenger models.

Trip Assignment

Highway assignment

Disaggregate dynamic simulation is seen as the future of highway assignment. While there has been significant research into the use of traffic simulation with travel demand models, the consensus of the cadre was that this integration has not yet proven to be practical. The first generation model, therefore, will continue to use a conventional static equilibrium highway assignment procedure, similar to the procedure used in the existing CMAP model. This will be a multi-class assignment procedure, with consideration of recent research into non-link based assignment processes and explicit consideration of intersection delay. Priced roads and managed lanes will be explicitly considered in the assignment procedure.

It is expected that ongoing research will yield a practical integration of an activity based travel demand model with dynamic traffic simulation in the near future. In such a process, individual auto tours generated by the activity based models, with their stop locations and times, would have their paths simulated by a traffic simulation program. The highway travel times resulting from the traffic simulation would then be reintroduced into the activity based demand model, providing more accurate level of service inputs to these models than can be achieved through the use of static network procedures.

Possible methods to perform this type of integrated model include products of the two Strategic Highway Research Program 2 (SHRP2) C10 projects, currently being developed for case studies in Jacksonville and Sacramento, and TRANSIMS deployment case studies being funded by the Federal highway Administration. While not being integrated with an activity based model, TRANSIMS is currently being tested in Chicago as part of a project conducted by Argonne National Laboratory. The recommended second generation procedure is to incorporate one of these procedures into the model as soon as one of them is deemed practical.

Transit assignment

The first generation transit assignment procedure will continue to be an aggregate process. The main enhancement compared to existing CMAP modeling procedures will be explicit consideration of transit vehicle capacities in the process.

While there has not been as much research into the integration of disaggregate transit simulation with activity based models, some work into this area has begun, notably in the SHRP2 C10 project in Sacramento. In this project, the transit tours generated by the Sacramento activity based model (DAYSIM) will be individually simulated, as will each individual transit vehicle run. Each tour segment, therefore, will be simulated in terms of the specific transit vehicles boarded, the boarding time, the stop location, the alighting time and stop, transfer times and locations, and walk or auto access and egress trips. This will result in more accurate mode choice and transit path building models, and more accurate results for transit operations planning and ridership forecasting. The recommended second generation procedure is to incorporate a procedure such as this into the model, probably at the same time the highway procedure is introduced.

Non-motorized travel

As discussed in Section 2.5, planning for non-motorized travel is a key issue for CMAP. The activity based models described in Section 3.4 will provide lists of tours and trips made by walking and bicycling, including origin-destination information. The model will be capable of estimating the effects of transportation level of service along with land use and demographic variables on the choice of whether to use non-motorized modes. It may also be possible to include variables that consider the effects of investments in non-motorized transportation facilities on mode choice. So the proposed activity base model structure provides needed analytical capability.

It is not common practice to assign walk and bicycle trips. Walk trips are particularly difficult to define paths for given that they do not generally share the facilities used by motorized vehicles and are usually short, requiring a very fine level of spatial resolution to accurately analyze. There is no good method for accurately assigning pedestrian volumes using a regional travel demand model. While it is recognized that there may be value in estimating facility (e.g. sidewalk) specific volumes to understand issues such as crowding, there seems to be little need to have assigned pedestrian volumes on every facility in the region. A comprehensive walk trip assignment process is therefore not part of the recommendation for either the first or second generation model.

Bicycle assignments might be more feasible; however, the benefit of assigning bicycle trips relative to its cost and difficulty is questionable. Origin-destination information for many bicycle trips would already be available from the activity pattern models, and this could be sufficient for such purposes as bikeway planning. Although recreational bicycle tours without destinations (i.e. people out for a bike ride beginning and ending at home) would not be modeled, there would be no way to determine the routes even if they were simulated. A comprehensive bicycle trip assignment process is therefore not part of the recommendation for either the first or second generation model.

Economic and Land Use Modeling

Integrated land use-transportation modeling

The development of an integrated land use-transportation model is a significant undertaking. Since it will require significant time and resources to develop, and it will be desirable to have a transportation model in place relatively soon (prior to the development of a fully integrated model system), it is recommended to consider such a model only as part of the second generation model. For the first generation model, existing CMAP procedures for creating demographic and land use forecasts will continue to be used.

It has long been recognized that land development and location choice decisions for businesses and individuals are affected by transportation accessibility, and therefore by transportation policy and investment decisions. Treating the locations of residences and businesses as exogenous inputs to travel models ignores these effects and therefore reduces accuracy in model results.

Integrated land use-transportation models have been developed to address this issue. While these models, including DELTA, UrbanSim, PECAS, TRANUS, and locally developed procedures, have been around for some time, there is still a wide variety of models used throughout the country (and the world). These models attempt to consider the various factors that affect land development and location choice decisions, including not only transportation accessibility, but also locations of existing development, economic effects, and development constraints.

The cadre agrees that the transportation profession has not reached any consensus on best integrated land use-transportation modeling practice. Bowman (2006) prepared a report for the Atlanta Regional Commission that discussed issues surrounding the existing models, including reliance on equilibrium assumptions despite observation that urban systems exist in dynamic disequilibrium, or attempts to model dynamic disequilibrium without the data and scientific methods to statistically capture the phenomenon. Because of these issues, and the fact that research will continue on these types of models during the period before the second generation model development will commence, no recommendation for a specific land use-transportation modeling procedure is included here. The following discussion focuses on some of the concepts relevant to these types of models and options for CMAP to consider when the time comes to make a decision how to proceed.

Equilibrium vs. disequilibrium models. There is no consensus on whether models that assume an equilibrium state at each particular point in time or those that assume disequilibrium but lack the data to predict disequilibrium states are more appropriate. Bowman (2006) notes that the disequilibrium models, including DELTA, UrbanSim, PECAS, and TRANUS are more widely used.

Endogenous incorporation of real estate prices in the model. Nearly all modern models include this feature, and it seems clear that consideration of prices is critical in understanding land development.

Economic input/output (I/O) modeling. Many of the existing models, including PECAS and TRANUS, are based on this concept while others, such as UrbanSim, are not. Again, there is no consensus on whether I/O models are superior to other model types in terms of accuracy in forecasts or policy analysis. CMAP should consider whether they have a preference for I/O type models during the period when the first generation model is under development.

Dynamic population development. A few models evolve population and household composition over time. This is consistent with the desire to have this feature in the second generation model. However, it is not a requirement for the integrated land use-transportation model to have this feature for it to be incorporated into the CMAP second generation travel model. CMAP should examine whether the population evolution features in existing land use models will be sufficient for the desired policy analysis purposes.

Freight modeling. Models such as PECAS and TRANUS explicitly include goods movement modeling, consistent with the economic I/O framework. The freight modeling needs for CMAP discussed above may or may not be sufficiently addressed by these existing procedures.

Integration with travel model. There is already experience in using some models, such as PECAS and UrbanSim, in connection with activity based travel models.

Economic modeling

The issue of economic modeling is closely related to both the land use and freight modeling topics. The existing land use models consider, to varying degrees, economic factors in location and land development choices. Models which are based on an economic I/O framework essentially would serve as economic models for the region (although effects external to the region could still be considered when using such a model). If a model such as PECAS was to be implemented by CMAP, it would be unnecessary to implement a separate economic modeling process.

CMAP has considered a modeling approach known as a Computable General Equilibrium (CGE) model (vom Hofe and Kawamura 2008), for which a case study has been performed in Chicago. The model is intended to be used to analyze transportation projects that affect freight movements and was tested for two hypothetical projects (a set of projects, in one case) in the Chicago region. It uses a social accounting matrix (SAM) for 1997 for the Chicago region, where industries are aggregated into 25 sectors. The SAM identifies seven transportation-related producing sectors and seven transportation-related commodities: Transportation Equipment, Water-Rail-Air Transportation, Freight Transportation and Warehousing, In-house Transportation, Miscellaneous Transportation Services, and Airports/Ports.

The cadre did not develop recommendations for specific economic models, or even if a separate economic model should be incorporated into the modeling process. Clearly such a model, if developed, would not be part of the first generation model.

Bibliography

Bowman, J.L. (2006). A Review of the Literature on the Application and Development of Land Use Models. ARC Modeling Assistance and Support 2006, Task 14: Land Use Model Research, Draft 2. Prepared for Atlanta Regional Commission.

Fischer, M.J., M.L. Outwater, L. Cheng, D.N. Ahanotu, and R. Calix (2005). An Innovative Framework for Modeling Freight Transportation in Los Angeles County. Working paper, available from the authors.

Hunt, J.D. and K.J. Stefan (2007). Tour-based microsimulation of urban commercial movements. Transportation Research Part B 41 (2007) 981–1013.

Samimi, A., A. Mohammadian, and K. Kawamura (2010). A Behavioral Freight Movement Microsimulation Model: Method and Data. Paper submitted for publication in the journal of Transportation Letters.

U.S. Environmental Protection Agency (2010). Technical Guidance on the Use of MOVES2010 for Emission Inventory Preparation in State Implementation Plans and Transportation Conformity. EPA-420-B-10-023.

vom Hofe, R. and Kawamura, K. (2008). Economic Impact Analysis of Goods Movement Using a CGE Modeling Framework: The Case of Chicago. Chiang Mai University Journal of Social Science And Humanities, (2008) Vol. 2(2).

Chapter 2: Data Development

Principal Author

John L. Bowman, Ph. D.
Bowman Research and Consulting
28 Beals Street, Brookline, MA 02446 USA
+1-617-232-8189 John_L_Bowman@alum.mit.edu http://JBowman.net

Other contributors

Kermit Wies, CMAP

Tom Rossi, Cambridge Systematics

Peter Vovsha, Parsons Brinckerhoff

Konstadinos Goulias, University of California, Santa Barbara

Ram Pendyala, Arizona State University

Summary

Introduction

This chapter addresses the data development requirements for implementing and using advanced models to inform planning and policy decisions in the Chicago region. In developing this plan we took the view that for CMAP, an agency charged with a regional planning function, data is part of the essential infrastructure. Data are expensive and slow to develop, and should also be systematically maintained, enhanced and archived.

The scope of data needs covers models of personal travel, as well as transport required for the delivery of goods and services. It addresses all modes of travel, motorized and non-motorized. It includes the models of demand for transport, models of traffic flows on the transport networks, and model systems that capture the interactions between demand for transport and the supply of transport facilities. It deals with all traffic within the region, including that which might flow into, out of and through the region.

All travel models require several types of data for a few main purposes, as summarized in Table 1. **Surveys**, such as the recent household activity and travel survey, are the main data source for ***model development***, providing information about travel behavior and preferences that is incorporated into the models' structure and equations. **Spatial data**, representing the nature of the activity and travel environment at the time of the surveys, is also needed for model development. Spatial attributes, such as the distribution of employment, and travel times and costs, are used to explain the behavior observed in the surveys. Typically, **network data** are also needed, in order to generate the spatial data related to travel times and costs.

Developed models require ***calibration and validation***. Calibration involves making simple adjustments to the model equations, causing them to match observed aggregate phenomena, such as traffic volumes, while preserving their ability to represent behavior. Once calibrated, the models should also be validated by applying them under conditions other than those used for development and calibration, so that the models' ability to realistically represent behavior can be demonstrated. Calibration and validation both require similar kinds of data. These include **traffic data**, to which the model predictions can be compared, as well as **spatial and network data** that serve as model inputs representing the conditions at the time that the traffic data were collected.

Of course the ultimate objective is to use the models for ***prediction***. This again requires **spatial and network data** that serve as model inputs. Inputs can represent forecasts or 'what if' scenarios for attributes such as population and employment, as well as conditions arising from policies or programs, such as transit route development or road pricing.

Table 1 Major types of data and their use for modeling

TYPES OF DATA	MODELING USES		
	Model development	***Calibration and validation***	***Prediction***
Surveys	✓		
Spatial data	✓ (survey years)	✓ (calibration and validation years)	✓ (projections)
Network data	✓ (survey years)	✓ (calibration and validation years)	✓ (projections)
Traffic data		✓	

CMAP's data status and development needs

The assessment of CMAP's current data status and data development needs followed the following basic sequence. First, Cadre member John Bowman developed a preliminary generic list of advanced models and the data they require, based on his own experience and research. This was adjusted over time as Cadre member Tom Rossi developed information related to CMAP's specific policy and modeling needs, and as feedback was received from other Cadre members. Then Bowman conducted research into CMAP's current data holdings, culminating in a four day residency at the CMAP offices, during which CMAP staff explained the data and needs from their perspective, and demonstrated how they develop and use it. From this, Bowman further modified the data needs, documented the current status, and developed recommendations for data development. Subsequently, minor adjustments were made in response to Cadre members' comments. The results are summarized below for each major category of data.

Survey Data

Several types of survey data are needed for model development, including not only household diary data, but also surveys of visitors to the region, transit riders, airport travelers, businesses conducting commercial movements in the region, and freight facilities. Supplemental survey or substitute data are also needed, including GPS and/or transponder traces, census data, commodity flow data and stated preference surveys. Among these, only the 2007 household survey, the US census microsample (collected by the US Census Bureau), and the Transearch commodity flow data are current enough to be useful. The rest are either obsolete or have never been collected. Therefore, substantial attention and resources are needed to establish and maintain a robust survey data program.

We recommend the implementation of a continuous survey program for most of the needed survey data types. CMAP has taken preliminary steps to implement such a program for the region's household travel survey, rather than waiting ten to twenty years to conduct another large scale survey. A continuous program offers the potential of eliminating survey start-up barriers, establishing and maintaining high survey quality, and including survey data as an important component of the region's data resource. It can also provide information about temporal changes that would be useful in the development of dynamic models, especially if the program includes a panel data component, in which some respondents are interviewed repeatedly over time. For this reason we recommend that serious consideration be given to incorporating a panel component into the household and business enterprise surveys. It is advisable to also ask retrospective questions in the survey to capture information about major change events.

CMAP should carefully consider how to divide survey management and administration responsibilities between CMAP and survey research firms. There may be reasons to increase CMAP management and staff roles in the survey program with the shift to a continuous program. For example, the previously mentioned benefits of a continuous survey might be best achieved if permanent staff was to conduct the survey interviews, and CMAP might prefer that they be CMAP employees. Thus, CMAP may need to increase staff level in order to support the program. However, the preliminary estimate of the out-of-pocket survey data budget assumes a steady state continuous program with a split between in-house and contracted responsibilities that is similar to the most recent household travel survey. It may be advisable to phase the continuous survey program in, one or two survey types at a time, so that adequate management and staff attention can be devoted to getting them successfully off the ground. It may also be advisable to complete larger samples per year during the first year or two of each survey type, so that an adequate sample size is available for timely model development, according to the model development schedule.

Spatial Data

Spatial data of many types are needed. These include GIS point and boundary data; spatially defined inventories of employment, residences, education facilities, parking, hotel and tourist attractions, and freight and commercial transport facilities. These data should be maintained continuously. They are used to calibrate newly developed or updated models, so snapshots are required representing survey data years, and for base years chosen for model development and calibration. They can also be used to provide accurate current input information when models are used to predict short-term responses assuming fixed spatial attributes.

One of the important developments in advanced travel forecasting models, which CMAP should adopt, is the use of disaggregate spatial data, made feasible by advancing computer technology. Regions began going down this path by subdividing their region into more and smaller traffic analysis zones (TAZ), or by dividing the TAZ into subzones. More recently, land use and transport models have been developed that take advantage of parcel data made available by tax assessors and detailed geographic information about employment locations, transport network and other spatial attributes. The advantage of disaggregate spatial data is that it greatly improves the measurement of factors that significantly affect travel decisions, such as the ease of walk access to transit and the attractiveness of activity destinations. This increases the quality of the model equations and subsequent predictions.

CMAP has already begun compiling an extensive parcel database and is therefore poised well to implement an activity-based (AB) model with parcel-level spatial resolution. Effective manual means are in place for collecting the data from the counties and combining it into one database. CMAP has staff capability and vision to automate the collection and archiving of this type of data. CMAP has begun to develop a file of parcel attributes associated with the household survey year, specifically to support activity-based model development. This will confirm whether or not parcel-level modeling is a viable option in the short-term and, if it is, it will let them proceed more quickly into model development.

CMAP will need to develop efficient procedures for generating future scenario parcel data for all attributes used in modeling. This has two aspects. The first aspect involves generating forecasts of spatially more aggregate values, such as households and employment, that are likely to come from economic and demographic forecasts or scenarios, or possibly land use models. Some land use models even generate forecasts at a parcel level. The second aspect involves implementing efficient procedures to disaggregate the needed data to the parcel point level. These procedures may use detailed parcel data from a base year, adjusting them to match the generated aggregate values. They could also involve the use of scenario development software that provides tools for quickly generating spatial attributes at a detailed geographic level to represent one or more development scenarios.

Network Data

Detailed representations of the transport networks are needed for developing, calibrating and applying models. They should be maintained continuously, and archived periodically. As with spatial data, the network data is used to calibrate newly developed or updated models, and might eventually be used to develop dynamic models, so snapshots are required representing survey data years, base years chosen for model development and calibration, and preferably all years.

CMAP currently has a method of maintaining the coded road and transit networks that enables rapid implementation of scenarios representing various years and facility development options for purposes of modeling. CMAP should devote additional resources to an ongoing program of improvement to these road and transit network representations, preserving the ability to rapidly implement scenarios for modeling. Speaking broadly, the improvements would provide greater spatial and temporal detail, enabling CMAP to implement advanced demand and traffic models that use this information. Special attention needs to be given to enhancing the network data architecture so that it can flexibly represent within-day changes in network attributes. For example, networks that can represent pricing of various types that change throughout the day are needed in order to model the effects of pricing policies.

Other specific objectives of the improvement program include transition to a geometrically accurate road network; inclusion of most or all streets; inclusion of attributes that affect non-motorized modes and ADA accessibility; representation of data elements required for dynamic traffic assignment and microsimulation (e.g., traffic controls); addition of parking and sidewalk features; representation of tolls and price attributes; and freight-related network attributes.

The priority and timing of the enhancements needs to be coordinated with the specific needs of the model development schedule. Enhancements that will probably be desirable for the first generation AB model development project include time-dependent tolls and pricing attributes, time-dependent truck restrictions (time, height, weight and noise), time-dependent transit networks with capacity information, facilities available only for cycling and/or walking, and on-street parking facilities (with capacity and price information).

Traffic Data

CMAP currently has access to extensive gateway count data, as well as highway link counts and speeds, and transit counts. It is likely that these can provide much of the data needed for calibration of the activity-based model system with static traffic assignment. However, archived data on traffic queues and turning movements, needed for dynamic traffic assignment and simulation, are not currently available. Also, detailed count data needed for the calibration of advanced models, such as counts by time of day, may not be fully archived. Some of these counts may be generated by the real-time ITS data collection program, but not adequately archived. Other counts may not be generated now at all.

A development effort should be carried out to define a set of traffic data to be used for calibration of a new AB model system, and for developing DTA and traffic microsimulation models. As part of the project, a set of calibration data should be developed and archived for the model development base years. In the course of developing the data, procedures should be implemented that make subsequent development of calibration data for other years as straightforward as possible. This includes the automatic cleaning and periodic archiving of data from the ITS monitoring systems that would otherwise be automatically deleted. At the same time, important gaps in the data should be identified and steps taken to improve the data collection process so that the gaps get filled. It is likely that this will require a new ongoing budget commitment.

Guide to the detailed report

The detailed data development report that follows contains three major sections: (1) Model Features and their Data Requirements, (2) Data Resource Requirements, and (3) Data Development Program. The first section examines data requirements from the perspective of needed models. Some data is needed by more than one model, so the description of that data requirement is repeated for each model that needs it. This is done intentionally, so that the reader can get a complete picture for a model of interest without skipping to other sections of the report.

The second section is written from the perspective of the data archivist. It organizes the requirements into several basic data categories (surveys, spatial, network, traffic and projections) listing and describing each required data resource only once, and explaining how the data are developed, maintained and used. Since this section is essentially a reorganization of the first section, there is substantial redundancy of data descriptions between the two sections.

The third section begins to outline a program for the ongoing development and maintenance of the data resources identified in the first two sections.

Model Features and Their Data Requirements

This section examines data requirements from the perspective of the needed models. For each model, this section briefly describes the model, and then it lists the model's required data resources, describing each resource and explaining how it is used. In some cases, where the model requirements are not clear enough to determine the data requirements, potential data requirements are identified and the uncertainty is noted.

Personal travel demand models

Demand for personal surface transport in the region is modeled using a variety of models for various aspects of the demand, including travel of residents within the region, travel that crosses the region's borders, and travel of visitors within the region.

Internal travel

These movements include all personal trips conducted in a given day (or days) by all residents of the region who remain within the region. They exclude trips made in commercial vehicles for work purposes. They include trips by motor vehicle, transit and non-motorized modes. The model generates a synthetic population and stochastically microsimulates the movements of each member of the population. Most of the existing activity-based (AB) models also simulate long-term choices, including at least work location, school location and auto ownership, although some of these outcomes are beginning to be simulated by land use models and supplied as inputs to the AB model.

Household diary survey

The survey interviews a stratified random sample of several thousand households. Sampled households provide information on the activities and trips of all household members for a 24-hour or longer time period, including origin, destination, purpose, travel mode, time and other details. The most complete surveys of this type also collect information about at-home activities, so that the AB model can capture relationships between at-home and out-of-home activities, and the factors causing people to stay home.

The data are used to estimate the model. They are also expanded to represent the total population, for purposes of model calibration.

A special sub-population of this survey may include a sample of non-institutionalized group quarters residents, especially military base personnel, college students and older residents living in group quarters. The sample frame for these persons may come from institutions to which these special sub-populations relate.

An additional component of this or a separate survey is a stated preference exercise in which the respondent provides choices in response to hypothetical scenarios. The resulting data are used to estimate values of time for the various categories of households, persons, tours and trips, for use in the model.

Another additional component of this survey is a GPS trace of person movements. Ideally, GPS traces would be integral parts of the survey data collection. Survey respondents would wear the GPS device on their assigned travel day(s). Then, the survey data collection software would use the GPS trace information to ease the respondent burden and enhance the completeness of their responses. The GPS results might also be used to develop models of route choice.

Existing AB models use cross-sectional household survey data and do not model long-term transitions dynamically. If the household survey includes a panel component, it may be possible to enhance the long-term models to represent household transitions over time, based on observed transitions among the households providing repeated responses over many years.

Geographic attributes of independent variables

Examples include employment, residence, school enrollment, parking and transport network. Versions of these data representing the regional situation in the year of the household survey are used to estimate the model. Versions representing the base year for model forecasting are used to calibrate the model. Versions representing each forecast year are needed for applying the model to generate forecasts.

Transportation level of service

Mode-specific zone-to-zone transport times, costs and other attributes. Versions of these data are used to estimate and calibrate the model. For model application, they are generated by the network models and are the primary information passing from the network model to the demand models in the integrated model system.

Control totals and microsample distributions

Census summary tables (ACS, CTPP) and micro sample of households and persons. These data are used to generate a synthetic population. They are also used for model calibration and application (e.g. commute mode and trip length distributions)

Zonal, sub-regional and/or regional control and validation values of household and person attributes. Some of these data are used to control the population synthesis process so that the synthetic population accurately represents the true population in certain important characteristics, such as income distribution. Others of these data can be used to determine how well the synthetic population matches the true population on certain uncontrolled characteristics. The control data are needed and used during both model calibration and application.

External travel

These movements include travel of residents making trips by auto or rail to locations outside the region, as well as travel of non-residents into and through the region.

In one modeling approach, the demand is predicted at gateway locations by special-purpose mode-specific trip generation models, and distributed in the region (and to other gateways) using singly constrained destination choice models. In a more advanced modeling approach, this travel would be modeled using a multi-state tour-based model developed from multi-state survey data. Data requirements of the simpler modeling approach are included here.

Highway survey

Count and survey of travelers entering and leaving region in autos and small trucks. The survey collects data from drivers entering and leaving the region. Questions are asked that enable estimation of models to predict the number, purpose, timing, and destination (origin) within the region of drivers entering (leaving) the region. The counts enable expansion of the survey to represent all trips entering and leaving at each gateway for purposes of model calibration, making them useful for model calibration and validation.

Intercity passenger rail survey

On-board count and survey of travelers entering and leaving region by rail. The survey collects data from rail passengers entering and leaving the region. Questions are asked that enable estimation of models to predict the number, purpose, timing, destination (origin) within the region, deboarding (boarding) station, and mode of egress (access) of passengers entering (leaving) the region by rail. The counts enable expansion of the survey to represent all trips entering and leaving at each gateway for purposes of model calibration.

Geographic attributes of independent variables

Examples include employment, residence, school enrollment, parking and transport network. Versions of these data representing the regional situation in the year of the surveys are used to estimate the model. Versions representing the base year for model forecasting are used to calibrate the model. Versions representing each forecast year are needed for applying the model to generate forecasts.

Geographic attributes of hotel room availability and tourist attractions. These data are used to estimate the destination choice model for non-residents visiting the region. Versions representing other years are needed and used for model calibration and application.

Transportation level of service

Mode-specific zone-to-zone transport times, costs and other attributes. Versions of these data are used to estimate and calibrate the model. For model application, they are generated by the network models and are the primary information passing from the network model to the demand models in the integrated travel model system.

Visitor travel

Tour-based model of within-region surface travel demand by leisure and business visitors. These movements include all trips conducted within the region in a given day by leisure and business visitors to the region. They include trips by motor vehicle, transit and non-motorized modes. The model stochastically microsimulates the movements of each visitor.

Visitor diary survey

The survey interviews a stratified random sample of regional business and leisure visitors. Sample is drawn from hotel guests and/or gateway intercepts. Sampled persons provide information on the nature and duration of the visit, and details of activities and trips for one 24-hour time period. The data are used to estimate a model similar to the model for residents. They are also expanded to represent the total visitor population, for purposes of model calibration and application.

An additional component of this or a separate survey is a stated preference exercise in which the respondent provides choices in response to hypothetical scenarios. The resulting data are used to estimate values of time for the various categories of visitors, for use in the model.

Geographic attributes of independent variables,

Examples include employment, residence, school enrollment, parking, transport network, hotel room availability and tourist attractions. Versions of these data representing the regional situation in the year of the visitor survey are used to estimate the model. Versions representing the base year for model forecasting are used to calibrate the model. Versions representing each forecast year are needed for applying the model to generate forecasts.

Transportation level of service

Mode-specific zone-to-zone transport times, costs and other attributes. Versions of these data are used to estimate and calibrate the model. For model application, they are generated by the network models and are the primary information passing from the network model to the demand models in the integrated model system.

Airport travel

These movements are made by persons traveling by air to and from the region's airports. The demand is predicted at each airport by special-purpose trip generation models, split by mode using a mode choice model, and distributed in the region using a singly constrained destination choice model. It would also be possible to model time-of-day choice for these trips.

If there are special event venues, such as sports arenas or convention centers, for which it is important to model the special transport system effects, then each of these venues could similarly have a special model of passenger travel to and from the venue. Developing such a model may require similar special surveys and counts, although some of this information might be available from the venue operator. Alternatively, it might be possible to adapt the airport model to accommodate this purpose.

Survey of airport travelers

The survey collects data from air passengers entering and leaving the region via the airport. Questions are asked that enable estimation of models to predict the trip purpose, destination (origin) within the region, mode of egress (access), and time of day. Similar surveys could be used to collect information about travel to and from major event venues in the region, for estimating origin and travel mode models.

Counts of air passengers

These data are supplied by the airport authority and used for estimation of the trip generation and timing model. Similar counts could be supplied by major event venue operators for calibrating trip generation models for major event venues.

Geographic attributes of independent variables

Examples include employment, residence, school enrollment, parking, neighborhood and transport network. Versions of these data representing the regional situation in the year of the surveys are used to estimate the model. Versions representing the base year for model forecasting are used to calibrate the model. Versions representing each forecast year are needed for applying the model to generate forecasts.

Additional geographic attributes of hotel room availability and tourist attractions are also needed. These data are used to estimate the destination choice model for non-residents visiting the region. Other versions are needed and used for model calibration and application.

Transportation level of service

Mode-specific zone-to-zone transport times, costs and other attributes. Versions of these data are used to estimate and calibrate the model. For model application, they are generated by the network models and are the primary information passing from the network model to the demand models in the integrated model system.

Freight models

Demand for freight and commercial surface transport in the region can be categorized into two main categories. The first category consists primarily of **intra-regional commerce**, in which the types of movements consist primarily of intra-regional tours made by goods and service vehicles. This category of demand can be modeled via tour-based models, similar to the activity-based or tour-based models being developed and used to model passenger transport within a region.

The second category of demand consists primarily of **inter-regional goods shipments**. These shipments often involve inter-modal transfers, shipment reconfigurations and warehousing at various points along a 'logistics chain' running from point of production to point of use. These movements can be national, continental and global in scope. Since Chicago is a continental crossroads of this kind of shipment, it may be desirable for CMAP to build models that have this large geographic scope, so that they can capture the effect of regional conditions on flows into, out of and through the region. This effect might be modeled via macro-scale economic and/or commodity flow models that are sensitive to regional conditions. It might also be modeled by large-scale freight models that take into consideration economic factors as well as the logistical choices made by shippers and carriers. For some types of industries, such as package delivery, the first and/or last movements in the chain consist of tour-based movements to and from regional distribution centers. These portions of the logistics chain can be modeled using the tour-based models described above for intra-regional commerce.

The design of freight and commercial models for the Chicago region is currently not clear. One or more projects are needed to clarify the design of these models and their data requirements. The following description of possible models and their associated data requirements represent a preliminary attempt to identify the data requirements that CMAP might encounter as it moves forward with freight modeling, but is by no means definitive. Indeed, some of the data described here may prove to be too difficult to collect, because of the difficulty of collecting data from vehicles en route or from freight companies.

Sources of material in this section include Donnelly (2005); Hunt et al. (2006); Hunt and Stefan (2007); the City of Calgary 2000 Commodity Flow Survey Report (2000); Kuzmyak (2009); Meheboob Ishani (2003); Wies, et al (2009); LLA Consultancy Ltd (2004); Fischer et al (2005).

Internal commercial vehicle tours

These movements consist of tours within the region conducted by commercial enterprises of various types and sizes for themselves, or by transportation and warehousing firms for their customers. They also include tours made by entities that dispatch vehicles to cover an area or to travel road links, rather than handle a specific shipment; examples include newspaper delivery, garbage and recycling pickup.

Business enterprise survey

Survey of business enterprises conducting commercial movements in the region. The commercial movement survey interviews a stratified random sample of several thousand business enterprises in the region. The interview is analogous to the household travel diary. Sampled establishments provide information on the movements of their entire fleet over a 24 hour time period, including origin, destination, purpose, fleet and commodity information. The data are used to estimate the model. They are also expanded by industry, size and location to represent the total population of commercial enterprises, for purposes of model calibration.

An additional component of this survey is a stated preference exercise in which the establishment provides choices in response to hypothetical scenarios. The resulting data are used to estimate values of time for the various categories of establishment, for use in the model.

Another additional component of this survey is a GPS trace of commercial vehicle movements. Ideally, GPS traces would be integral parts of the survey data collection. GPS devices would trace vehicle movements on the assigned travel day(s). Then, the survey data collection software would use the GPS trace information to ease the respondent burden and enhance the completeness of their responses. The GPS results might also be used to develop models of route choice.

The cost of this survey per respondent will be comparable to, but larger than for the household survey because it takes more time to contact the right person at the firm and to help them complete the survey correctly.

In order to conduct the survey of business enterprises it is necessary to develop a list of the establishments. This list is used as a sample frame (master list) to randomly select establishments for the interview. It is also used to determine how many firms are in each of the model's categories (by type and number of employees), so that the resulting sample can be expanded to represent the entire population. There isn't necessarily a single good data source for this, although the US Census Bureau's employment data products (such as OnTheMap) may be useful.

Geographic attributes of independent variables

Employment, residence and school enrollment are needed. As with the personal travel model, these data are used for model estimation, calibration and application.

Transportation level of service

Mode-specific zone-to-zone transport times, costs and other attributes. As with the personal travel model, these data are used for model estimation, calibration and application.

External goods shipment

Models of inter-regional goods shipments are needed. In Chicago, much regional traffic supports national and international flow of goods. Modeling of these flows in and out of the region's transportation gateways, and within the region, is of high importance. However, this is a new area of modeling for Chicago and, more broadly, for public sector agencies in North America. This paragraph lays out a potential description of the types of models involved. The demand may be generated at gateway locations in the form of mode-specific commodity flows by national or continental models representing economic processes and logistical decisions of shippers and carriers. These models determine the volume, mode, and timing of goods movements into and out of the region's gateway locations (air, rail and water terminals, as well as major truck entry points, exit points and distribution centers). Flows to and from the rail, air and water terminals that involve truck movements for the first or final link in the logistics chain are converted to truck vehicle flows and distributed from those locations via tour-based models for some types of goods, and trip-based models for others. Truck flows through the region are also distributed via trip-based models. The exact data requirements of these models depend on the exact nature of the models to be developed. However, most models chosen to handle inter-regional goods shipment will require the following types of data.

Commodity flows

Public and private commodity flow data are needed. These data are used for freight model development and calibration. They include the proprietary Transearch database (by IHS Global Insight), and possibly additional publicly available data, such as the US Census Bureau Commodity Flow Survey and the FHWA Freight Analysis Framework (FAF) data, among others.

The commodity flows must also be developed for forecast years, and these can be purchased from Global Insight as part of the Transearch database. However, limitations in their forecasting procedures may require the development of procedures or models to modify the forecast commodity flows in response to regional conditions, or may require the development of complete alternatives to the Transearch forecasts. For example, the Transearch forecasts assume that base year mode splits are maintained in the forecast year for each commodity type.

Port, intermodal, terminal and warehouse facilities

These data include the location and transport-related functions of all major freight and commercial traffic facilities in the region for all modes of transport. They provide the basis for counts and surveys, as well as for development and application of freight and commercial transport models representing movements within the region of inter-regional shipments.

Counts and surveys at port, intermodal, terminal and warehouse facilities are also needed. The survey collects data needed to estimate the number of truck movements to and from each port and intermodal facility, and the destinations of those tours and trips (within the region and external gateways). The counts enable expansion of the survey to represent all trips to and from the facility for purposes of model calibration.

Highway points of entry

Counts and survey of medium and large trucks entering and leaving the region. The survey collects data about trucks entering and leaving the region. Questions are asked that enable estimation of models to predict the number and destination (origin) within the region of trucks entering (leaving) the region. The counts enable expansion of the survey to represent all trips entering and leaving at each gateway for purposes of model calibration.

Transponder records, GPS traces and/or dispatch records of truck movements will facilitate the survey of these vehicles. Transponder records, GPS traces and/or dispatch records supplied by carriers and their dispatchers could provide an alternative to surveys conducted at ports, inter-modal facilities and gateways. If a representative sample of vehicles was captured via transponder (used for electronic tolling) or GPS (used for carrier fleet tracking) or dispatch records, then these records could be used to estimate the number of truck movements and their destinations. Counts would still be needed to expand the sample data for purposes of model calibration.

Geographic attributes of independent variables

Examples include employment, residence and school enrollment. As with the personal travel model, these data are used for model estimation, calibration and application.

Transportation level of service

Mode-specific zone-to-zone transport times, costs and other attributes. As with the personal travel model, these data are used for model estimation, calibration and validation.

Freight network databases

National rail, truck, water and air networks. If CMAP decides to invest in freight transport models with national or continental scope, then it will be necessary to develop a representation of the freight network on the same scale spanning all modes included in the model.

Network models

These models route, on representations of the region's surface transport networks, the trips predicted by the demand models. They also supply needed information to the demand models about travel times, costs and network-based policies.

Regional static (highway) traffic assignment model

This is a model like the existing traffic assignment model, although it would likely be enhanced.

Coded road network for static traffic assignment

These data consist of a network of links and nodes specified in a format required by the roadway traffic assignment model, representing the network for one or more base years as well as for each forecast year scenario. They are used for model development, calibration and application. The network is tested, maintained and enhanced in an ongoing program.

Traffic count data

These data consist of link counts by time of day for links on important cordons identified in the region, and for a sample of additional links in the network. They are used for calibrating and validating the assignment models, and the integrated demand-supply model systems. Counts are categorized by vehicle class, direction and time of day.

Traffic link speed data

These data consist of automatically collected link speeds by time of day for links on important corridors identified in the region, and for a sample of additional links in the network. They are used for validating the assignment model.

Traffic OD travel time data

These data consist of survey-based, trace vehicle or GPS-based OD travel times by time of day for ODs using important corridors in the region, and for a sample of other OD pairs in the region. They are used for validating the assignment model.

Pedestrian count data

These data consist of pedestrian link counts by time of day for links in important pedestrian zones and corridors identified in the region. They are used for validating the demand and assignment models.

Bicycle count data

These data consist of bicycle link counts by time of day for links on important bicycle corridors identified in the region. They are used for validating the demand and assignment models.

Truck route data

Network-based information on truck routes, restrictions and tolls for model estimation and application is included in this data.

Regional transit assignment model

This is a model like the existing transit assignment model, although it would likely be enhanced.

Coded transit network for transit assignment

These data consist of a network of links, nodes, lines, routes and schedules specified in a format required by the transit assignment model, representing the network for one or more base years as well as for each forecast year scenario. They are used for model development, calibration and application. The network is tested, maintained and enhanced in an ongoing program.

Transit count data

These data consist of boardings, alightings, vehicle frequencies and loads by time of day for the important stations, lines and links in the network. They are used for validating the transit assignment model. They may also be useful for calibrating the transit submode choice model. Much of these data should be available from the transit operators, although some additional counts may be needed.

On-board transit survey

These data consist of origin, destination and route information collected from transit riders. They would be used to adjust transit ridership to account for transfers unaccounted for by the transit operators' automatically collected count data.

Regional dynamic (highway) traffic assignment model

Coded road network for dynamic traffic assignment

These data consist of a network of links, nodes and traffic controls specified in a format required by the dynamic traffic assignment model, representing the network for one or more base years as well as for each forecast year scenario. They are used for model development, calibration and application. The network data are tested, maintained and enhanced in an ongoing program. Important enhancements over the static assignment network include a more complete and physically accurate representation of links, intersections and loading points, as well as inclusion of traffic controls and pricing features, and representation of link/node pricing and restrictions by time of day.

Traffic count data

These data consist of link counts by time of day for links on important cordons identified in the region, and for a sample of additional links in the network. They are used for calibrating and validating the assignment models, and the integrated demand-supply model systems. Counts are categorized by vehicle class, direction and time of day. An important distinction between this and the count data required for static assignment is that counts of turning movements and counts by detailed time-of-day period are more important.

Traffic link speed data

These data consist of automatically collected link speeds by time of day for links on important corridors identified in the region, and for a sample of additional links in the network. They are used for validating the assignment model.

Traffic OD travel time data

These data consist of survey-based, trace vehicle or GPS-based OD travel times by time of day for ODs using important corridors in the region, and for a sample of other OD pairs in the region. They are used for validating the assignment model.

Traffic queue data

These data consist of queue size and time delays by turning movement and outgoing link at important roadway nodes and choke points that experience queuing in the region, as well as at a sample of nodes throughout the region. They are used for validating the queuing phenomena captured by the dynamic traffic assignment and microsimulation models.

Dynamic traffic microsimulation model

Coded road network for dynamic traffic microsimulation

These data consist of a network of links, lanes, nodes and traffic controls specified in a format required by the microsimulation model, representing the network for one or more base years as well as for each forecast scenario. They are used for model development, calibration and application. The network data are tested, maintained and enhanced in an ongoing program. Important enhancements over the dynamic traffic assignment network include a more complete and physically accurate representation of links, lanes, intersections, loading points, and traffic controls for the subregion being modeled.

Traffic count data

These data consist of link counts by time of day for links on important cordons identified in the region, and for a sample of additional links in the network. They are used for calibrating and validating the assignment models, and the integrated demand-supply model systems. Counts are categorized by vehicle class, direction and time of day. An important distinction between this and the count data required for DTA is that the counts are required for the more detailed network representation.

Traffic link speed data

These data consist of automatically collected link speeds by time of day for links on important corridors identified in the region, and for a sample of additional links in the network. They are used for validating the assignment model. An important distinction between this and the speed data required for DTA is that the speeds are required for the more detailed network representation.

Traffic OD travel time data
These data consist of survey-based, trace vehicle or GPS-based OD travel times by time of day for ODs using important corridors in the region, and for a sample of other OD pairs in the region. They are used for validating the assignment model.

Traffic queue data
These data consist of queue size and time delays by turning movement and outgoing link at important roadway nodes and choke points that experience queuing in the region, as well as at a sample of nodes throughout the region. They are used for validating the queuing phenomena captured by the dynamic traffic assignment and microsimulation models. An important distinction between this and the data required for DTA is that they are required for the more detailed network representation.

Bicycle and pedestrian models

Coded non-motorized network for demand model
These data consist of an enhancement to the roadway network to include a representation of links and nodes available only to bicycles and/or pedestrians, as well as bicycle and pedestrian restrictions on all other links. These basic data are needed even if there is no special assignment of bicycles and pedestrians, for purposes of producing shortest path travel time skims for use in the demand models.

Coded road network for bicycle and pedestrian route choice and assignment
These data consist of an enhancement to the non-motorized network to include a representation of link and node attributes needed by the bicycle and/or pedestrian route choice and network models, such as sidewalk availability and width, bike lane availability and width, and grade information. These data are used for model development, calibration and application.

Model systems of demand and supply
The models described above are not very useful by themselves. They need to be integrated into model systems that deal comprehensively with demand and supply, capturing the interactions on the transport network of all aspects of personal and commercial travel. The integrated model systems have additional data requirements associated primarily with model system calibration and validation. However, the categories of required data are already included in the above sections for each of the component models.

Data Resource Requirements
Many of the needed data resources satisfy the needs of more than one model. This section lists and describes the data resources themselves, and explains how they are developed, maintained and used. In some cases, where the model requirements are not clear enough to determine the data requirements, potential data requirements are identified and the uncertainty is noted.

Survey Data
These data are collected periodically, or as part of a continuous survey data program, and used for estimation of model coefficients, as well as for model calibration. This section also includes survey-like data that are purchased from external sources. At the end of this section, Table 1 lists the needed survey data and, for each item, identifies the models that need it and for what purposes. The purposes

include (1) model development and estimation of coefficients used in the model formulas; (2) model calibration and base year validation; (3) model application for forecasting and scenarios.

Household diary

The survey interviews a stratified random sample of several thousand households. Sampled households provide information on the activities and trips of all household members for a 24-hour or longer time period, including origin, destination, purpose, travel mode, time and other details. The most complete surveys of this type also collect information about at-home activities, so that the AB model can capture relationships between at-home and out-of-home activities, and the factors causing people to stay home.

The data are used to estimate the model. They are also expanded to represent the total population, for purposes of model calibration.

A special sub-population of this survey may include a sample of non-institutionalized group quarters residents, especially college students and older residents living in group quarters. The sample frame for these persons may come from institutions to which these special sub-populations relate.

An additional component of this or a separate survey is a stated preference exercise in which the respondent provides choices in response to hypothetical scenarios. The resulting data are used to estimate values of time for the various categories of households, persons, tours and trips, for use in the model.

Another additional component of this survey is a GPS trace of person movements. Ideally, GPS traces would be integral parts of the survey data collection. Survey respondents would wear the GPS device on their assigned travel day(s). Then, the survey data collection software would use the GPS trace information to ease the respondent burden and enhance the completeness of their responses. The GPS results might also be used to develop models of route choice.

Existing AB models use cross-sectional household survey data and do not model long-term transitions dynamically. If the household survey includes a panel component, it may be possible to eventually enhance the long-term models to represent household transitions over time, based on observed transitions among the households providing repeated responses over many years. A "second best" alternative would be to ask retrospective questions in the survey to capture information about major life change events.

External travel

Surveys of travelers entering and leaving region. These surveys are needed if CMAP chooses to use models that distribute border-crossing travel. If CMAP chooses to use a multi-state tour-based model instead, then multi-state data similar to that described for the regional AB model would be required. If these surveys prove too difficult or costly to conduct, then they might be replaced by GPS traces of movements that would enable accurate estimation of these flows as background demand for calibration of the overall travel forecasting model system.

Autos and small trucks

Surveys of travelers entering and leaving region in autos and small trucks. The survey collects data from drivers entering and leaving the region. Questions are asked that enable estimation of models to predict the number, purpose, timing, and destination (origin) of drivers entering, leaving and passing through the region. It is unlikely that these travelers can be recruited via roadside intercept. (However, for an example of such a survey, see Parsons Transportation Group, Inc., 2001). Rather, the recruitment may need to use vehicle license numbers (e.g., see Alliance Transportation Group, Inc., 2009) or, for some gateways, electronic toll records. Alternatively, it might be possible to obtain samples from phone vendors of GPS traces for users of GPS-equipped mobile phones, simplifying the survey itself, or eliminating the need for a survey altogether.

Medium and large trucks

The survey collects data about trucks entering and leaving the region. Questions are asked that enable estimation of models to predict the number and destination (origin) within the region of trucks entering (leaving) the region. The need for and exact nature of this data is uncertain. It would most likely be needed if CMAP chose a reasonably simple modeling approach for inter-regional freight that focused primarily on the portion of those movements occurring within the region.

Transponder records, GPS traces and/or dispatch records of truck movements will facilitate this survey. Transponder records, GPS traces and/or dispatch records supplied by carriers and their dispatchers could provide an alternative to surveys conducted at ports, inter-modal facilities and gateways. If a representative sample of vehicles was captured via transponder (used for electronic tolling) or GPS (used for carrier fleet tracking) or dispatch records, then these records could be used to estimate the number of truck movements and their destinations. Counts would still be needed to expand the sample data for purposes of model calibration.

Passenger rail

On-board survey of travelers entering and leaving region by rail. The survey collects data from rail passengers entering and leaving the region. Questions are asked that enable estimation of models to predict the number, purpose, timing, destination (origin) within the region, alighting (boarding) station, and mode of egress (access) of passengers entering (leaving) the region by rail. The counts enable expansion of the survey to represent all trips entering and leaving at each gateway for purposes of model calibration.

Visitor diary

The survey interviews a stratified random sample of regional business and leisure visitors. Sample is drawn from hotel guests and/or gateway intercepts. Sampled persons provide information on the nature and duration of the visit, and details of activities and trips for one 24-hour time period (or more). The data are used to estimate a model similar to the model for residents. They are also expanded to represent the total visitor population, for purposes of model calibration and application.

An additional component of this or a separate survey is a stated preference exercise in which the respondent provides choices in response to hypothetical scenarios. The resulting data are used to estimate values of time for the various categories of visitors, for use in the model.

Airport travelers

The survey collects data from air passengers entering and leaving the region via the airport. Questions are asked that enable estimation of models to predict the trip purpose, destination (origin) within the region, mode of egress (access), and time of day. Similar surveys could be used to collect information about travel to and from major event venues in the region, for estimating origin and travel mode models.

Business enterprises

Survey of business enterprises conducting commercial movements in the region. The commercial movement survey interviews a stratified random sample of several thousand business enterprises in the region. The interview is analogous to the household travel diary, and is used to develop a tour-based model of freight and commercial movements within the region. The need for this data depends on finalizing a model development strategy that includes such a model. Sampled establishments provide information on the movements of their entire fleet over a 24 hour time period, including origin, destination, purpose, fleet and commodity information. The data are used to estimate the model. They are also expanded by industry, size and location to represent the total population of commercial enterprises, for purposes of model calibration.

An additional component of this survey is a stated preference exercise in which the establishment provides choices in response to hypothetical scenarios. The resulting data are used to estimate values of time for the various categories of establishment, for use in the model.

Another additional component of this survey is a GPS trace of commercial vehicle movements. Ideally, GPS traces would be integral parts of the survey data collection. GPS devices would trace vehicle movements on the assigned travel day(s). Then, the survey data collection software would use the GPS trace information to ease the respondent burden and enhance the completeness of their responses. The GPS results might also be used to develop models of route choice.

The cost of this survey per respondent will be comparable to, but larger than for the household survey because it takes more time to contact the right person at the firm and to help them complete the survey correctly.

In order to conduct the survey of business enterprises it is necessary to develop a list of the establishments. This list is used as a sample frame (master list) to randomly select establishments for the interview. It is also used to determine how many firms are in each of the model's categories (by type and number of employees), so that the resulting sample can be expanded to represent the entire population. There isn't necessarily a single good data source for this, although the US Census Bureau's employment data products (such as OnTheMap) may be useful.

Port, intermodal, terminal and warehouse facilities

The survey collects data needed to estimate the number of truck movements to and from each port and intermodal facility, and the destinations of those tours and trips (within the region and external gateways). The need for and exact nature of this data is uncertain, although it would likely be needed for both the intra-regional tour-based model and the inter-regional logistics chain model.

Commodity flow data

These data are used for freight model development and calibration. They include the proprietary Transearch database (by IHS Global Insight, see Global Insight, 2009), or possibly publicly available data, such as the US Census Bureau Commodity Flow Survey and the FHWA Freight Analysis Framework (FAF) data, among others.

On-board transit survey

These data consist of origin, destination and route information collected from transit riders. They would be used to adjust transit ridership to account for transfers unaccounted for by the transit operators' automatically collected count data, and to provide additional information needed for calibrating the transit submode choice model.

Table 2: Survey Data Usage in Advanced Modeling

	Survey Data Resource	Model development/estimation	Model calibration/validation	Model application/forecasting
1	Household diary survey	AB passenger demand	AB passenger demand	
2	Survey of travelers entering and leaving region in autos and small trucks	External highway passenger demand	External highway passenger demand	
3	On-board survey of travelers entering and leaving region by rail	External rail passenger demand	External rail passenger demand	
4	Visitor diary survey	Leisure & business visitor demand	Leisure & business visitor demand	Leisure & business visitor demand
5	Survey of airport travelers	Airport (and event) demand	Airport (and event) demand	
6	Survey of business enterprises conducting commercial movements in the region	Freight and commercial demand	Freight and commercial demand	
7	Survey of port, intermodal, terminal and warehouse facilities	Freight demand	Freight demand	
8	Survey of medium and large trucks entering and leaving the region	Freight demand	Freight demand	
9	Transponder records, GPS traces and/or dispatch records of truck movements	Freight and commercial demand	Freight and commercial demand	
10	Census summary tables (ACS, CTPP) and micro sample of households and persons		AB passenger demand	AB passenger demand
11	Public and private commodity flow data	Freight demand	Freight demand	
12	On-board transit survey		Transit assignment AB passenger demand	

Spatial Data

These data are maintained continuously. They are used to calibrate newly developed or updated models, and might eventually be used to develop dynamic models of activity opportunities, so snapshots are required representing survey data years, base years chosen for model development and calibration, and preferably for all years. They can also be used to provide accurate current input information when models are used to predict short-term responses assuming fixed spatial attributes.

In the case of point or parcel inventories, not only are data values (e.g. employment) associated with the parcel itself needed, but also GIS techniques are required to generate corresponding values of each data item for buffer zones of various sizes, such as ¼ mile radius, surrounding the point or parcel centroid.

For examples of spatial data requirements for activity-based models in other regions, see Parsons Brinckerhoff et al, 2007 (for Jerusalem), Bowman and Bradley, 2008 (for Sacramento) and Bradley et al, 2008 (for PSRC Seattle).

At the end of this section, Table 2 lists the needed spatial inventory data and, for each item, identifies the models that need it and for what purposes.

GIS point and boundary files

A point file identifies the geographic location of each parcel centroid (or zone centroid, etc.) The other spatial attributes, such as households and employment, are associated with these centroids.

A boundary file can be used to aggregate to a given level any spatial data that is available at a more disaggregate geography. For example, utility hookup points might be aggregated to the parcel level using a parcel boundary map layer. A boundary file can also be used for displaying attributes stored with the associated points.

Residential/housing inventory

These data include estimates of households, non-institutional group quarters residents and possibly housing units by type. Of special interest are inventories of university dormitory and off-campus housing, and other non-institutional group quarters, because their high concentration at few locations can substantially affect model performance.

Employment/jobs inventory

These data include estimates of jobs (at the work location) by industry (such as Industrial, Retail, Office, Food, Education, Medical, Services, Government) or type of employment. Information about the number of establishments by industry and size (floor space and number of employees) may also be useful for the validation of land use models.

Schools/education facilities inventory

These data include estimates of school enrollment (at the classroom location) by grade level, and perhaps by private vs. public school.

Parking inventory

These data include estimates of the off street parking capacity and price by type of parking (daily public, hourly public, private employer, etc). It can also include an inventory of on-street parking, although this would be better maintained as link attributes on a full street network.

Hotel and tourist attraction inventory

These data include locations of hotels and major tourist attractions, as well as estimates of hotel rooms by hotel type, and estimates of visitor capacity of major tourist attractions.

Inventory of port, intermodal, terminal and warehouse facilities

These data identify the location and transport-related functions of all major freight and commercial traffic facilities in the region for all modes of transport. They provide the basis for counts and surveys, as well as for development and application of freight and commercial transport models.

Aerial Photography

Aerial photography can be used to assign land use codes, and also to validate the residential, employment and other inventory estimates.

Transportation Network GIS attributes

These data consist of point/parcel attributes representing attributes of the full street transportation network, usually in buffer zones surrounding parcel centroids. Example attributes include the number of dead-end links, the number of 3-link intersections, the number of 4+ link intersections, the number of on-street parking spaces, the percentage of links with sidewalk, the number of, and distance to, the transit stops of each transit type.

Table 3: Spatial Data Usage Advanced Modeling

	Spatial Data Resource	Model development/estimation and calibration/validation
1	GIS point and boundary files	All point/parcel-based models
2	Residential/housing inventory	AB passenger demand External highway passenger demand Leisure & business visitor demand External rail passenger demand Airport (and event) demand Commercial demand
3	Employment/jobs inventory	AB passenger demand External highway passenger demand Leisure & business visitor demand External rail passenger demand Airport (and event) demand
4	Schools/education facilities inventory	AB passenger demand External highway passenger demand Leisure & business visitor demand External rail passenger demand Airport (and event) demand Commercial demand
5	Parking inventory	AB passenger demand External highway passenger demand Leisure & business visitor demand External rail passenger demand Airport (and event) demand

	Spatial Data Resource	Model development/estimation and calibration/validation
6	Hotel and tourist attraction inventory	Leisure & business visitor demand External rail passenger demand Airport (and event) demand Commercial demand
7	Inventory of port, intermodal, terminal and warehouse facilities	Freight demand
8	Aerial Photography	Used to develop other inventory data
9	Transportation Network GIS attributes (regenerate annually)	AB passenger demand External highway passenger demand Leisure & business visitor demand External rail passenger demand Airport (and event) demand Commercial demand

Transport Network Representations

These are detailed representations of the transport networks that are needed for developing, calibrating and applying the network and demand models. They are maintained continuously, and archived periodically.

At the end of this section, Table 3 lists the needed network data and, for each item, identifies the models that need it and for what purposes.

Full street network

This network includes a representation of all links and nodes representing all passageways available to all motorized and non-motorized modes. It also includes link and node attributes relevant for measuring availability, attractiveness and impedance to passage and parking by all of those modes. It is used as a source of information for the coded networks used by models and/or might be used directly by some models. Versions of the network are established representing important time points in the past, for purposes of model development and calibration. Versions of the network may also be established for forecasting purposes, each representing a specific policy scenario and point in time.

Coded transit network for transit assignment

These data consist of a network of links, nodes, lines, routes and schedules specified in a format required by the transit assignment model, representing the network for one or more base years as well as for each forecast year scenario. The network is tested, maintained and enhanced in an ongoing program.

Coded road network for static traffic assignment

These data consist of a network of links and nodes specified in a format required by the roadway traffic assignment model, representing the network for one or more base years as well as for each forecast year scenario. The network is tested, maintained and enhanced in an ongoing program.

Coded road network for dynamic traffic assignment

These data consist of a network of links, nodes and traffic controls specified in a format required by the dynamic traffic assignment model, representing the network for one or more base years as well as for each forecast year scenario. The network data are tested, maintained and enhanced in an ongoing program. Important enhancements over the static assignment network include a more complete and physically accurate representation of links, intersections and loading points; inclusion of traffic controls and pricing features, and representation of link/node pricing and restrictions by time of day.

Coded road network for dynamic traffic microsimulation

These data consist of a network of links, lanes, nodes and traffic controls specified in a format required by the microsimulation model, representing the network for one or more base years as well as for each forecast scenario. The network data are tested, maintained and enhanced in an ongoing program. Important enhancements over the dynamic traffic assignment network include a more complete and physically accurate representation of links, lanes, intersections, loading points, and traffic controls for the subregion being modeled.

Coded bicycle and pedestrian road network and additional attributes

These data consist of an enhancement to the roadway network (full street or traffic assignment) to include a representation of links and nodes available only to bicycles and/or pedestrians, as well as bicycle and pedestrian restrictions on all other links. They also include link and node attributes representing factors that affect bicycle and pedestrian choices, such as sidewalk availability and width, bike lane availability and width, and grade information. The data are needed for the development and application of mode choice models, and potentially also for bicycle and/or pedestrian route choice and network models.

Truck, rail and water freight and commercial route data

These data provide network-based information on routes, restrictions and tolls for truck trip assignment model application.

National rail, truck, water and air networks

If CMAP decides to invest in freight transport models with national or continental scope, then it will be necessary to develop a representation of the freight network on the same scale spanning all modes included in the model.

Table 4: Table of Network Data Usage in Advanced Models

<table>
<tr><th></th><th>Network Data Resource</th><th>Model development/estimation</th><th>Model calibration/validation</th></tr>
<tr><td>1</td><td>Full street network</td><td>Used to develop point/parcel and buffer attributes</td><td>Used to develop point/parcel and buffer attributes</td></tr>
<tr><td>2</td><td>Coded transit network for transit assignment</td><td rowspan="6">Used to generate the spatial attributes of travel time and cost for model estimation</td><td>Transit assignment</td></tr>
<tr><td>3</td><td>Coded road network for static traffic assignment</td><td>Static traffic assignment</td></tr>
<tr><td>4</td><td>Coded road network for dynamic traffic assignment</td><td>Dynamic traffic assignment</td></tr>
<tr><td>5</td><td>Coded road network for dynamic traffic microsimulation</td><td>Traffic microsimulation</td></tr>
<tr><td>6</td><td>Coded bicycle and pedestrian road network and additional attributes</td><td>AB passenger demand
Static traffic assignment
Bike/ped assignment</td></tr>
<tr><td>7</td><td>Truck, rail and water freight and commercial route data</td><td>Static traffic assignment</td></tr>
</table>

	Network Data Resource	Model development/estimation	Model calibration/validation
8	National rail, truck, water and air networks		Inter-regional freight

Traffic data

These data provide information about the speeds and flows of various types on the transportation network, and are used to calibrate newly developed or updated models. They are collected in a continuous count program.

At the end of this section, Table 4 lists the needed traffic data and, for each item, identifies the models that need it and for what purposes.

Traffic count data

These data consist of link counts by time of day for links on important cordons identified in the region, and for a sample of additional links in the network. They are used for calibrating and validating the assignment models, and the integrated demand-supply model systems. Counts are categorized by vehicle class, direction and time of day. For freight modeling, counts must be categorized by detailed freight and commercial vehicle class. For DTA, counts of turning movements and counts by detailed time-of-day period are important. For traffic microsimulation, counts are required for the more detailed network representation used in microsimulation.

Traffic link speed data

These data consist of automatically collected link speeds by time of day for links on important corridors identified in the region, and for a sample of additional links in the network. They are used for validating the assignment model.

Traffic OD travel time data

These data consist of survey-based, trace vehicle or GPS-based OD travel times by time of day for ODs using important corridors in the region, and for a sample of other OD pairs in the region. They are used for validating the assignment model.

Traffic queue data

These data consist of queue size and time delays by turning movement and outgoing link at important roadway nodes and choke points that experience queuing in the region, as well as at a sample of nodes throughout the region. They are used for validating the queuing phenomena captured by the dynamic traffic assignment and microsimulation models.

Transit count data

These data consist of boardings, alightings, vehicle frequencies and loads by time of day for the important stations, lines and links in the network. They are used for validating the transit assignment model. Much of these data should be available from the transit operators, although some additional counts may be needed.

Pedestrian count data

These data consist of pedestrian link counts by time of day for links in important pedestrian zones and corridors identified in the region. They are used for validating the demand and assignment models.

Bicycle count data

These data consist of bicycle link counts by time of day for links on important bicycle corridors identified in the region. They are used for validating the demand and assignment models.

Mode-specific zone-to-zone transport times, costs and other attributes

Versions of these data are used to estimate, calibrate and apply demand models that depend on estimates of network impedances of various types. They are generated by the network models and are the primary information passing from the network models to the demand models in the integrated demand-supply model systems.

Counts of travelers entering and leaving the region by road and rail

Counts enable expansion of surveys or GPS traces of the same traveler type, to represent all trips entering and leaving at each gateway, making them useful for model calibration and validation.

Counts of air passengers

These data are supplied by the airport authority and used for estimation of the trip generation and timing model. Similar counts could be supplied by major event venue operators for calibrating trip generation models for major events venues.

Counts at ports, intermodal facilities, terminals and warehouses

The counts enable expansion of the corresponding survey to represent all trips to and from the facility for purposes of model calibration.

Count of medium and large trucks entering and leaving the region

The counts enable expansion of the corresponding survey or gps traces to represent all trips entering and leaving at each gateway for purposes of model calibration.

Table 5: Traffic Data Usage in Advanced Models

	Traffic Data Resource	Model development/estimation	Model calibration/validation
1	Traffic count data		Static traffic assignment Dynamic traffic assignment Traffic microsimulation
2	Traffic link speed data		Static traffic assignment Dynamic traffic assignment Traffic microsimulation
3	Traffic OD travel time data		Static traffic assignment Dynamic traffic assignment Traffic microsimulation
4	Traffic queue data		Dynamic traffic assignment Traffic microsimulation
5	Transit count data		Transit assignment
6	Pedestrian count data		AB passenger demand Static traffic assignment

	Traffic Data Resource	Model development/estimation	Model calibration/validation
7	Bicycle count data		AB passenger demand Static traffic assignment
8	Mode-specific zone-to-zone transport times, costs and other attributes	AB passenger demand External highway passenger demand Leisure & business visitor demand External rail passenger demand Airport (and event) demand Freight and commercial demand	AB passenger demand External highway passenger demand Leisure & business visitor demand External rail passenger demand Airport (and event) demand Freight and commercial demand
9	Counts of travelers entering and leaving region by road and rail		External highway passenger demand External rail passenger demand
10	Counts of air passengers		Airport demand
11	Counts at ports, intermodal facilities, terminals and warehouses		Freight demand
12	Count of medium and large trucks entering and leaving the region		Freight demand

Data projections

Data projections (forecasts and assumptions) are used to generate the inputs needed by the transport models for forecasting.

At the end of this section, Table 5 lists the needed data projections and, for each item, identifies the models that need it for model application and forecasting.

Economic and demographic projections

Economic and demographic projections are generated at the most disaggregate level possible. It would be expected that in many cases reasonable forecasts might be made for transport analysis zones or aggregations of them, but in some cases they might be only at the county or regional level. However, the advent of parcel-based scenario development software tools makes it possible to develop assumption-based scenarios at the point or parcel level.

These projections might be developed by (a) regional economists, using the tools at their disposal, (b) simple land use models or scenario development tools that rely heavily on assumptions for alternative future scenarios, or (c) sophisticated integrated economic and land use models that represent the complex dynamic economic interactions of regional economies.

Households and group quarters

These projections are used to control the residential population synthesis process so that the synthetic population accurately represents the true population in certain important characteristics, such as income distribution. Others of these data can be used to determine how well the synthetic population matches the true population in certain uncontrolled characteristics. The control data are needed and used during both model calibration and application. The synthetic population is used as the population for which the activity-based residential travel demand model predicts travel demand of residents. The projections are also used to then generate the zonal, parcel or point-level estimates of households and group quarters residents, which are used as explanatory variables in the demand models. These are fundamental outputs of the economic forecasts, which should take into consideration immigration and emigration activity, as well as natural trends within the in-place population.

Leisure and business visitors

These projections are used to control the visitor population synthesis process so that the synthetic population accurately represents the population in certain important characteristics, such as travel purpose, mode of entry, destination sub-region, duration, and value of time. The synthetic population is used by the tour-based visitor travel demand model. These data might be derived from the economic forecasts of business activity.

Inter-regional freight flows

These projections provide mode-specific (and in some cases route-specific) OD estimates of commodity flows into, out of and through the region. They are used by the freight demand and network models. These data can be purchased from IHS Global Insight as part of their Transearch product. However, limitations in their forecasting procedures may require the development of procedures or models to modify the forecast commodity flows in response to regional conditions, or may require the development of complete alternatives to the Transearch forecasts. For example, the Transearch forecasts assume that base year mode splits are maintained in the forecast year for each commodity type.

Inter-regional passenger flows

These projections provide estimates of the daytime travel into, out of and through the region by non-residents living close enough to the region to make day-trips into (and back out of) the region for work or other purposes. They are used as control totals for the models predicting the resulting flows within the region. These projections are likely to be assumption-based modifications of estimated base year flows, taking into account expectations for economic and demographic changes in regions bordering the CMAP region.

Business establishment population and employment

These projections are used to control the business establishment population synthesis procedure so that the synthetic population accurately represents the true population in certain important characteristics, such as industry and firm size. The synthetic population is used as the population for which the tour-based commercial travel demand model predicts travel demand of business establishments. The projections are also used to generate the zonal, parcel or point-level estimates of employment, which are used as explanatory variables in the demand models. These projections should be derived from economic forecasts of business activity.

Schools/education facilities

These projections correspond directly to the school enrollment inventory, identifying the projected locations of all school classroom facilities and the projected enrollment (at the classroom location) by grade level, and perhaps by private vs. public school. These are used to populate the point/parcel/zonal values of school enrollment used as explanatory variables in the travel demand models.

Spatial data projections

The spatial data projections are analogous to the base year inventories, but represent scenarios used as input for transport model forecasting. The economic and demographic projections provide a base for these more detailed projections, and procedures must be developed and used to generate the needed spatial detail. In the simplest case, the projections represent proportional projections of the base year inventory data, adjusted to match the more aggregate economic and demographic projections. Scenario development software, such as iPlace3s, or land use forecasting model systems, such as PECAS or UrbanSim, may also be employed. In either case, supplemental automated procedures may be needed to get the data into the form needed by the travel forecasting models.

GIS point and boundary files

This file identifies the geographic location of each point or parcel centroid. Projection into the future may involve subdividing large parcels to represent development that is modeled or assumed.

Residential/housing projections

These forecasts include estimates of households, non-institutional group quarters residents and possibly housing units by type. Of special interest are inventories of university dormitory and off-campus housing, and other non-institutional group quarters, because their high concentration at few locations can substantially affect model performance. Zone-level estimates may come from the population synthesizer, with supplemental models used to distribute households within each zone according to base-year distribution of housing stock by type, as modified for future year scenarios.

Employment/jobs projections

These data include estimates of jobs (at the work location) by industry (such as Industrial, Retail, Office, Food, Education, Medical, Services, Government) or type of employment. These may be provided by an economic and land use model, or may be generated from a synthetic population of businesses, in a procedure analogous to the residential and housing projections.

Schools/education facilities

These projections correspond directly to the school enrollment inventory, identifying the projected locations of all school classroom facilities and the projected enrollment (at the classroom location) by grade level, and perhaps by private vs. public school. These are used to populate the point/parcel/zonal values of school enrollment used as explanatory variables in the travel demand models. These projections should be derived from and consistent with the household projections.

Parking availability and cost projections

These projections may be generated by parking demand and supply submodels in an economic and land use model, by freestanding parking supply and demand models, or by scenario development software.

Hotel and tourist attractions

These data include locations of hotels and major tourist attractions, as well as estimates of hotel rooms by hotel type, and estimates of visitor capacity of major tourist attractions. These can be modified to match economic and land use model results, or via scenario development software.

Port, intermodal, terminal and warehouse facilities

These projections identify the projected locations and transport-related functions of all major freight and commercial traffic facilities in the region for all modes of transport. They can be scenario-based modifications of base-year inventories.

Land use projections

Land use models and/or scenario analysis tools can be used to modify the distribution of land uses from the base year.

Transportation network GIS attributes

These projections are derived from projected networks using the same tools that derive them for base year networks. The values consist of parcel centroid attributes representing attributes of the full street transportation network, usually in buffer zones surrounding the parcel centroid. Example attributes include the number of dead-end links, the number of 3-link intersections, the number of 4+ link intersections, the number of on-street parking spaces, the percentage of links with sidewalk, the number of transit stops, and the distance to the nearest transit stop of each transit type.

Transportation network projections

Transportation network projections represent future networks associated with infrastructure development and policy scenarios. They are identical in format to the base year network representations described above. They can be maintained as part of a master network representation that includes date-stamps of attributes that are expected to change over time. Alternatively they could be derived as modifications to the base year networks using the same tools that are used to build and maintain the base year networks themselves. See the above section on Transportation network representations for detailed descriptions of the various networks.

Table 6: Projections in model application and forecasting

		Data Resource	Model application/forecasting
Economic and demographic projections	1	Households and group quarters	AB passenger demand External highway and rail passenger demand Visitor and airport demand Commercial demand
	2	Leisure and business visitors	Leisure & business visitor demand
	3	Inter-regional freight flows	Freight demand
	4	Inter-regional passenger flows	External highway and rail passenger demand
	5	Business establishment population and employment	AB passenger demand External highway and rail passenger demand Visitor and airport demand Freight and commercial demand
	6	Schools/education facilities	AB passenger demand External highway and rail passenger demand Visitor and airport demand Commercial demand
Spatial Data Projections	1	GIS point and boundary files	
	2	Residential/housing projections	AB passenger demand External highway and rail passenger demand Visitor and airport demand Commercial demand
	3	Employment/jobs projections	AB passenger demand External highway and rail passenger demand Visitor and airport demand Freight and Commercial demand
	4	Schools/education facilities projections	AB passenger demand External highway and rail passenger demand Visitor and airport demand Commercial demand

		Data Resource	Model application/forecasting
	5	Parking availability and cost projections	AB passenger demand External highway passenger demand Visitor and airport demand
	6	Hotel and tourist attractions	Visitor and airport demand Commercial demand
	7	Port, intermodal, terminal and warehouse facilities	Freight demand
	8	Land use projections	AB passenger demand External highway and rail passenger demand Visitor and airport demand Commercial demand
	9	Transportation network GIS attributes	AB passenger demand External highway and rail passenger demand Visitor and airport demand Commercial demand
Transportation network projections	1		Static traffic assignment Transit assignment Dynamic traffic assignment Traffic microsimulation

Data Development Program

This section begins to outline a program for the ongoing development and maintenance of the data resources needed for an advancing modeling practice at CMAP. For each broad data category, this section takes stock of the current status of relevant data at CMAP and describes a recommended data program. For the survey data program, where substantial new cash outlays are likely, preliminary cost estimates are provided. The other data programs consist primarily (although not entirely) of increased investment in, or reallocation of, staff time. These estimates are not provided, but they may be substantial.

Survey Data

Table 6 lists the required data resources and briefly notes the current status of each one.

Table 7: Survey Data Current Status

	Data Resource	Status	Year developed	Notes
1	Household diary survey	In active use	2007	10,600 households, small GPS and SP
2	Survey of travelers entering and leaving region in autos and small trucks	Ignored	2000	Survey execution failed.
3	On-board survey of travelers entering and leaving region by rail	Obsolete	80s and 90s	Amtrak
4	Visitor diary survey	obsolete	1970s and 80s	Usually related to special events
5	Survey of airport travelers	Obsolete	70s and 80s	O'Hare and Midway
6	Survey of business enterprises conducting commercial movements in the region	No		1986 truck survey
7	Survey of port, intermodal, terminal and warehouse facilities	No		
8	Survey of medium and large trucks entering and leaving the region	Ignored	2000	Survey execution failed.
9	Transponder records, GPS traces and/or dispatch records of truck movements	No		
10	Census summary tables (ACS, CTPP) and micro sample of households and persons	In active use		PUMA used in population synthesis
11	Public and private commodity flow data	In active use		Transearch
12	Transit on-board survey			Limited; transit operators provide these

Among the survey data resources identified above as important for advanced model development at CMAP, only the 2007 household survey and the US census microsample (collected by the US Census Bureau) are current enough to be useful. The rest are either obsolete or have never been collected. Therefore substantial attention and resources are needed to establish and maintain a robust survey data program.

CMAP has taken preliminary steps to implement a continuous survey program for the region's household travel survey, rather than waiting ten to twenty years to conduct another large scale survey. An ongoing continuous program offers the potential of eliminating survey start-up barriers, establishing and maintaining high survey quality, and including survey data as an important component of the region's data resource. It can also provide information about temporal changes that would be useful in the development of dynamic models, especially if the program includes a panel data component, in which some respondents are interviewed repeatedly over time. It is advisable to also ask retrospective questions in the survey to capture information about major change events. An important aspect of the continuous survey program is the regular archiving of temporally coincident non-survey data, such as spatial attributes, network models, travel times and costs, and fuel costs. This should include all attributes that are necessary for model development. Without the accompanying non-survey data, the survey data would not be useful for model development.

Most of the other needed survey data resources are also well suited to a continuous survey program approach. These include group quarters and visitor activity and travel diaries; airport ground access/egress survey; transit, commuter rail and inter-city rail on-board surveys; and business establishment, port, intermodal, terminal and warehouse surveys (for freight and commercial travel behavior). The survey methods should include state of the art GPS devices that enhance data quality and reduce respondent burden.

Table 7 provides a rough preliminary estimate of the budget for services performed by outside survey research firms, assuming a steady state continuous program, with a split between in-house and contracted responsibilities that is similar to the most recent household travel survey. CMAP should carefully consider how to divide survey management and administration responsibilities between CMAP and survey research firms. There may be reasons to increase CMAP management and staff roles in the survey program with the shift to a continuous program. For example, the previously mentioned benefits of a continuous survey might be best achieved if permanent staff were to conduct the survey interviews, and CMAP might prefer that they be CMAP employees. Thus, CMAP may need to increase staff level in order to support the program.

It may be advisable to phase the programs in one or two at a time, so that adequate management attention can be devoted to getting them successfully off the ground. It may also be advisable to complete more surveys per year during the first year or two so that an adequate sample size is available for timely model development, according to the model development schedule.

Table 8: Survey data program

	Data Resource	Acquisition Program	Budget
	Survey Data		
1	Household diary survey	continuous survey	$400K/yr (1600 completes * $250)
3&12	On-board survey	continuous survey	$100K/yr (1000 completes * $100)
4	Visitor diary survey	continuous survey	$125K/yr (500 completes * $250)
5	Survey of airport travelers	continuous survey	$50K/yr (300 completes * $150)
6	Survey of business enterprises conducting commercial movements in the region	continuous survey	$250K/yr (500 completes * $500)
7	Surveys at port, intermodal, terminal and warehouse facilities	continuous survey	$150K/yr (1500 completes * $100)
9 2&8	Transponder records, GPS traces and/or dispatch records of truck and passenger movements or gateway intercept surveys	negotiate acquisition from telephone carriers, carriers and/or third party logistics firms	$100K/yr
10	Census summary tables (ACS, CTPP) and micro sample of households and persons	continued acquisition from US Census Bureau	
11	Public and private commodity flow data	continued purchase of Transearch data negotiate acquisition of data from carriers and/or third party logistics firms	$100K/yr
	Supplemental stated preference surveys	as needed for special purposes, and/or attached to continuous surveys	$100K/yr
	Total		Approximately $1.4 million per year

Spatial Data

Table 8 lists the required spatial data resources and briefly notes the current status of each one.

Table 9: Spatial Data Current Status

	Data Resource	Status	Year developed	Notes
	Parcel data overall	In active use	Current	County Assessors In active development for modeling purposes
1	GIS point and boundary files	In active use	Current	Large selection
2	Residential/housing inventory	Available	Current	CMAP Indicators Project
3	Employment/jobs inventory	In active Use	Current	Employment Security
4	Schools/education facilities inventory	Available	Current	CMAP Indicators Project
5	Parking inventory	In active use	Current/Incomplete	CMAP Congestion Management Process
6	Hotel and tourist attraction inventory	Available	Current	Assembled for 2016 Olympics Bid (but doesn't provide room count)
7	Inventory of port, intermodal, terminal and warehouse facilities	Available	Current	2040 Freight Plan
8	Aerial Photography	Available	Current	
9	Transportation Network GIS attributes (regenerate annually)	Available	Current	Navteq

The spatial data should be maintained continuously. They are used to calibrate newly developed or updated models, and might eventually be used to develop dynamic models of activity opportunities, so snapshots are required representing survey data years, base years chosen for model development and calibration, and preferably for all years. They can also be used to provide accurate current input information when models are used to predict short-term responses assuming fixed spatial attributes.

CMAP has already begun compiling a parcel database and is poised well to implement an activity-based (AB) model with parcel-level spatial resolution. It has in place effective manual means of collecting the data from the counties and combining it into one database. It also has staff capability and vision to automate the collection and archiving of this type of data.

CMAP has begun to develop a file of parcel attributes associated with the household survey year, specifically to support AB model development. This will confirm whether or not parcel-level modeling is a viable option in the short-term and, if it is, it will let CMAP proceed more quickly into model development. Known needs, which can begin to be addressed immediately, are to (a) acquire parcel data for the Indiana and Illinois counties within the model's geographic scope but outside the 7-county core of the CMAP region; (b) develop automated methods of matching disaggregate employment data (which tends to be located on street centerlines) with the parcels associated with those centerline addresses; (c) develop software scripts to automatically generate data items representing attributes of buffer areas around each parcel; and (d) implement procedures for archiving snapshots of the data at regular intervals. Once the data needed for AB model development is in hand, it will be necessary to obtain additional data needed for other advanced models, such as hotel and tourist attraction inventories for visitor models, and freight facilities for freight models.

CMAP will need to develop efficient procedures for generating future scenario parcel data for all attributes used in modeling. This has two aspects. The first aspect involves generating forecasts of spatially more aggregate values, such as households and employment, that are likely to come from economic and demographic forecasts or scenarios, or possibly land use models. Some land use models even generate forecasts at a parcel level. The second aspect involves implementing efficient procedures to generate the needed data at the parcel point level. These procedures may use detailed parcel data from a base year, adjusting them to match needed aggregate control values. They could also involve the use of scenario development software that provides tools for quickly generating spatial attributes at a detailed geographic level to represent one or more development scenarios.

Network Data

Table 9 lists the required network data resources and briefly notes the current status of each one.

Table 10: Network Data Current Status

	Data Resource	Status	Year developed	Notes
1	Full street network	Available	Current	Navteq
2	Coded transit network for transit assignment	Available	Current	a.m. peak LOS only
3	Coded road network for static traffic assignment	Available	Current	8 time periods (stick network)
4	Coded road network for dynamic traffic assignment	No		
5	Coded road network for dynamic traffic microsimulation	Available	Current	Static network converted for use by TRANSIMS, but important details are missing (lanes, turning restrictions, intersection details, etc)
6	Coded bicycle and pedestrian road network and additional attributes	Available (bicycle only)	Current	Bicycle facilities and accommodations only.
7	Truck, rail and water freight and commercial route data	Available	Current	2040 Freight Plan
8	National rail, truck, water and air networks			

Detailed representations of the transport networks are needed for developing, calibrating and applying the network models. They should be maintained continuously, and archived periodically. As with spatial data, the network data is used to calibrate newly developed or updated models, and might eventually be used to develop dynamic models, so snapshots are required representing survey data years, base years chosen for model development and calibration, and preferably all years.

CMAP currently has a method of maintaining the coded road and transit networks that enables rapid implementation of scenarios representing various years and facility development options for purposes of modeling in EMME. The current network has approximately 45K links and 16K nodes.

It would be desirable for CMAP to devote additional resources to an ongoing program of improvement to these road and transit network representations. Objectives of the improvement program should include transition to a geometrically accurate road network, inclusion of most or all streets that could be included or excluded for purposes of assignment modeling, inclusion and identification of facilities available to non-motorized modes, representation of data elements required for dynamic traffic assignment and traffic microsimulation (e.g., lanes, traffic control devices, restrictions), addition of parking and sidewalk features, representation of tolls and other facility price attributes, a more complete representation of time-dependent features (ultimately to support dynamic assignment with time-sensitive attributes), and the freight-related network attributes.

It would be desirable to develop comprehensive data on sidewalk coverage and accessibility on a full street network, and derive from them parcel attributes that indicate availability of continuous pedestrian and mobility-impaired accessibility (in addition to distance) to nearest transit stops of each transit submode. Likewise, mobility-impaired accessibility for transit transfers could be incorporated into the transit network data. These could be used in the new demand and assignment models and/or used to measure accessibility in the region for mobility-impaired residents.

The priority and timing of the enhancements needs to be coordinated with the specific needs of the model development schedule. Enhancements that will probably be desirable for the first generation AB model development project include time-dependent tolls and pricing attributes, time-dependent truck restrictions (time, height, weight and noise), time-dependent transit networks with capacity information, facilities available only for cycling and/or walking, and on-street parking facilities (with capacity and price information). If this network is expanded to include all streets, then full-street data on transit stops and intersections surrounding parcel points, which is needed for AB model development, could be maintained in this database also.

Traffic Data

Table 10 lists the required network data resources and briefly notes the current status of each one.

Table 11: Traffic Data Current Status

	Data Resource	Status	Year developed	Notes
1	Traffic count data	Available	Current	state, county and municipal sources
2	Traffic link speed data	Available	Current	ITS sources; Expressways only. Archive project underway.
3	Traffic OD travel time data	no		ITS sources, development underway
4	Traffic queue data	no		
5	Transit count data	Yes	Current	Source: service board reports
6	Pedestrian count data	Obsolete	1990s	Downtown only
7	Bicycle count data			
8	Mode-specific zone-to-zone transport times, costs and other attributes	Available	Current	From skims
9	Count of travelers entering and leaving region by road and rail	Available	Current	State classification counts Amtrak ridership data
10	Counts of air passengers	Available	Current	Department of Aviation data
11	Counts at ports, intermodal facilities, terminals and warehouses	Available	Current	2040 Freight Planning work
12	Count of medium and large trucks entering and leaving the region	Available	Current	State classification counts

CMAP currently has access to extensive gateway count data, as well as highway link counts and speeds, and transit counts. It is likely that these can provide much of the data needed for calibration of the activity-based model system with static traffic assignment. However, archived data on traffic queues and turning movements, needed for dynamic traffic assignment and simulation, are not currently available. Also, detailed count data needed for the calibration of advanced models, such as counts by time of day, may not be fully archived. Some of these counts may be generated by the real-time ITS data collection program, but not adequately archived. Other counts may not be generated now at all.

A development effort should be carried out to define a set of traffic data to be used for calibration of a new AB model system, and for developing DTA and traffic microsimulation models. As part of the project, a set of calibration data should be developed and archived for the model development base years. In the course of developing the data, procedures should be implemented that make subsequent development of calibration data for other years as straightforward as possible. This includes the automatic cleaning and periodic archiving of data from the ITS monitoring systems that would otherwise be automatically deleted. At the same time, important gaps in the data should be identified and steps taken to improve the data collection process so that the gaps get filled. It is likely that this will require a new ongoing budget commitment.

Data Projections

Table 11 lists the required data projections and briefly notes the current status of each one. Eventually, models of land use and demographic change should be used in some cases to make some of these projections.

Table 12: Current Status of Data Projections

	Data Resource	Status	Year developed	Notes
	Economic and demographic projections			
1	Households and group quarters	Available	Current	CMAP scenario-based forecasting
2	Leisure and business visitors	No		Some base data available
3	Inter-regional freight flows	Available	Current	Assumed 1% per year
4	Inter-regional passenger flows	Available	Current	Assumed 1% per year
5	Business establishment population and employment	Available	Current	CMAP scenario-based forecasting
6	Schools/education facilities	Available	Current	No experience using with modeling. http://www.isbe.state.il.us/research/pdfs/public_school_enrollment.pdf
	Spatial Data Projections			
1	GIS point and boundary file	Available	Current	
2	Residential/housing projections	No		Only socioeconomic variables are projected
3	Employment/jobs projections	Yes	Current	CMAP Scenario forecasts
4	Schools/education facilities projections	No		Not at small geography
5	Parking availability and cost projections	No		
6	Hotel and tourist attractions	No		
7	Port, intermodal, terminal and warehouse facilities	No		scenario-based
8	Land use projections	No		
9	Transportation network GIS attributes	Yes	Current	Recommend transportation facilities are coded into static model networks
1	**Transportation network projections**			maintained within network database

Bibliography

Alliance Transportation Group, Inc. (2009). MAG and PAG 2008 External Travel Study Final Report (2009). Project Report.

Bradley, Bowman & Castiglione (2008). PSRC Activity Model Work Plan Report. Project Report.

Bowman, John L and Mark A Bradley (2008). SACSIM/08 Activity-Based Travel Forecasting Model for SACOG, Featuring DAYSIM—the Person Day Activity and Travel Simulator, Technical Memo Number 10--DaySim08 Documentation, Draft 9. Model Reference Documentation.

City of Calgary 2000 Commodity Flow Survey Report (2000).

Donnelly, Rick (2005). A Freight Modeling Action Plan for the Atlanta Regional Commission. Project Report.

Global Insight (2009). Development of the TRANSEARCH Database. Product Documentation.

Hunt, JD, KJ Stefan and AT Brownlee (2006). An Establishment-Based Survey of Urban Commercial Vehicle Movements in Alberta: Survey Design, Implementation and Results. Presented at the 2006 Transportation Research Board Annual Meeting.

Hunt, J.D. and K.J. Stefan (2007). Tour-based microsimulation of urban commercial movements. Transportation Research Part B 41 (2007) 981–1013.

Fischer, Michael J., Maren L. Outwater, Lihung Luke Cheng, Dike N. Ahanotu, Robert Calix (2005) An Innovative Framework for Modeling Freight Transportation in Los Angeles County. Working paper, available from the authors.

Kuzmyak, J. Richard (2009). NCHRP Synthesis 384: Forecasting Metropolitan Commercial and Freight Travel: A Synthesis of Highway Practice. Transportation Research Board, Washington, D.C.

Meheboob Ishani (2003). Edmonton Region Truck/Commodity Survey Report and Results, July 2003

Parson Brinckerhoff, Inc, NuStats and Mark Bradley Research & Consulting (2007). Development of a Tour-Based Transportation Demand Model for the Jerusalem Metropolitan Region, Inception Report: Part II, "Socio-Economic and Land-Use Data". Project Report

Parsons Transportation Group, Inc. (2001) Phoenix External Travel Survey Executive Summary. Project Report.

Survey on Hong Kong Goods Vehicle Trip Characteristics 2003, Final Report (2004). LLA Consultancy Limited. Project Report.

Wies, Kermit, Tom Murtha and Brett Baden (2009). Constructing a Policy Responsive Analysis Model for a Freight-Rich Region. Working abstract.

Chapter 3: Computing Environment

Principal Author

Peter Vovsha
Parsons Brinckerhoff
1 Penn Plaza, 2nd Floor, New York, NY 10019
212-465-5511

Vovsha@PBworld.com

Other contributors

Kermit Wies, CMAP

John Bowman, Bowman Research & Consulting

Tom Rossi, Cambridge Systematics

Konstadinos Goulias, University of California, Santa Barbara

Ram Pendyala, Arizona State University

Computational Features, Algorithms, and Types of Calculations

A travel demand model represents a sequence of calculations structured by meaningful travel dimensions. There are two major approaches for structuring a demand model – the traditional *trip-based* (frequently referred to as 4-step) and an advanced *activity-based*. Trip based 4-step models were the modeling technique developed during the 1970s for most MPOs in US, and they constitute a majority of travel models in practice even today.

In recent years, advanced Activity-Based microsimulation models (ABs) have been developed and applied in practice, currently constituting the majority of newly-developed models for large MPOs (already more than 10 such models in practice). The model features and data requirements for an advanced AB model are briefly restated in this section with an emphasis placed on their computing implications and considerations for establishing a robust programming, data management and computer hardware configuration.

Special attention is paid to the relationship between model complexity, computing capacity, and resulting runtime taking into account the size and complexity of the Chicago Metropolitan Region that is the third biggest metropolitan region in the US.

There is variety to particular AB model designs applied in practice – see [*Vovsha Bradley & Bowman, 2005; Bradley & Bowman, 2006; Davidson et al, 2007*] for comprehensive surveys of the existing ABs. Despite the variations in technical details between existing AB model systems, there are common features across all models representing core concepts of the AB model paradigm. These features include:

- A *tour-based structure* where the tour – a closed chain of trips starting and ending at a base location (usually home, school or workplace) – is used as the main unit of modeling travel. This structure preserves consistency across trips included in the same tour by travel dimensions such as destination, mode, and time of day. Further, the whole spectrum of travel dimensions (mode, destination, and time of day) related to non-home-based travel can be properly linked to home-based travel. This feature has a strong impact on the computing process.
- An *activity-based platform* that implies that modeled travel is derived within the general framework of the daily activities undertaken by households and persons. This allows for the consistency of the typological, spatial, and temporal dimensions of individual activity patterns, the substitution between in-home and out-of-home activities, modeling the duration of activities in a coherent framework with trip departure and arrival times with a fine level of temporal resolution, intra-household interactions, and other aspects pertinent to activity analyses.
- A *microsimulation* modeling technique that is applied at the fully-disaggregate level of persons and households, which converts activity and travel related choices from fractional-probability model outcomes into a series of "crisp" decisions among the discrete choices; this method of model implementation results in more realistic model outcomes, with output files that look much like actual travel/activity survey data. This feature has a primary impact on the computing process.

The combination of these three features proves to be a powerful platform for constructing operational model structures that incorporate multiple advanced techniques from the behavioral research that had been largely unused within the 4-step modeling paradigm. In the following sections, we will describe the first and third fundamental features in detail since they have many implications on the AB programmatic implementation and computing process.

The typical model structure and relevant travel dimensions modeled in each type of model are shown in **Figure 1**.

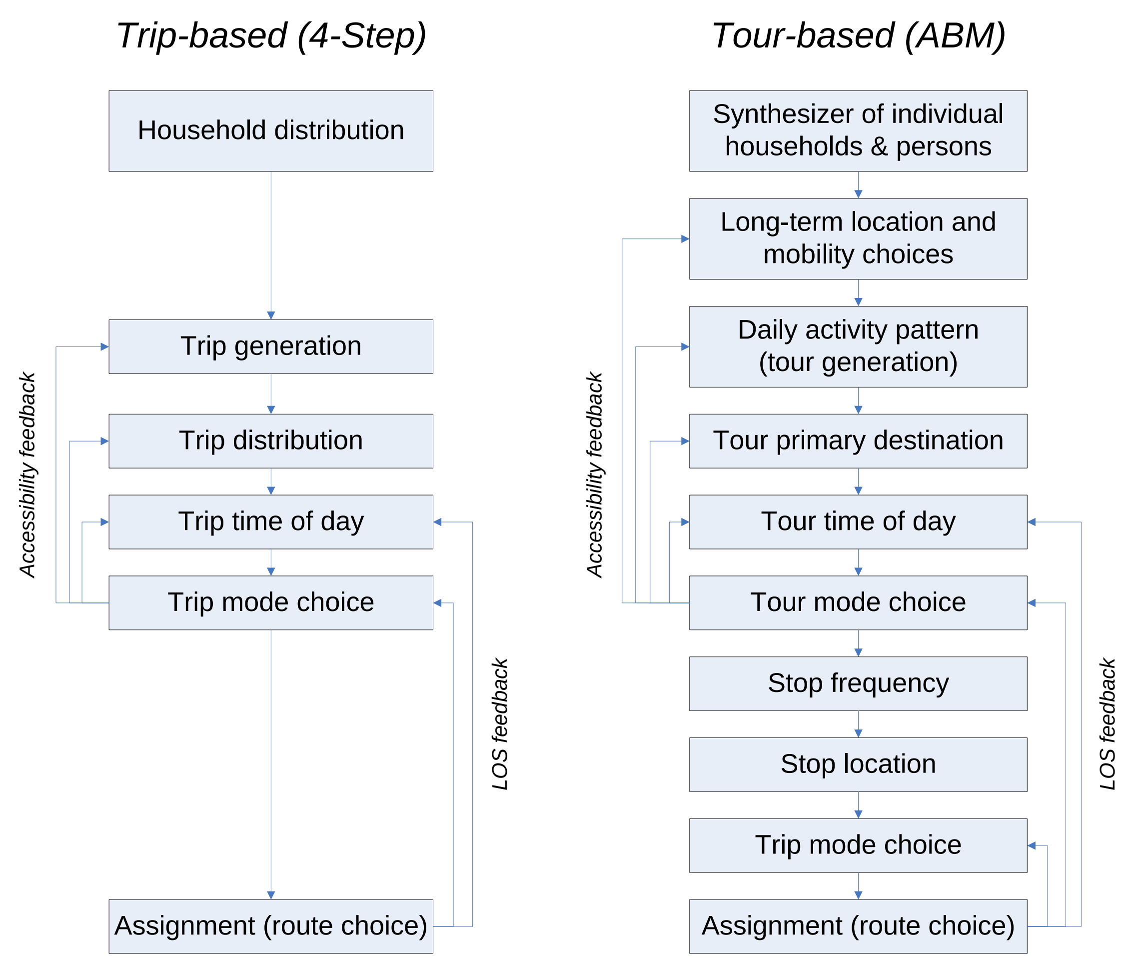


Figure 1: Typical Structure of Demand Model

A demand model represents a computerized travel simulation system where demand generation is integrated with network simulation, and equilibrium travel times and costs are sought. This equilibrium feature, which is technically implemented by feedback of level-of-service (LOS) variables representing travel time and cost to the demand generation stages, is essential. LOS variables are normally encapsulated in composite accessibility measures when are fed back to upper-level choices.

It should be noted that the model structures shown above are the most typical in practice but there are many alternative modeling paradigms that are more evolutionary in nature and a great deal of technical variation across different models.

Microsimulation

The key technical aspect that makes ABs operational is microsimulation, where a synthetic population is constructed and discrete choices are selected from probability distributions for each model component. The microsimulation modeling technique is applied at the fully-disaggregate level of persons and households. Microsimulation can be thought of as an extreme case of sample enumeration where the sample is extended to represent the whole population. This method of model implementation results in the model outcomes that look much like an actual travel/activity survey implemented for the entire population. Conversely, the outcomes of traditional trip-based travel models are only meaningful after fractional probabilities of all choice alternatives are aggregated.

The explicit modeling of travel for a full list of individual households and persons in the region does not mean that the objective of the model system design is to pinpoint each individual's behavior. The disaggregate level of detail is only maintained as a *more robust way to predict aggregate travel statistics* that are the focus of transportation planning. Better aggregate prediction is ensured by means of a more realistic and consistent representation of the underlying individual travel choices.

The microsimulation modeling paradigm is not bound just to transportation modeling. It is recognized today as an important general principle in operation research and decision-making science. There are many ways to implement a microsimulation system with different levels of aggregation and sophistication in terms of behavior of individual particles (or "agents" if the individual behavior is modeled with a certain reflection on the intelligence of each particle and independence in decision-making). In this report, microsimulation is used in a general sense of modeling individual agents rather than continuous flows or other entities. It can be a complete enumeration (1 modeled object representing 1 real object) is the most common case with ABs in practice, sample enumeration (1 modeled object representing 100 real objects) is also frequently used in ABs in practice to speed up the process of the first global iterations, or repeated enumeration (100 modeled objects represent 1 real object under different conditions) may also be used. The behavior of each modeled object can be described by deterministic rules or by probabilistic models that would require drawing from distributions. These distributions can be also created in different ways, with or without choice models.

The microsimulation used in many AB applications is a process in which the attributes of objects are modified based on environmental conditions but the fundamental rules of the behavior do not change. There is another type of microsimulation process in which "intelligent" agent simulation gives each object a few fundamental rules and then permits its behavior to evolve. This is a fundamentally different way of microsimulating the real world and is the primary subject of the artificial intelligence discipline. It certainly has an appeal for transportation models. In fact, learning and spatial cognition models may benefit from this type of approach. Certain attempts to explore this avenue have been undertaken by the Eindhoven University in the Netherlands that resulted in the ALBATROSS model. We currently reserve this approach for future research.

Microsimulation has several principal advantages over the standard fractional-probability approach in the context of travel demand modeling – see [*Vovsha Petersen & Donnelly, 2002*] for a more detailed discussion:

- *Computational savings* in storage and manipulation of large fractional-probability arrays; the microsimulation method typically requires far fewer calculations than an aggregate travel model, due to the elimination of fractional trips.
- Modeling of realistic individual *decision-making chains* with situational variables and constraints; situational variables are not known at the beginning of simulation, their values are defined in the process and correspond to the decisions modeled prior to the given one.
- Accounting for unobserved heterogeneity in travel behavior by *randomization of the model parameters*, for example, Value of Time. If model is applied in a micro-simulation fashion its parameters do not have to be deterministic, they can be drawn from a known distribution.

To illustrate the *computational savings*, consider the microsimulation example of tour destination and mode choice in **Figure 2**. On the left-hand side, a standard fractional-probability approach is shown where each destination and mode combination is assigned a fractional probability (by means of a choice model) that has to be stored as a multidimensional array. On the right-hand side, the microsimulation technique is demonstrated. It starts with a calculation of probabilities by means of the same choice model as for the fractional-probability approach. However, further on, the fractional probabilities are converted into discrete or “crisp” choices (by Monte-Carlo or other method). As the result, the choice tree for each individual tour is pruned and the multidimensional array of probabilities is converted into a list of tour records with the chosen attributes (destination and mode).

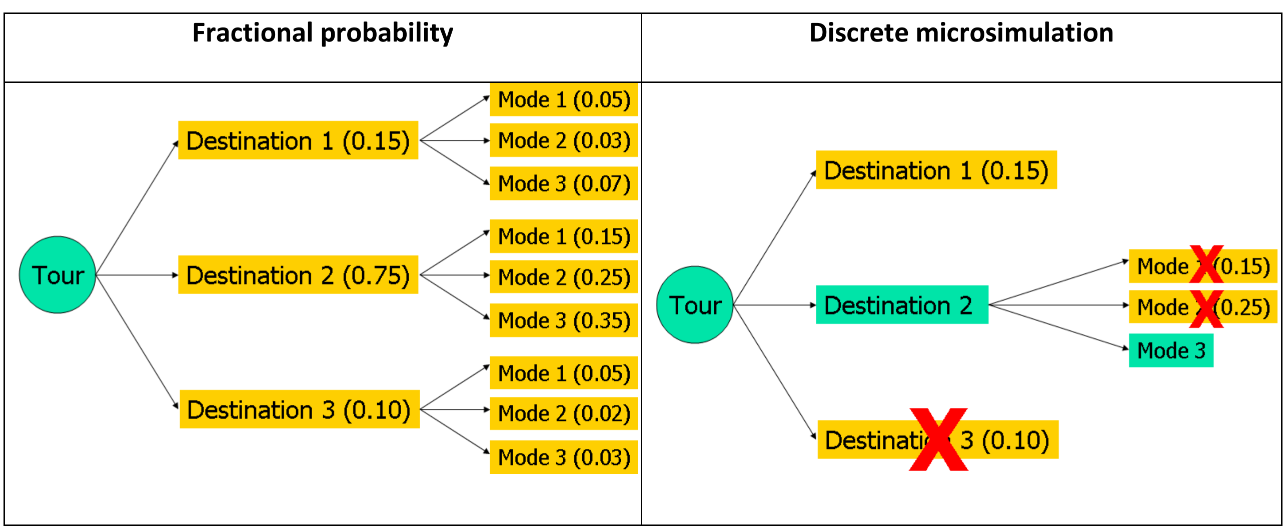


Figure 2: Discrete Microsimulation Vs. Fractional Probability

An example application of the Monte-Carlo simulation technique is shown in **Figure 3**. In this example, the choice is between four car-ownership alternatives (0, 1, 2, or 3+ cars) applied for a given household. A random number is generated (0.3897) and related to the cumulative probabilities of the alternatives to be chosen. The first alternative which cumulative probability exceeds the random number is considered as the chosen one (the 2-car alternative).

Autos	Utility	Exp(Utility)	Probability	Cumulative Probability
0	0	1.0000	0.0570	$\boxed{0.0570}$
1	1.7	5.4739	0.3122	0.3692
$\boxed{2}$	2	7.3891	0.4215	0.7907
3+	1.3	3.6693	0.2093	1.0000
Sum		17.5323	1.0000	

> 0.3897 ?

Random Number Draw = 0.3897
= 2 autos

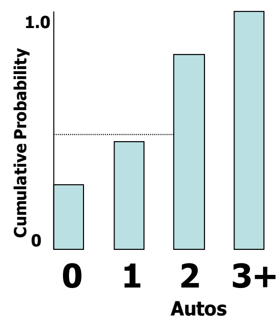
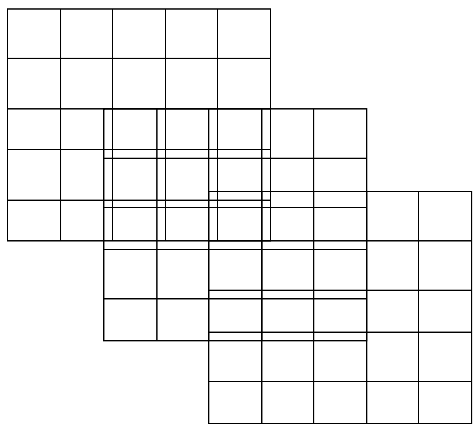


Figure 3: Monte Carlo Simulation Example – Auto Ownership

The key factor that allows for substantial computational savings (even while considering more travel segments) is that the microsimulation model works with a list of tours and trips, while a traditional trip-based travel demand model stores and calculates probabilities for each cell in origin-destination matrices, one for each market segment, for which many cells contain only a small fraction of trips.

The main advantage of microsimulation associated with compactness of the generated data structure is illustrated in **Figure 4**. In the matrix-based structure associated with trip-based 4-step models, each set of calculations (for example, mode choice) has to be replicated for each matrix cell (OD-pair). Travel or population segmentation by trip purpose, household income, car ownership, or any other attribute, results in duplicating the entire matrix structure for each segment. Introducing more detailed markets (that is necessary to address the growing number of projects and policies) results in direct multiplication of the number of calculations and one quickly runs into the so-called “curse of dimensionality”; i.e. the interaction of all variables in a model must be accommodated, even when introducing the smallest change in conditions. Contrary to this, in the micro-simulation structure, one set of calculations is implemented for each agent (household, person, or tour) and the size of the entire model is predetermined by the size of the modeled population and region. Each market segment is essentially represented as an additional column that supplies an additional data item for calculations but the number of calculations does not grow with the number of segments.

Trip-Based Models



• One set of calculations per cell
• Each market segment = new set of trip tables
• **More markets = more calculations**

Micro-simulation

HID	PID	AUT	INC	WRK	GEN	AGE	EMP
1	1	1	3	1	0	24	1
1	2	1	3	0	1	23	0
1	3	1	3	0	1	3	0
2	1	2	4	2	0	32	1
2	2	2	4	2	1	34	1
3	1	3	2	2	0	49	1
3	2	3	2	2	1	47	1
3	3	3	2	2	1	15	0
3	4	3	2	2	0	12	1

• One set of calculations per agent
• Each market segment = new column
• **More markets = no additional calculations**

Figure 4: Microsimulation Vs. Trip Matrix Structure

Implications for the core model system structure are presented in **Figure 5**. The conventional 4-step model results in a multi-dimensional array of fractional probabilities where all population and travel dimensions are multiplied. The microsimulation techniques based on a sequence of Monte-Carlo realizations results in an additive rather than multiplicative structure that is much more effective and computationally efficient. In operational implementation, the AB model looks very much as a growing database of travel details added sequentially to the synthetic population records.

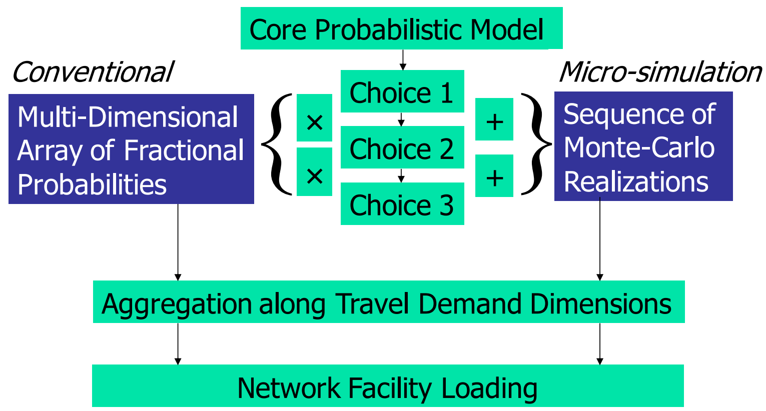


Figure 5: Application of Microsimulation in Model System

A general desire in behavioral modeling is to consider a large a number of segments by relevant household (income, car ownership, lifecycle, composition) and person characteristics (worker/student status, gender, age). The more behavioral segments considered, the more likely and prominent is the imbalance between geographic and typological outcomes using the fractional-probability approach. Consider a realistic number of zones – 2,000 and realistic number of modes – 6, with the total of 2 million tours (typical dimensions for a 2-million person region). In the fractional-probability model this would mean an array of 2,000×2,000×6=24,000,000 cells to process, while in the micro-simulation framework, there are only 2,000,000 individual tours to process, for a 12-to-1 savings in computations assuming no market segmentation.

The differences in computations become even more significant if, for example, 27 travel segments are applied (3 travel purposes by 3 income groups by 3 car ownership groups). In the fractional-probability structure these segments will be applied as additional dimensions multiplied by the number of OD cells and modes, resulting in more than 600 million cells. Using microsimulation, the number of calculations is unchanged; there would still be 2,000,000 tours. The key difference is that for each small segment and mode, the standard structure would still operate with a full OD matrix (4 million cells) despite the fact that most of the cells in this matrix would have a very small fraction of a trip. The computation burden associated with fractional probability aggregation explains why 4-step models have always been developed with a relatively small number of travel segments. ABs developed in practice include many more travel purposes (6-9 in some of the models) and socio-economic variables (7-8 person types, gender, age, presence of children in the household composition, etc) than traditional 4-step models – see [*Bradley & Bowman, 2006*]. Thus, ABs provide a balance between geographic segmentation and travel segmentation while 4-step models are characterized by an excessive geographic coverage at the expense of limited explanation of travel determinants.

Microsimulation also provides the ability to reflect the effects of previously made choices on subsequent decisions. These *decision-making chains* are usually expressed in various Boolean variables that are used in both model estimation and application, and create linkages across various choice models. In some cases, this takes the form of hierarchical nesting where several choice dimensions are considered in a joint model (usually, nested logit) and appropriate composite log-sum measures are carried from the lower levels to upper levels of the hierarchy. In other cases, the linkage has a more complicated form that cannot be reduced to a single nested structure. A good example of a strong variable of this type is a Boolean indicator of a "preschool child staying at home" (for example, because of a sickness on the given day or as a permanent household arrangement). This variable has an extremely strong impact on work and non-work tour frequency for adult household members [*NYMTC, 2004; MORPC, 2005*].

Another form of a chained decision-making relates to physical *time-space constraints* on the individual daily travel agenda. For example, if the travel associated with a work tour takes a great deal of time (in addition to the work activity duration), this may limit other home-based tours. In the ABs developed for Columbus, Atlanta, and the San-Francisco Bay Area, special "situational time pressure" indices are used that are calculated as the ratio of the residual time window (left after the higher-priority tours have been scheduled) divided by the number of generated (lower-priority) tours that have not been scheduled yet. This situational variable has very significant effects in both destination and time-of-day choice for the given tour.

Finally, an additional advantage of microsimulation relates to a natural incorporation of choice models with *random coefficients* (of mixed logit type). The interest in use of distributed parameters comes primarily from studies of highway pricing, where unobserved heterogeneity in willingness to pay can significantly affect traffic and revenue forecasts. In an aggregate model framework, application of a random-coefficient model would require integration for each OD pair and segment, making such models practically infeasible. In the microsimulation framework, there is virtually no extra computation burden for using distributed parameters, as model parameters for each traveler can be drawn from a specified distribution. This technique of individual parameter variation has been successfully incorporated in ABMs and applied for pricing studies in San-Francisco (distributed Value of Time) and New York (license plate rationing schemes where certain vehicles are prohibited from entering the priced zone on certain days depending on the last digit of the license plate number). These examples are discussed in more detail in subsequent sections.

Because Monte-Carlo methods are used to draw discrete alternatives from probability distributions, the results of the microsimulation models are subject to stochastic variability. The magnitude of the Monte-Carlo variation observed in the San-Francisco County model and ways to overcome it are described in [*Castiglione Freedman & Bradley, 2003*]. Alternative analytical procedures that reduce or even completely eliminate Monte-Carlo variation and speed up the convergence of the model system to the network equilibrium are reported by [*Vovsha Donnelly & Gupta, 2008*] for the New York model.

Market and Behavior Segmentation

Microsimulation opens a way to apply unlimited segmentation with respect to population types and travel segments. Typical person type segments are listed in **Table 13** and typical travel segments are listed in **Table 14**. These segments have been applied in most ABs in practice. For comparison, trip-based 4-step models normally include only 2-3 crude person types (workers, non-workers, and children) and 4-5 travel segments (home-based work, home-based school, home-based university, home-based other, and non-home based).

Table 13: Typical Segmentation by Person Type Applied in ABs

NUMBER	PERSON-TYPE	AGE	WORK STATUS	SCHOOL STATUS
1	Full-time worker	18+	Full-time	None
2	Part-time worker	18+	Part-time	None
3	Non-working adult	18 – 64	Unemployed	None
4	Non-working senior	65+	Unemployed	None
5	College student	18+	Any	College +
6	Driving age student	16-17	Any	Pre-college
7	Non-driving student	6 – 16	None	Pre-college
8	Pre-school	0-5	None	None

In addition to this population segmentation, any person or household variables (age, gender, income, ethnicity, etc) can be applied in an AB if supported by the Population Synthesizer. The difference is that these person types are normally considered as the most basic that justify a full segmentation of person-based models (like person Daily Activity-Travel Pattern) while the other variables are used in the utility functions for a partial segmentation.

Also, in addition to the segmentation by travel purpose shown in **Table 14**, many AB sub-models are segmented by travel party (solo vs. joint). Again, the AB structure is fully flexible with respect to which dimensions to use for a full segmentation of certain models (for example, tour/trip destination choice, time-of-day choice, and mode choice) and which dimensions to use for a partial segmentation. This decision is frequently based on statistical analysis rather than *a priori* established rules.

Unlimited population and travel segmentation is one of the major inherent advantages of individual microsimulation framework compared to aggregate 4-step models.

Table 14: Typical Segmentation by Activity Type Applied in ABMs

TYPE	PURPOSE	DESCRIPTION	CLASSIFICATION	ELIGIBILITY
1	Work	Working at regular workplace or work-related activities outside the home.	Mandatory	Workers and students
2	University	College +	Mandatory	Age 18+
3	High School	Grades 9-12	Mandatory	Age 14-17
4	Grade School	Grades K-8	Mandatory	Age 5-13
5	Escorting	Pick-up/drop-off passengers (auto trips only).	Maintenance	Age 16+
6	Shopping	Shopping away from home.	Maintenance	5+ (if joint travel, all persons)
7	Other Maintenance	Personal business/services, and medical appointments.	Maintenance	5+ (if joint travel, all persons)
8	Social/Recreational	Recreation, visiting friends/family.	Discretionary	5+ (if joint travel, all persons)
9	Eat Out	Eating outside of home.	Discretionary	5+ (if joint travel, all persons)
10	Other Discretionary	Volunteer work, religious activities.	Discretionary	5+ (if joint travel, all persons)

Variation and Randomization of Individual Parameters

In addition to explicit segmentation by population type and travel market segment, that is an expression of *observed heterogeneity*, microsimulation offers a convenient framework for a new set of techniques that express *unobserved heterogeneity*. These techniques in general cannot be applied in an aggregate 4-step framework. One of them is Individual Parameter Variation (IPV) technique. The IPV technique was successfully used for incorporation of probabilistic Value of Time (implemented in the San-Francisco ABM) and license plate rationing (implemented in the New York ABM). These two particular examples of IPV technique are discussed in the sub-sections below. IPV can be used to model incidental choices such as types of payment media, individual discounts, or other cost-related attributes that vary unpredictably across individuals. The alternative to IPV is an explicit model segmentation that quickly runs into infeasible number of segments.

IPV can also be applied for network simulations based on individual particles like advanced Dynamic Traffic Assignment (DTA) models. It is also recognized as a powerful general principle for integrated AB-DTA models that is recognized as the most promising general avenue in transportation modeling.

Example of Random Value of Time

For highway pricing studies, it is important to address a distribution of Value of Time (VOT) across multiple users. An explicit segmentation by travel purpose, income, and other variables does not solve this problem since there is always a significant situational variation within each segment that relates to the urgency of the particular trips, schedule constraints, and importance of being on time. The importance of the distribution shape is illustrated in **Figure 6**.

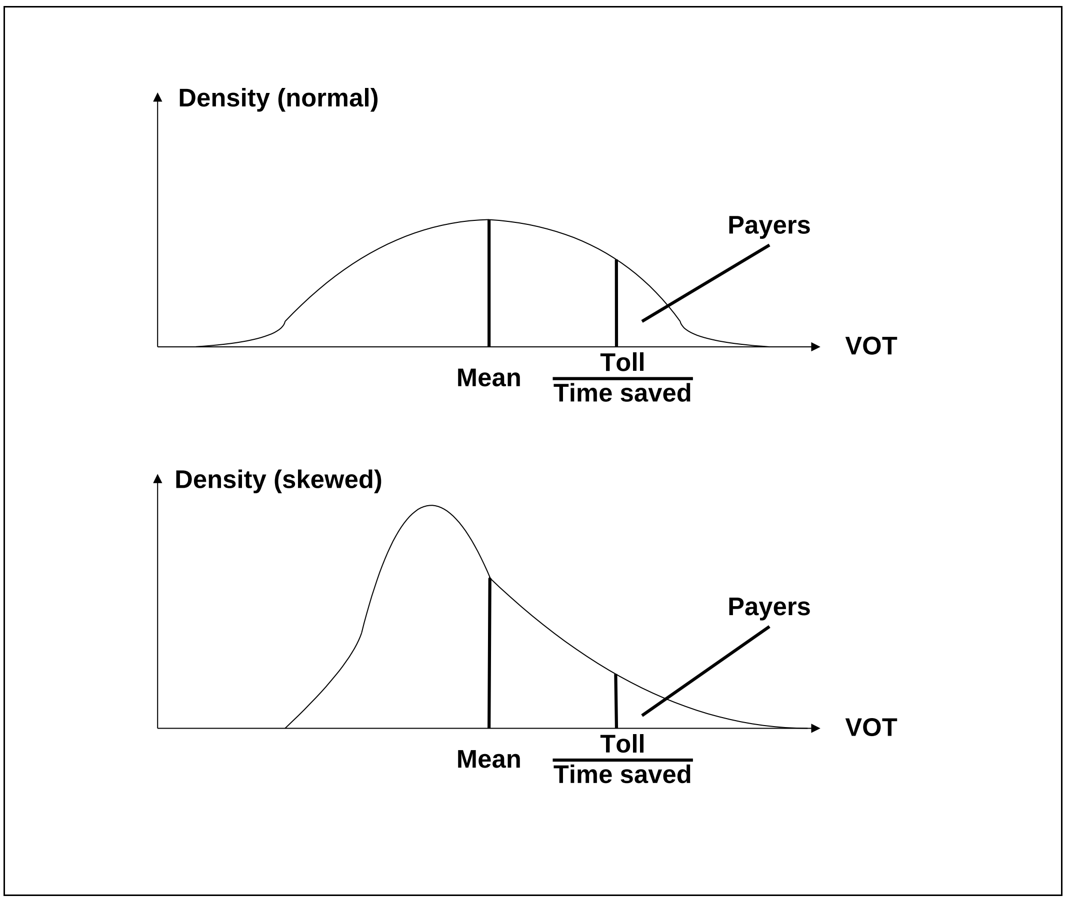


Figure 6: Importance of Consideration of Probabilistic Value of Time (VOT)

Toll road users are travelers whose VOT is greater than the ratio of toll to time savings. It can be seen that knowing only the average VOT is not enough to predict the number of toll road/lane users. The shape of the distribution has a strong impact on the number of potential users. A simple symmetric distribution (like normal distribution) has generally been recognized as not the best approximation of the real-world distribution. Some more elaborate forms skewed to the left are more behaviorally appealing reflecting on the fact that there should be a long tail corresponding to the users with very high situational VOT.

Random VOT was incorporated in the recent version of San-Francisco AB prepared specifically for pricing studies in the Bay Area. The Stated Preference (SP) survey data were used to estimate VOT distributions for use throughout the model stream. This was done by estimating a joint mode and departure time choice model, with mixed logit form, rather than nested logit form. Mixed logit is important in this case because it allows for estimation of a distribution on a coefficient, rather than just the mean value. In this case, a distribution was estimated on the travel time variable, asserting an asymmetric lognormal form. The cost coefficients are estimated as standard, non-distributed, coefficients segmented by income. The resulting model is shown in **Figure 7.**

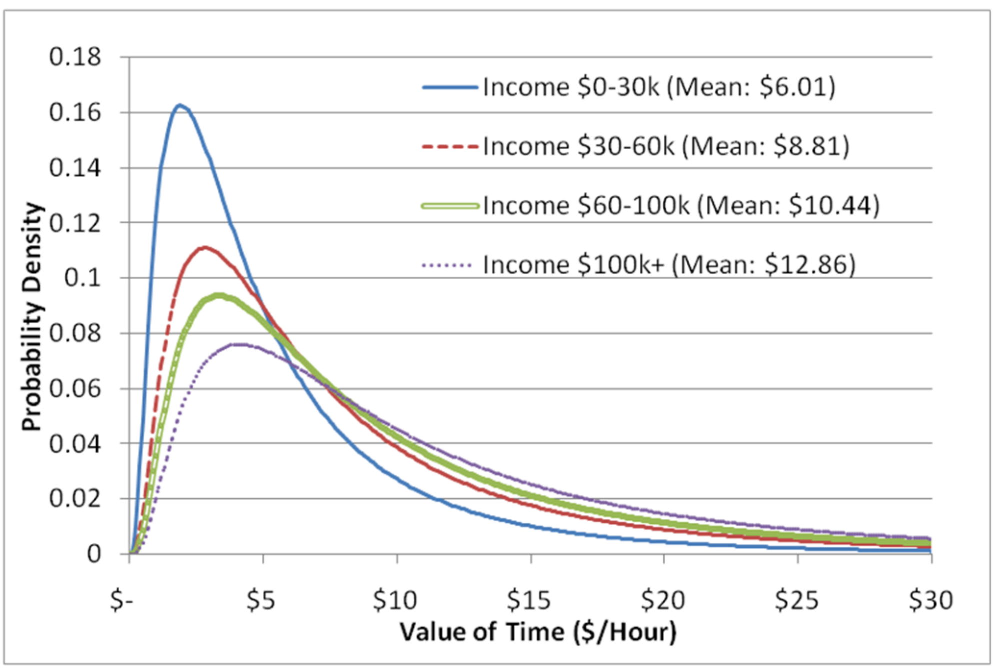


Figure 7: VOT Distributions Estimated from Mixed Logit

These distributions are used in the microsimulation process. Each individual user (travel tour) is randomly assigned a VOT from the corresponding distribution. Thus, in the process of simulation of the entire population, the actual variation of VOT across users is properly accounted. Contrary to that, aggregate 4-step models suffer from the aggregation bias when the actual variety of users is collapsed to a few segments with a single average VOT for each segment.

Example of License Plate Rationing

License Plate Rationing is a new travel management policy that represents a challenge to modelers and requires an advanced IPV technique. The essence of License Plate Rationing is that a certain percentage of vehicles (for example, 10% or 20%) are subject to a no-drive-to-CBD policy based on the last digit of license – see **Figure 8**. This type of policy cannot be addressed with a 4-step model, but an advanced microsimulation framework opens a way to effectively model it.

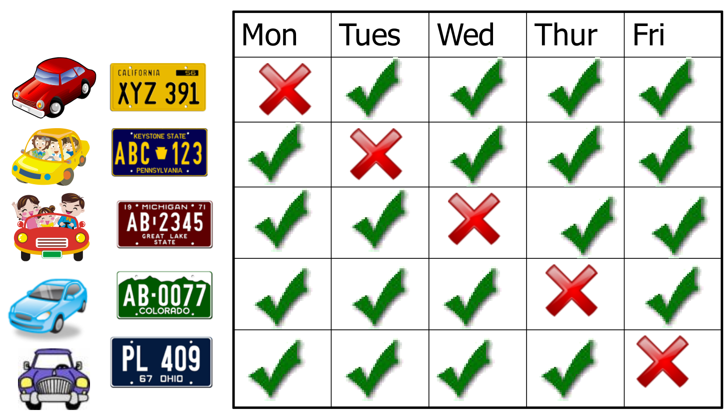


Figure 8: What is License Plate Rationing?

The corresponding modeling technique essentially falls into the general category of IPV which is one of the most powerful advantages of microsimulation. In contrast to the aggregate 4-step models where any variation in parameters requires an explicit segmentation of the entire trip table by all combined categories, microsimualtion allows for any variation in individual parameters, either in the form of predetermined categorized segmentation, or in the form of randomly drawing from a distribution accounting for situational variability. It can be incorporated at practically no cost in terms of model complexity. The IPV technique can be applied to any behavioral parameter used in the demand model.

In the context of License Plate Rationing, the Individual Parameter Variation principle is applied through the Household Auto Availability model – see **Figure 9**.

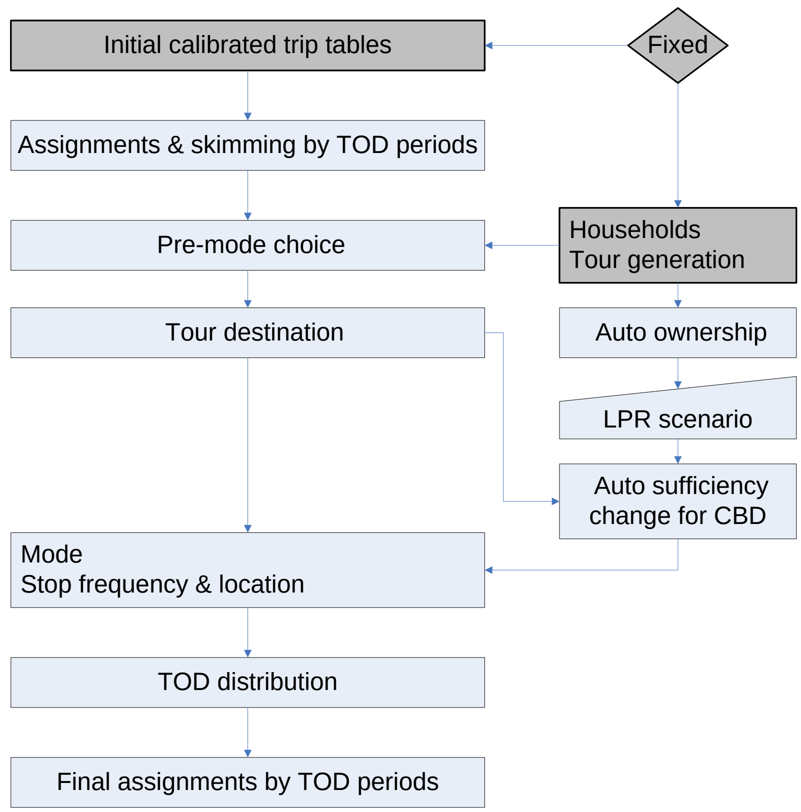


Figure 9: New York ABM Application for License Plate Rationing

In the microsimulation process, for each household some cars are randomly tagged as unavailable for travel to the Congestion Pricing Zone (CPZ) based on the rationing policy that defines the probability of disabling a car as shown in **Figure 10**. This affects the household car sufficiency variable (number of cars minus number of workers) that has a strong impact on mode choice, as well as on choices of frequency and location of intermediate stops for the given tour.

For each car in each household, a random number between 0 and 1 is generated independently and compared to the rationing percentage. In the example, 20% of cars are subject to the ban, thus, all random numbers that are less than 0.2 are tagged as "unavailable" autos. After the tagging, the number of autos available for driving into CPZ is recalculated and the household car-sufficiency index (number of autos minus number of workers) is recalculated accordingly. This adjusted car-sufficiency index is further used for modeling travel choices (specifically, mode and time of day).

		Not-CPZ		Random #'s for tagging				To CPZ	
HH#	Wkrs	Autos	Car Suff	a1	a2	a3	a4	Autos	Car Suff
1	2	3	1	0.914	0.190	0.245		2	0
2	1	1	0	0.988				1	1
3	1	2	1	0.246	0.487			2	1
4	1	1	0	0.121				0	-1
5	2	4	2	0.375	0.878	0.165	0.341	3	1
6	2	2	0	0.080	0.660			1	-1

For Tours Not to Restricted Area **For Tours to Restricted Area**

Figure 10: License Plate Rationing: Car Availability by Destination

In the model application at this stage, we assume that there is no impact on tour frequency choice and primary tour destination choice. This makes the comparison across scenarios easier since the same subset of tours with the destination in CPZ that are affected by the Rationing is fixed. Using a household car-sufficiency variable rather than person car availability allows for an accounting of interchangeable vehicle allocation and use within the household.

Tour-Based Techniques and Challenges

Typical Tour Choice Structures and Dimensions

One of the main features of tour-based modeling technique is *bi-directionality* of tours. This makes it principally different from the one-directional trip-based technique that is a foundation of the 4-step modeling paradigm. This aspect has important implications for the computational stream of the model. The associated dimensionality and computational issues has to be taken into account, and specifically, in the context of travel dimensions directly affected by bi-directional LOS variables (travel time and cost).

The related choice models include tour-level choices (destination, time-of-day, mode, stop frequency, and stop location) and trip-level choices (mode, departure time). The most complicated part is associated with ensuring a proper consistency and conditioning between tour-level and trip-level choices although certain important issues have to be taken into account already at the tour level. Considering the entire tour (round trip in the simplest case) not only simply doubles the number of LOS variables compared to a trip-level model but also requires a much more elaborate consideration of the combinatorial issues of choice alternatives for mode and time-of-day choice.

Bi-directionality of Travel – Example of Tolls

One of the important aspects of bi-directionality associated with tours relates to modeling tolls differential by direction and time-of-day (as the most frequent case with congestion pricing) but not necessarily symmetric (i.e. outbound AM toll in the peak direction does not have to be equal to inbound PM toll in the peak direction). This situation may not be frequent for fixed toll schedules but it is very common for dynamic pricing that is recognized as the most promising technology of highway pricing. This situation is illustrated in **Figure 11** where a corridor is considered between the Congestion Pricing Zone (CPZ) and outside areas.

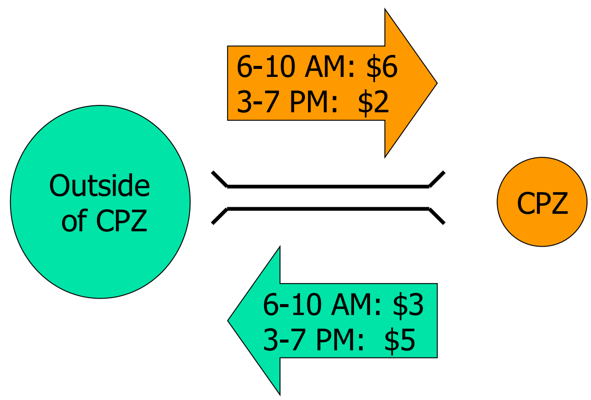


Figure 11: Example of Bi-directional Tolls

As the result of this pricing scheme, different users will experience different daily tolls as shown in **Table 15** depending on the time-of-day combination of the outbound and inbound commutes. We have to consider five time-of-day periods because of the toll schedule: 1) earlier than 6AM, 2) from 6 to 10 AM, 3) from 10 AM to 3PM, 4) from 3PM to 7PM, and 5) later than 7PM. In order to properly account for the total toll paid in both directions we have to consider all 15 feasible combinations of outbound and inbound periods (where inbound time cannot be earlier than the outbound time). Thus, modeling true tolls for each user group represents a conceptual and computational challenge. All 15 segments have to be explicitly defined in the model and the users have to be broken into the corresponding groups. This essentially means that a tour-based time-of-day choice model has to be applied (in combination with the other models including mode and route choice).

Table 15: True Tolls Paid by Commuters

Outbound time	Inbound time	Toll, $
Earlier than 6AM	Earlier than 6AM	
Earlier than 6AM	6-10AM ($3)	3
Earlier than 6AM	10AM-3PM	
Earlier than 6AM	3-7PM ($5)	5
Earlier than 6AM	Later than 7PM	
6-10AM ($6)	6-10AM ($3)	9
6-10AM ($6)	10AM-3PM	6
6-10AM ($6)	3-7PM ($5)	11
6-10AM ($6)	Later than 7PM	6
10AM-3PM	10AM-3PM	
10AM-3PM	3-7PM ($5)	5
10AM-3PM	Later than 7PM	
3-7PM ($2)	3-7PM ($5)	7
3-7PM ($2)	Later than 7PM	2
Later than 7PM	Later than 7PM	

Modeling true tolls and associated LOS variables with an aggregate trip-based 4-step model is impossible. More specifically, it is impossible to ensure any reasonable level of consistency across trip distribution, mode choice, and time of day choice with regard to real toll values. In the 4-step framework, commuters will see a limited range of toll values for each trip (up to $6 in the outbound direction and up to $5 in the inbound direction). With a tour-based microsimulation model it is still difficult to ensure a full consistency, but a much better job can be done with a proper accounting for the entire range of tolls up to $11. The corresponding computational side involves a joint choice of tour mode and time of day with a large number of alternatives (several dozens of modes by several hundred of combined time-of-day alternatives).

Treatment of Space and Spatial Choices

An important trend in ABs is the move toward the use of finer spatial and temporal detail. In terms of spatial detail, there are three options that should be considered for the CMAP advanced model: 1=TAZ (that is currently used in the CMAP 4-step model and results in thousands of zones), 2=Smaller geographic unit that relates to non-motorized and transit accessibility (as described below for the San-Diego AB; this results in dozens of thousands of zones), 3=Parcel (as described below for the Sacramento AB; this results in hundreds of thousands of zones). In terms of temporal detail, the main choice is between an AB that operates in discrete space with 30-min resolution and employs discrete choice technique for time-of-day choice (as in ABs developed for Atlanta, Sacramento, Bay Area, and San-Diego) and continuous-time ABs that employ duration models for time-of-day choice (as planned for the Southern California AB that will be the first AB of this type in practice). Both important issues (level of spatial resolution and level of temporal resolution) still remain open for CMAP.

The ultimate goal is to operate with continuous time and space as this is the way travel choices are made by people in reality. In operational models, this is expressed in using smaller units of geography and smaller time increments in the discrete space. Finer detail has many benefits, all arising from the increased accuracy of measurements: impedance between locations (especially important for transit access/egress and non-motorized travel) and attractiveness of locations themselves (i.e. an exact mix of land-uses, population, employment, special generators, or parking lots in each location). However, using greater detail presents major computing and data challenges. At least two specific issues must be addressed with respect to a finer level of spatial resolution:

- *Specific treatment of spatial choices that involve a large number of alternatives*. Effective sampling and data-processing procedures have been developed for each of possible spatial levels. In particular, destination choice models account for 50% of the run time of the core demand model (not including network assignments) primarily because of the need to calculate time-consuming logsums over modes and time-of-day periods for each destination. Effective two-stage sampling procedures have been tested. At the first stage, 100-200 destinations can be quickly sampled based on the distance and land-use size variables only. At the second stage, 20-30 destinations are selected out of 100-200 based on a simplified logsum measure (in addition to distance and size variable). At the final stage, a full model is applied for a small subsample of 20-30 “good” destinations only.

- *Effective manipulation of large multidimensional arrays*, including tour and trip matrices, LOS skims (one-directional and bi-directional), and joint household/person distributions. Direct manipulation and storage of two-dimensional (origin-destination) matrices is only feasible at the TAZ level. For smaller units and parcels, special algorithms for calculation of LOS attributes on the fly have to be applied. Fortunately, the core microsimulation structure of most of the advanced models offers good opportunity for on-the-fly calculations as discussed below in the examples for San-Diego and Sacramento ABMs.

Treatment of Geographic Space

This section describes an approach to addressing 'the tyranny of zones' problem in travel demand forecasting models, within the context of an AB. The existing trip-based modeling system developed for San Diego Association of Governments (SANDAG) utilizes multiple layers of spatial analysis, including a typical TAZ system for representation of auto times and costs, and a spatial classification scheme of approximately 33,000 polygons for representation of trip ends, transit and non-motorized accessibilities, and land-use data. One of the unique aspects of this system is that transit modal availabilities, times, and costs are explicitly based on the locations of trip ends vis-à-vis the transit stops that serve them. This ensures that the access and egress times to transit stops are consistent with the skimmed in-vehicle times and costs, and allows the model to represent trade-offs between walk time and transfers, as often occurs for feeder bus access to rail. The traditional method of utilizing zonal-based LOS coupled with aggregate zonal percent walk for TAZs is obviated by this approach.

The SANDAG model encompasses San Diego County, which is approximately 4,500 square miles, and was home to 3,173,407 million persons in January 2009. The modeling area is represented by a TAZ system of 4,600 zones, and a finer-detailed spatial system of 33,353 polygons, referred to as Master Geographic Reference Areas (MGRAs), as shown in **Figure 12**.

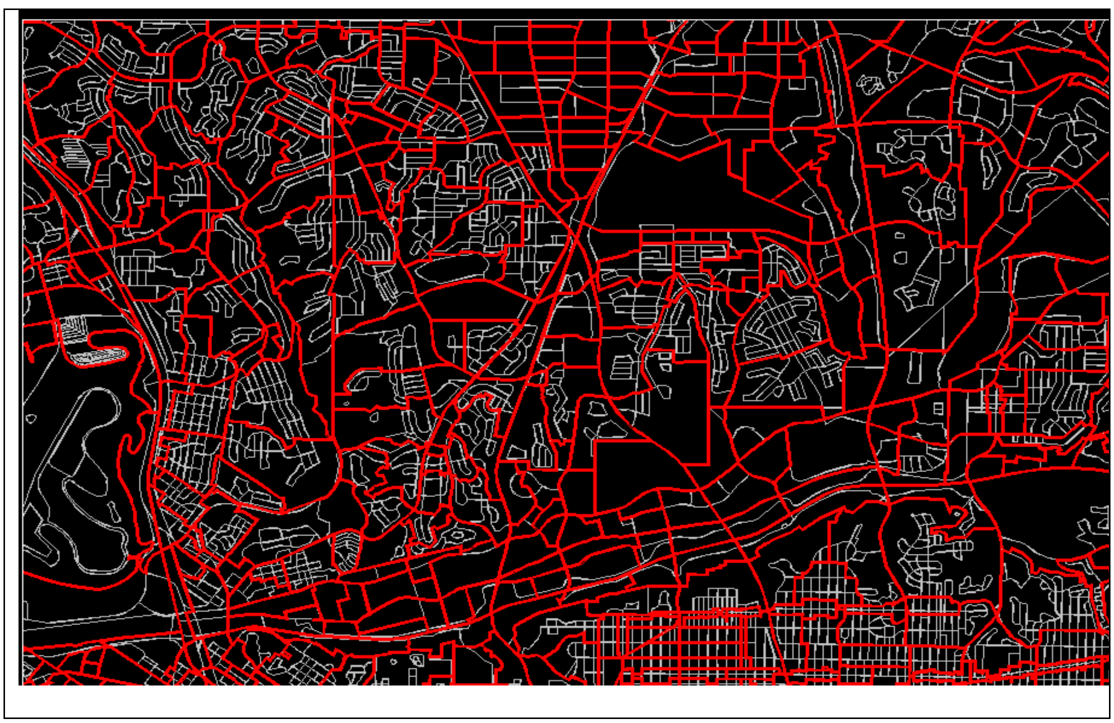


Figure **12**: MGRAs and TAZs (San-Diego ABM)

The TAZ system is used to skim auto networks and assign auto demand. Transit and non-motorized travel is handled differently. The transit network is coded with explicit representation of transit stops as 'dummy zones', or Transit Access Points (TAPs) as shown in Figure **13**. TransCAD is used to skim stop-to-stop travel times and costs, including the standard in-vehicle time, first wait, transfer wait, etc. Walk access to/from transit stops is calculated between MGRA centroid and transit stop using GIS methods that take into account physical constraints and slope (Figure **14**). Then the utility of travel by each stop-stop combination is computed on-the-fly within the model application software. This approach ensures that the access times that are computed using GIS and used in accessibility calculations are consistent with the in-vehicle times that are calculated using the TransCAD transit skimming procedure. This is a significant improvement over other ABs that utilize detailed location data and assume that the nearest transit stop to each parcel is consistent with the LOS matrix skimmed at a zonal level. In the SANDAG approach, no such simplifying assumptions are made, and trade-offs regarding walking distance to transit versus in-vehicle time and transfers are explicit and accurate.

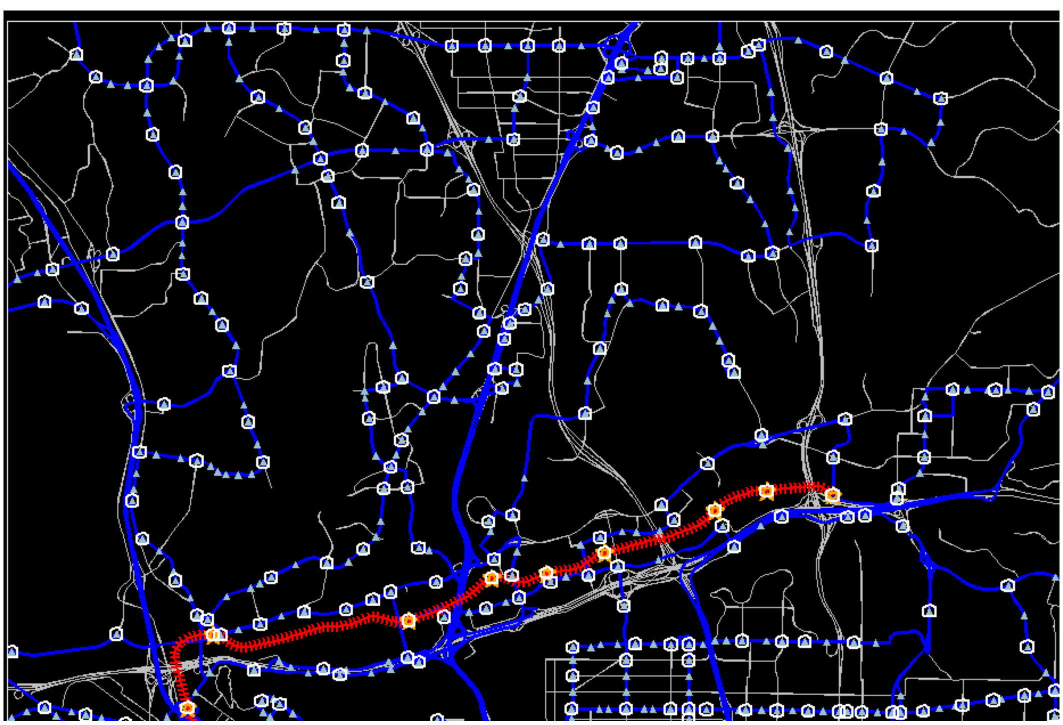


Figure **13**: Transit Network, Stops, and Transit Access Points

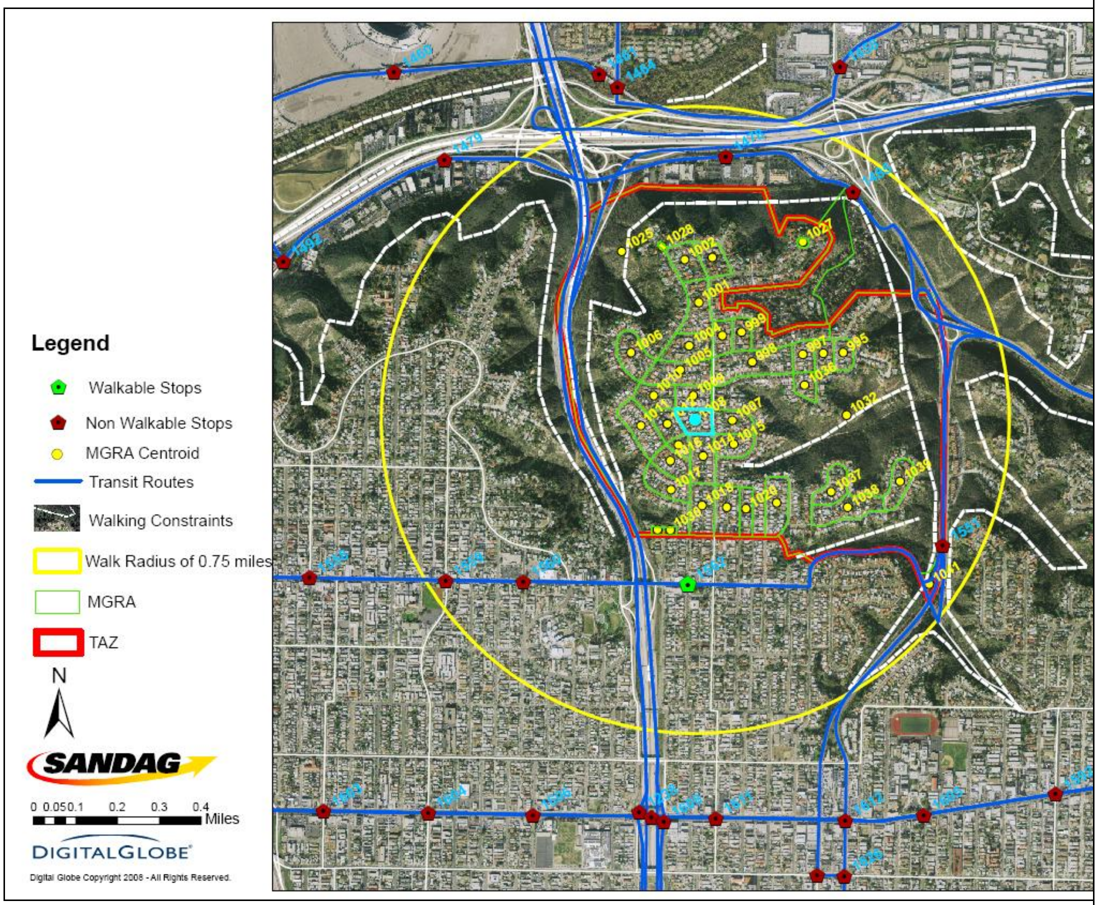


Figure **14**: Walking Constraints

Figure **15** shows an example of the explicit trade-offs that the model considers. In this figure, there is a choice between walking to a bus that offers direct service between the origin and destination MGRA, versus walking a short distance from the origin MGRA to a feeder bus that provides access to an LRT station, and versus walking further directly to the origin LRT station. Utilities are calculated for all potential choices in the mode choice application software, as well as auto access utilities (not shown). Auto access times are calculated based on the TAZ of the trip origin (according to its MGRA) and the TAZ of the boarding transit stop/parking lot. Non-motorized utilities, including walking and biking, are also represented at the MGRA level. All of these lower-level calculations are utilized in a nested logit mode choice model, and the logsum of the model is taken for calculation of subsequent multi-modal accessibility measures.

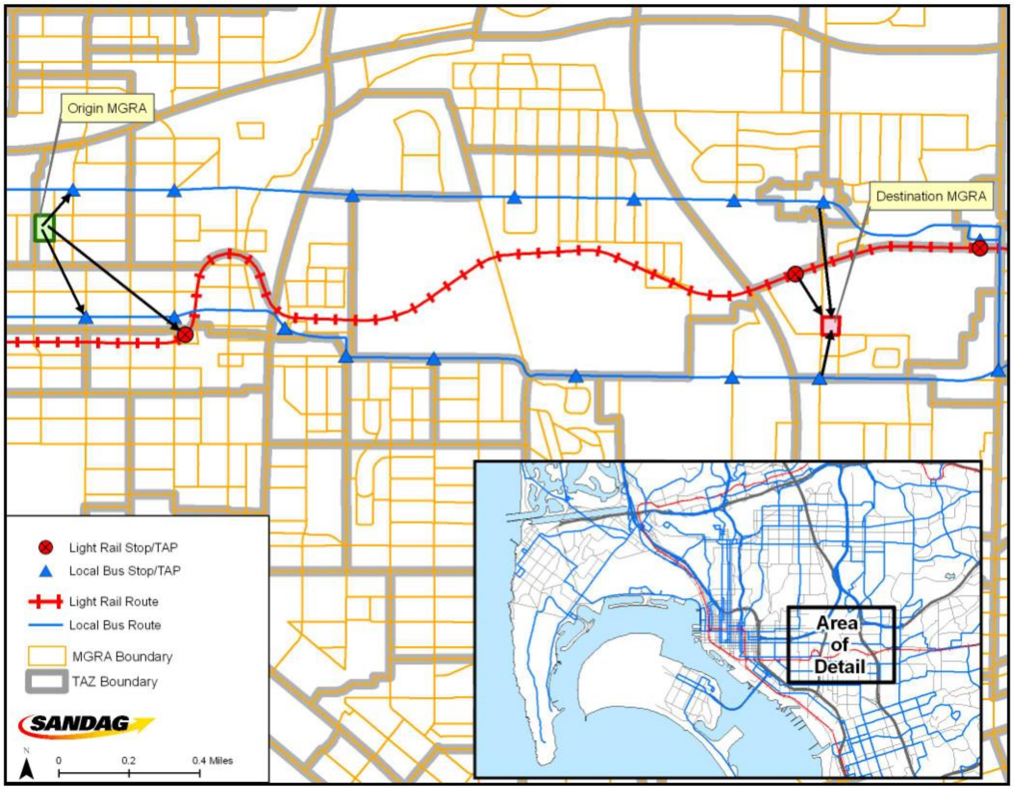


Figure 15: Transit Path Examples (San-Diego ABM)

Treatment of Space (Sacramento ABM)

A key feature of the DaySim line of ABs, introduced first with the Sacramento version, is the option to use parcel level detail for quantifying spatial attributes and defining location choices. Since there are many hundred thousand parcels, it is computationally infeasible to calculate skim matrices for all possible parcel pairs. The solution for this that has been implemented in the first two versions of DaySim is to use TAZ-to-TAZ skims for basic measurement of travel times and costs, and to adjust and supplement these measures in two ways, taking advantage of parcel information on each end of the

Origin-Destination pair. For walk access to transit, parcel attributes measuring the distance to transit stops of various types are used to approximate the walk access and egress times for transit access. These significantly improve on traditional measures of walk access time.

For other impedance measures, DaySim computes two measures of proximity at parcel level. One is a parcel-to-parcel orthogonal distance (the sum of the "X" and "Y" coordinate distance separating two parcels). The second is a conventional TAZ-to-TAZ distance skim, comparable to skims for four-step, TAZ-based models. Based on orthogonal distance estimate, the two measures of proximity are formulaically combined. For nearby parcels, the parcel-to-parcel distance is weighted heavily; for distant parcels, the TAZ-to-TAZ distance skim is weighted heavily. By using this combined approach, unique measures of parcel-to-parcel distance are computed, reflecting the "true" proximity to a greater degree than do TAZ-to-TAZ skims alone. These are used to adjust TAZ-to-TAZ travel times as well, using a pivot procedure. The technique has been shown to significantly improve the measurement of impedance for short trips [*SACOG, 2008*].

The use of parcel data presents additional computational challenges that must be addressed. First, some parcel attributes are measured in terms of buffer zones surrounding the parcel centroid, such as employment within ¼ mile of the parcel centroid. The large number of parcels makes these attributes time consuming to calculate. GIS techniques are used to calculate them as rapidly as possible. Second, in destination choice models, sampling of alternatives is used because of the large number of parcel alternatives. Nevertheless, it is necessary to calculate travel times to all choice locations in order to determine sampling probabilities. A two-stage sampling technique is used. The first stage samples zones using TAZ-to-TAZ travel times, so that travel times must only be calculated for all zone pairs (instead of all parcel pairs). The second stage samples parcels within zone, using only parcel size information, which is less costly because size depends only on the destination location, not on the origin and destination pair. When the model is run, sampling probabilities are calculated up front and the results are then used repeatedly during the simulation, saving computation time.

Another important aspect of spatial and temporal detail is in the representation of traffic conditions and level of service across the regional network at various times of day. Applied ABs to date have all used static equilibrium traffic assignment at the TAZ-to-TAZ spatial level, for four or five different periods of the day. With the eventual shift to dynamic traffic assignment (DTA) and traffic microsimulation, there will be much greater detail available for the travel demand models. Taking advantage of that detail, however, presents significant software challenges. The DaySim developers have begun to tackle those challenges in two different DaySim-related projects. The first project was one funded by TMIP to integrate DaySim with the Transims Router for the Sacramento Region [*Resource Systems Group, et al, 2010a*]. In that project, DaySim still uses TAZ-to-TAZ travel time skims from Transims, but has been adapted to store and use skims for 22 different time periods in the day, with periods as short as 30 minutes during the peaks. In the SHRP 2 C10A project, DaySim is being adapted further to work with both the Transims Router and Microsimulator, for both Burlington, VT and Jacksonville [*Resource Systems Group, et al, 2010b*]. In that project, DaySim will also accommodate more spatial detail in the network travel time skims, with matrices for as many as 20,000 different "activity locations", depending on hardware memory configuration.

Flexible Temporal and Spatial Disaggregation

As computing power improves, it becomes feasible to generate travel times and costs in network models that are spatially and temporally more and more detailed than traditional TAZs for five time periods in the day. Likewise, it becomes feasible to hold more travel time and cost data in memory while the AB software runs and to accommodate the computational cost of using these data. Also, while some agencies that use these models can generate and use parcel information, others can provide information only at the zone or subzone level. The enhanced version of DaySim (Version 3) will accommodate these realities by introducing a parameterized, scalable method of defining and using spatial and temporal information. The design includes the following features:

Allows locations to be defined at either of two levels of spatial aggregation: TAZ or any finer level of detail called here, for convenience, parcel level.

Allows two different levels of spatial aggregation for LOS matrices: TAZ or a finer level of detail called here, for convenience, Skim Locations (SL).

Allows two different levels of temporal aggregation for LOS matrices—a basic level currently used by the MPO (typically 4 or 5 periods in the day), or more detailed (such as 22 periods in the day, used in the SACOG Transims application).

Allows the user to specify the level of spatial aggregation (TAZ or SL) and temporal aggregation (basic or detailed) for the storage and use of each LOS matrix (different modes and LOS variables may use different levels).

Implements procedures that can use one matrix that is defined with very fine time and space detail, such as auto travel time or distance, to adjust (via a pivot procedure) attributes retrieved from other matrices that are defined with less detail.

Makes the parameterized definitions of the more detailed “parcels”, SLs and time periods easily scalable, so that more detail can be added as memory and run-time considerations allow.

These techniques are aimed at (a) accommodating the level of spatial and temporal aggregation that is feasible institutionally and technically, (b) maximizing the amount of spatial and temporal information derived by DaySim from the available data and network simulation methods, and (c) enabling the level of detail to increase over time as it becomes available and computing power increases.

Internal Database and Types of Objects

The microsimulation paradigm goes hand-in-hand with the Object-Oriented Programming (OOP) paradigm. In AB model implementation, we are dealing with such objects as households, persons, tours, and trips that are related to each other by the inheritance principle of OOP as shown in **Figure 16**. This is a very useful view on the internal structure of the model and associated database. The corresponding data fields contain household, person, tour, and trip attributes with the ability to invoke any data attribute that belongs to the parent object. For example, any person-level choice model can use the household income as a variable while any tour-level model can use the person age as a variable. The methods (functions) for these computerized objects relate to the corresponding choice dimensions that essentially manipulate data fields. For example, each tour has an important attribute – primary destination that is calculated by applying a destination choice model.

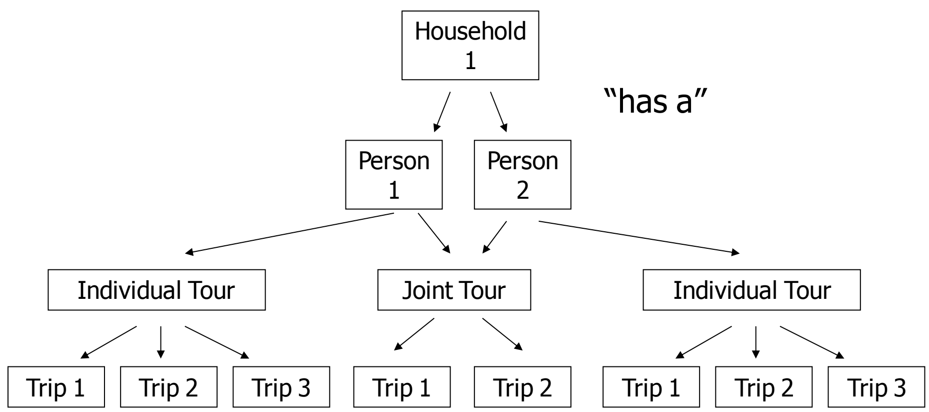


Figure 16: OOP Composition Pertinent to ABM

In addition to conceptual and programming elegancy, the object structure of AB modeling is an important vehicle for multithreading and distributing of calculations that is essential for an effective and efficient implementation of an AB for a large region like the CMAP modeled area. In particular, it is important to mention that most of the choice models and associated calculations in the core AB model are implemented independently for each household (this is not true, however, for network assignments where different ways of decomposition should be applied). This is a very beneficial feature and several AB implementations (in Columbus, Atlanta, and Bay Area) have already taken a full advantage of it by distributing packets of households across multiple CPUs as described in the sections on software implementation below.

Practical Logsum Computation

An emphasis from the beginning of this line of models, which has been enhanced over time, is realistic model integration. The careful use of well-specified logsum variables enhances the realistic sensitivity of all models in the system to changes in transport conditions, especially travel time and cost.

Unfortunately, the strength of the logsum variable as a composite measure rests in a feature that makes it computationally expensive, and essentially infeasible with very large and detailed hierarchical model systems. In order to model the highest level outcome, utilities of all alternatives in the entire hierarchy must be computed. When applied across a hierarchy of models at the person-day, tour and trip levels, each of which can have hundreds of alternatives, there are billions of possible combinations, making the use of full nested model logsums infeasible. Two computational techniques have been applied in ABMs in practice to achieve a reasonable level of upward vertical integrity with a manageable amount of computation.

The first technique (that was successfully applied in the Sacramento DaySim ABM) is to avoid the use of a logsum when applying an upper level model by treating as given a conditional outcome that is not known, and would otherwise require the calculation of a logsum from all possible conditional outcomes. The assumed conditional outcome is selected by a Monte Carlo draw using approximate probabilities for the conditional outcome. Rather than making every simulated outcome sensitive to variability in the conditional outcome, sensitivity is achieved across the population through the variability of outcome in the Monte Carlo draws. This technique is used to incorporate time-of-day sensitivity into the tour mode choice logsums used in the tour destination choice models. In this way, the destination choice models are sensitive to variations in transport level of service and spatial attributes across all possible combinations of time-of-day and mode, with the affects approximately weighted by the joint time-of-day and mode choice probabilities.

The second technique (that was successfully applied in the Sacramento DaySim and San-Diego CT-RAMP ABs) is to calculate an approximate, or aggregate, logsum. It is calculated in the same basic way as a true logsum, by calculating the utility of multiple alternatives, and then taking expectation across the alternatives by calculating the log of the sum of the exponentiated utilities. However, the amount of computation is reduced, either by ignoring some differences among decision-makers, or by calculating utility for a carefully chosen subset or aggregation of the available alternatives. The approximate logsum is pre-calculated and used by several of the model components, and can be re-used for many persons. The categories of decision-makers and the aggregation of alternatives are chosen so that in all choice cases an approximate logsum is available that closely approximates the true logsum. In essence, this is a sophisticated ad hoc measure that is intended to achieve most of the realism of the true logsum at a small fraction of the computational cost. Two kinds of approximate logsums are used, an approximate tour mode-destination choice logsum and an approximate intermediate stop location choice logsum. Aggregate logsums as “accessibility” variables can also be specified and used in various other ways, but the overall concept and software approach remains the same.

Transportation Network Procedures

Activity-based models like traditional trip-based models should be integrated with network assignment and skimming procedures that produce Level-of-Service (LOS) data needed that serve as supply-side variables in the demand model. There are two main options for CMAP to consider in this regard: 1) conventional Static Traffic Assignment (STA) based on User Equilibrium, and 2) Dynamic Traffic Assignment (DTA). These procedures normally take 60-70% of the time needed for a model production run (i.e. they constitute the more significant time-taking component of advanced models). DTA is theoretically more robust but is not yet realistic operationally on a scale of the CMAP modeling region although it might become realistic in the near future. A short term option should include a set of STAs (preferably run for each hour) with a linkage between them. A long-tem option which we recommend to consider while designing a new travel model and freight model is a fully disaggregate and integrated microsimulation demand and network model system based on DTA.

In general, with respect to network assignment procedures, CMAP will likely choose one of the existing commercial software packages provided by INRO, Caliper, Citilabs, or PTV. All of these vendors today provide both STA assignment tools for regional planning purposes (EMME, TransCAD, Voyager, Visum) and DTA/microsimulation tools for corridor/subarea analysis (Dynameq, TransModeller, Avenue, Vissim) with the tendency to close the gap between STA and DTA in terms of network size and runtime. In addition to major transportation software vendors, there are credible DTA packages developed by universities (like DynaSmart or Dynus-T). It should be noted, that switching from STA to DTA only resolves the issue with highway assignment and skimming. Transit assignments and skimming would still require a transportation planning package although some DTA developers (Dynus-T) are planning to incorporate transit simulations in the DTA package.

With respect to software development, it is important to recognize a principal difference between the demand-side AB and supply-side DTA. This dictates a difference in approaches that CMAP should consider when managing decisions to buy, adopt, adapt, hire a consultant to develop a custom software, or develop software in-house. Advanced demand AB models cannot currently be implemented using script languages of transportation packages. Thus, the AB software is normally developed by consultants using general-purpose program languages (like C/C++ or Java). Some vendors are trying to penetrate the market and offer a script language for AB, but this has not been utilized yet in practice. Contrary to that, network simulation software, because of its more rigid operational constraints, is almost always purchased from a commercial vendor.

DTA is characterized by a sophisticated but generic algorithm (not customized for a specific application). The DTA algorithm itself has a relatively small number of parameters with recommended default values. Custom model estimation of route choice is not typically required. DTA calibration in practical applications relates mostly to engineering treatment of physical characteristics such as lane width, posted speed, etc. Contrary to that, ABM is characterized by a less sophisticated (from the mathematical standpoint) but more specific algorithm tailored to the specific regional conditions. It includes a large number of widely-derived behavioral choice sub-models and parameters to estimate and calibrate.

Achieving Demand-Network Equilibrium with ABM

There are several theoretical challenges associated with the achievement of network equilibrium with ABs [*Vovsha Donnelly & Gupta, 2008*]:

- Running the same model twice with exactly the same set of inputs and the same starting random seed would reproduce the results exactly; the reason for practical non-convergence is that small variations in LOS variables can produce significant local disturbance in the chain of demand-related choices. So the traditional way of implementing multiple iterations by feeding back LOS variables will not work within the AB modeling approach.

- There are two types of Monte-Carlo effects which have significant and distinctive impacts on global convergence – one relates to microsimulation with a fixed structure of modeled agents and choices, and the other relates to microsimulation where the structure of modeled agents and choices is dependent on the results of microsimulation of prior choices in the model chain. Microsimulation of choices with a fixed structure is the simpler case where convergence can be achieved by simple iteration since averaging the LOS variables will ultimately reproduce not only the aggregate shares but even individual choices. Microsimulation of choice model chains with structural impacts can have discontinuity and abrupt responses to even small variations in inputs; in this case it is difficult to predict the level of convergence and to analytically estimate the Monte-Carlo component of variation.

There are essentially two practical ways to ensure convergence (assuming that a fixed-point equilibrium solution exists):

- *Enforcement*. These methods are specific to microsimulation and designed to ensure convergence of "crisp" individual choices by suppressing or avoiding Monte-Carlo variability. These methods are currently only at an early stage of theoretical development, with some empirical strategies being tested.
- *Averaging*. These methods have been borrowed from conventional 4-step modeling techniques, but can be also used with microsimulation as long as they are applied to iterated LOS variables and/or synthetic trip tables generated by the microsimulation process.

"Enforcement" Methods to Achieve Equilibrium

Several ways to "enforce" convergence at the individual level have been suggested and tested in practice. From the theoretical perspective they can be broken into three groups:

- *Re-using the same random numbers or starting random seeds* for certain choices that would ensure that the choice will be replicated if no change occurs to the inputs,
- *Gradual freezing* of portions of households or travel dimensions from iteration to iteration,
- *Analytical discretizing* of probability matrices instead of Monte-Carlo simulation.

Re-Using the Same Random Numbers / Seeds

One of the reasons for instability of the demand microsimulation that manifests itself, even if the LOS variables converge, is that the microsimulation generates "crisp" choices from and on the top of the probabilities generated by choice models. This is usually done by generating a random number for each choice and relating this number to the choice probabilities (analogous to Monte-Carlo roulette). Even if the probabilities become constant, the Monte-Carlo variability alone causes a random fluctuation of the individual microsimulation results, although the results are quite stable at a sufficient level of aggregation.

To avoid Monte-Carlo variability, random numbers associated with each choice can be generated in advance, stored and re-used when the model is applied for different scenarios and across different iterations. In this case, it is enough to store a seed that would automatically generate the same sequence of random numbers. Application of this simple strategy, however requires a structural stability of the agents and their decision chains embedded in the model structure. In comparatively simple model structures, the list of simulated agents (households, persons, tours, and trips) and their choice alternatives are fixed from iteration to iteration and only associated choices probabilities fluctuate with changes in network times and costs. In this structure, it can be shown that convergent probabilities (as a function of convergent LOS variables) would ensure convergence of the individual choices. The same random number will be always applied to the same agent and choice dimension. Examples of structurally stable microsimulation models include tour, mode, and destination choice with a fixed set of generated tours.

However, this strategy becomes problematic for more complex decision chains where there are structural impacts of prior choices on subsequent choices in the model chain. For example, the daily activity pattern model that generate tours (i.e. create a list of tours by type for every person) is in itself a simulation model. At each iteration, it may generate a different set of tours for the same person. The subsequent mode & destination model would then be applied for a different set of agents; thus, freezing a seed for each person will not help in mode and destination choice. It would generate the same sequence of random numbers for each person, but these random numbers would be applied to different agents and associated choices.

Theoretically, structurally stable decision chains can be ensured by considering the maximum possible number of agents created at each stage of microsimulation and reserving a random seed for each of them. For example, if the maximum modeled number of work tours per day generated by a worker is two, we could create a placeholder for the 1st and 2nd tour random seed, and just not use it if the tour was not actually created. This, however, could create quite a huge system of placeholders in advanced ABs, since they include numerous structural components, such as multiple stops on each half-tour.

Re-using random seeds in different forms has been applied in most ABs in current practice.

Gradual Freezing of Portions of Households or Travel Dimensions

This is a set of empirical procedures that is based on a predetermined strategy of progressively freezing certain portions of the simulated agents over the course of global iterations (i.e. fixing the corresponding choice outcomes until the end of model application). In general, decisions are required for the following components of this strategy:

- Principal sequence of agents to freeze that can be, for example:
    - Subsets of households with all related choices,
    - Certain travel dimensions with all households considered (like tour generation, destination choice, time-of-day choice, mode choice, etc).
- Steps in progressing through global iterations; for example, one can envision 6 global iterations with freezing additional 20% of households at each iteration, starting from the third iteration. More, specifically, this would result in the following strategy:
    - 100% of households simulated in the 1st iteration,

  - 100% of households in the 2nd iteration (i.e. all households are re-simulated),
  - 80% of households in the 3rd iteration (20% of households are frozen),
  - 60% of households in the 4th iteration (another 20% of households are frozen),
  - 40% of households in the 5th iteration (another 20% of households are frozen),
  - 20% of households in the 6th iteration (another 20% of households are frozen),
- Principles for choice of the frozen and re-simulated households:
  - Purely random (with some possible geographic stratification).
  - Based on some criterion that reflects on "unstable travel conditions"; for examples households/persons/tours/trips that are characterized by a high level of congestion would be better to re-simulate multiple times.

Gradual freezing is always effective. It does not mean, however, that a true fixed-point solution is achieved. It relies on the reasonability of the technical strategy that can be established only after multiple trials. In some cases, such as model application for FTA New Starts projects, certain travel dimensions can be fixed across all compared scenarios that simplify the choice of strategy.

Gradual freezing can also be combined with a gradual unfreezing of portions of agents in order to economize and reduce run time during early iterations when LOS variables are crude and unstable anyway. If accuracy can be partially sacrificed, there is no need to simulate each household in the list. Only a subset of household can be simulated with corresponding weights applied before the network simulation stage. This allows for a quick and reasonable build-up for LOS skims before the time-taking full-blown simulation is applied.

Analytical Discretizing of Probability Matrices

Analytical discretizing represents a method for converting fractional-probability outputs of choice models into "crisp" choices as an alternative to the Monte-Carlo technique.

Analytical discretizing has two major advantages over the Monte-Carlo technique:

- *Full replication* of the model outcome with fixed inputs, i.e. if we run the discretizing procedure several times with the same choice model, with fixed input variables, the results will be identical, while the Monte-Carlo technique will be characterized by inherent variability (so-called "Monte-Carlo error") of the results.
- *Logical elasticity* of the aggregate model outcome with respect to the input variables that is identical to the elasticity of the parent choice model. The Monte-Carlo technique does not guarantee logical elasticity, and fixing seeds for random number does not help in this respect. The expected responses of the parent choice model can be "eaten" by the Monte-Carlo error that will make the model response illogical.

It is interesting to note, that in real terms, discretizing in application is the opposite of estimation in choice modeling. In the estimation procedure, discrete outcomes are given in the form of observed choices, and fractional probabilities are generated in order to replicate the observed choices as closely as possible. In applying the discretizing procedure, the fractional probabilities are given by the core choice model and "crisp" choices are generated in order to replicate the modeled probabilities as closely as possible. Thus, discretizing can be thought of as restoring the observations that would be most plausible for the given choice probabilities provided by the applied models.

We introduce the following notation:

$n \in N$	=	observations in the model estimation, realizations in the model application,
$i \in C_n \in C$	=	choice alternatives available for each observation / realization taken from universal set of alternatives,
$\delta_{in} = (0,1)$	=	Boolean indicator on the observed / modeled choice of alternative for each observation / realization.
$P_n(i)$	=	modeled choice probability for each alternative and observation / realization.

It is assumed that in both estimation and application of the model the "crisp" choices and fractional probabilities are subject to the logical constraints:

$$\sum_{i \in C_n} \delta_{in} = 1, \quad \sum_{i \in C_n} P_n(i) = 1 \text{ for all observations / realizations.} \tag{1}$$

The choice model **estimation** is done by maximizing the (log) likelihood function over choice probabilities (parameters of the choice model) while the observed choices are given:

$$\max_{\{P_n(i)\}} L = \sum_{n \in N} \sum_{i \in C_n} \delta_{in} \ln P_i(i). \tag{2}$$

The discretizing procedure in the model **application** is done by maximizing the entropy function over the "crisp" choices while the fractional probabilities are given by the core choice model:

$$\max_{\{\delta_{in}\}} E = -\sum_{n \in N} \sum_{i \in C_n} \delta_{in} \ln \frac{\delta_{in}}{P_n(i)} = \sum_{n \in N} \sum_{i \in C_n} \delta_{in} \ln P_n(i). \tag{3}$$

The discretizing approach considers the whole matrix of probabilistic outcomes of the core choice model and tries to find the structurally closest matrix of discrete numbers that also fits to the marginal totals of original matrix. Rows of the matrix correspond to observations / realizations. Columns of the matrix correspond to the choice alternatives. The marginal totals are readily interpreted. Row totals are all equal to 1 by the condition (1). The column totals correspond to the aggregate shares of alternatives. For a logit choice model with a full set of alternative-specific constants, the aggregate shares predicted by the model are equal to the observed shares for all alternatives:

$$\sum_{n \in N} P_n(i) = \sum_{n \in N} \delta_{in} = A_i. \tag{4}$$

It can be seen that likelihood optimization (2) and discretizing (3) refer to the same objective function under the same set of constraints (1, 4) but that the maximum is achieved with respect to the different subsets of variables. The optimization problem associated with the model estimation (maximize (2) given (1)) in the MNL case results in the convex problem that can be solved by the steepest descending method. The constraint (1) is guaranteed by the form of the choice model and constraint (4) is guaranteed by the alternative-specific constants. The optimization problem associated with the discretizing (maximize (3) given (1) and (4)) results in a linear programming (LP) problem of the so-called transportation type. An important property of the transportation problem is that discrete marginal totals guarantee a discrete solution.

Several properties of the discretizing procedure should be noted:

- Without constraints (4), maximization of (3) would result in the trivial choice of the alternative with maximum probability for each observation to be converted to the "crisp" choice.
- When the discretizing procedure is applied for the choice model outcomes for the same set of observations that was used in the model estimation, the value of the objective function can only be improved versus the likelihood achieved in the model estimation. The observed "crisp" choices form one of the possible solutions in the feasible region of the LP problem associated with discretizing.
- The better the core model is in terms of likelihood function (i.e. the closer is the modeled probabilities to the observed choices) then the closer the discretizing outcome will be to the observed choices. Indeed, the (log) likelihood function has a theoretical maximum value of zero that corresponds to an ideal model with probabilities equal to the observed choice indicators. The closer the estimated model is to this ideal, the less room left for the further improvement of the likelihood in the discretizing procedure.
- The discretizing procedure guarantees unbiasedness of the solution in an aggregate sense by virtue of the constraint (4).

The analytical discretizing procedure described here has been successfully applied for several sub-models of the Columbus (MORPC) AB Model – car ownership and activity allocation choices. In addition to the stability of outcomes that is beneficial for convergence to the network equilibrium, discretizing has also proved to be significantly superior in terms of replication of the observed individual choices when compared to a standard Monte Carlo method. In particular, when the MORPC car-ownership choice model was applied for the Household Survey sample of 5,560 households, a simple Monte-Carlo approach yielded about 50% of households simulated with the exact observed number of cars, with another 40% of households having one car less or more than the observed number, and another 10% having some larger deviations. The discretizing procedure improved these statistics to 60%, 35%, and 5%.

Averaging Methods to Achieve Equilibrium

Averaging is a universal tool that can be applied for continuous outcomes of any iterative process. If a fixed-point solution exists, an averaging strategy like Method of Successive Averages (MSA) will always find it, although it might require multiple iterations. If the equilibrium can be formulated as an optimization problem in view of the assumed simplicity of the demand functions, much more effective analytical procedures than MSA can be applied. However, if the demand model is too complicated to be written as an explicit optimization problem, MSA represents the only viable option to ensure a fixed-point solution. MSA has many possible technical variations. Any sequence of numbers S_k would suffice if it satisfies two basic conditions:

$$\lim_{k\to\infty} S_k = 0 \text{ and } \lim_{k\to\infty}\left(\sum_{m=1}^{k} S_m\right) = \infty . \tag{5}$$

In particular, the following MSA modifications are frequently used:

- $S_k = 1/k$ (literally corresponds to the term MSA),
- $S_k = 1/\sqrt{k}$ (may exhibit a faster convergence if the starting point is far off).

The essence of MSA is to smooth up the outcome of an iterative procedure where n denotes iteration, in the following way:

$$D_k = (1 - S_k)D_{k-1} + S_k\tilde{D}_k , \tag{6}$$

where:

$\tilde{D}_k$ = raw outcome of iteration k ,

D_k = smoothed outcome of iteration k that is fed back to the next iteration.

There are four possible ways in which MSA can be applied to ensure convergence of a demand model combined with network simulation, by averaging:

- Demand trip tables,
- Link volumes,
- Link times,
- Matrix LOS variables / skims.

As shown in **Figure 17**, these methods also can be effectively combined for ABMs, although it is not necessary for simple 4-step models.

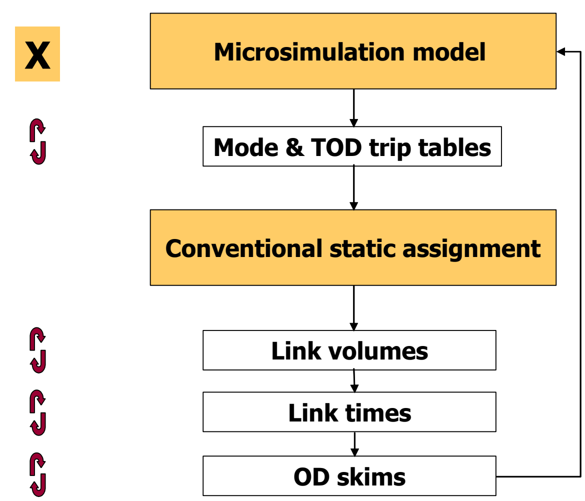


Figure 17: Equilibration and Feedback Options with ABMs

In many applications, a microsimulation demand model can be considered as a trip table generator followed by conventional assignment and skimming procedures. In typical planning situations, the model users are interested in aggregate outcomes of the microsimulation and do not usually track individual record details. With aggregation of microsimulation outcomes to Origin-Destination demand flows, equilibrium strategies with a microsimulation model are not principally different from equilibrium strategies applied with conventional models, with the exception that there is a different and more sophisticated way for generation of the raw trip table in each iteration. In terms of methods for averaging model outputs/inputs, the following should be noted:

- Original output of microsimulation procedure (individual household / person / tour / trip characteristics) cannot be meaningfully averaged between iterations since it represents a unique set of discrete values associated with a different list of agents at each iteration.
- Trip tables can be averaged in the same way as for conventional models.
- LOS skims can be averaged in the same way as for conventional models. There are three different technical ways to average LOS skims:
    - Directly average Origin-Destination skim matrices.
    - Average link times and then skim Origin-Destination LOS matrices.
    - Average link volumes, calculate corresponding link times, and then skim Origin-Destination LOS matrices (preferred method).

- Averaging link volumes is a better strategy in that it results in a faster convergence for over-congested networks, since link volumes are more stable than link travel times which are derived from an exponential function of link volumes.

Equilibrium Techniques and Experience from the New York AB Model

The New York Metropolitan Region represents a comparable example for CMAP regarding the challenge of equilibrating the modeling demand and network procedures.

- Very high levels of congestion that constitute a setting where a perfect convergence are difficult to achieve even with an aggregate model.
- Huge number of persons (20,000,000), size of the regional network (4,000 traffic analysis zones), and multi-class trip tables (7×4,000×4,000) that result in significant model run time, even for a single global iteration.
- Full variety of possible behavioral responses of travelers to changing LOS variables (switching modes, destinations, time-of-day) that contributes to instability and non-convergence.

The findings of numerous tests that comprise this research implemented with the New York model can be summarized in the following way:

- The most effective convergence strategy has been found to be a parallel MSA applied for both trip tables and link volumes producing synthetic LOS skims based on these link volumes for each subsequent iteration.
- A good level of convergence can be achieved with respect to network link volumes and aggregate county-to-county trip tables (28×28).
- In practical terms, the first 3 or4 global iterations result in a reasonable equilibrium state, while implementing additional 5 or 6 iterations brings only a marginal improvement.
- It is clearly seen that the further improvement in convergence cannot be achieved without overcoming average Monte-Carlo error, and that further refinement of the procedure is bound to an effective handling of Monte-Carlo variability through enforcement.

Examples of convergence statistics from these tests are shown in **Figure 18** for trip tables, and in **Figure 19** for highway link flows. Various feedback strategies were tested with 9 global iterations implemented for each strategy. Each global iteration included a run of the demand model and full set of period-specific assignments. The convergence statistics relate to the difference between outputs of two successive iterations; thus they can be calculated from the second iteration on. The following different feedback strategies are reported in the paper:

- *Direct* – full update of trip tables and LOS skims with no averaging.
- *MSA* – full update of trip tables and standard MSA for link volumes.
- *Root MSA* – full update of trip tables and "square root" MSA for link volumes.

- *MSA Trip* – parallel standard MSA applied for trip tables and link volumes.

The left-hand side of **Figure 18** relates to the AM period that spans 4 hours from 6:00 AM to 10:00 AM. The right-hand side of **Figure 18** relates to the Midday period that spans 6 hours from 10:00 AM to 4:00 PM. It should be noted that in contrast to many other metropolitan areas with a prominent peak of travel in the AM period, there is no principal difference in the hourly amount of travel and associated congestion between AM and Midday periods in the New York Metropolitan Region. Both periods (as well as the PM period) are characterized by significant congestion levels in the main corridors leading to and from the metropolitan core – Manhattan – and specifically on the bridges and tunnels connecting Manhattan to the rest of the region.

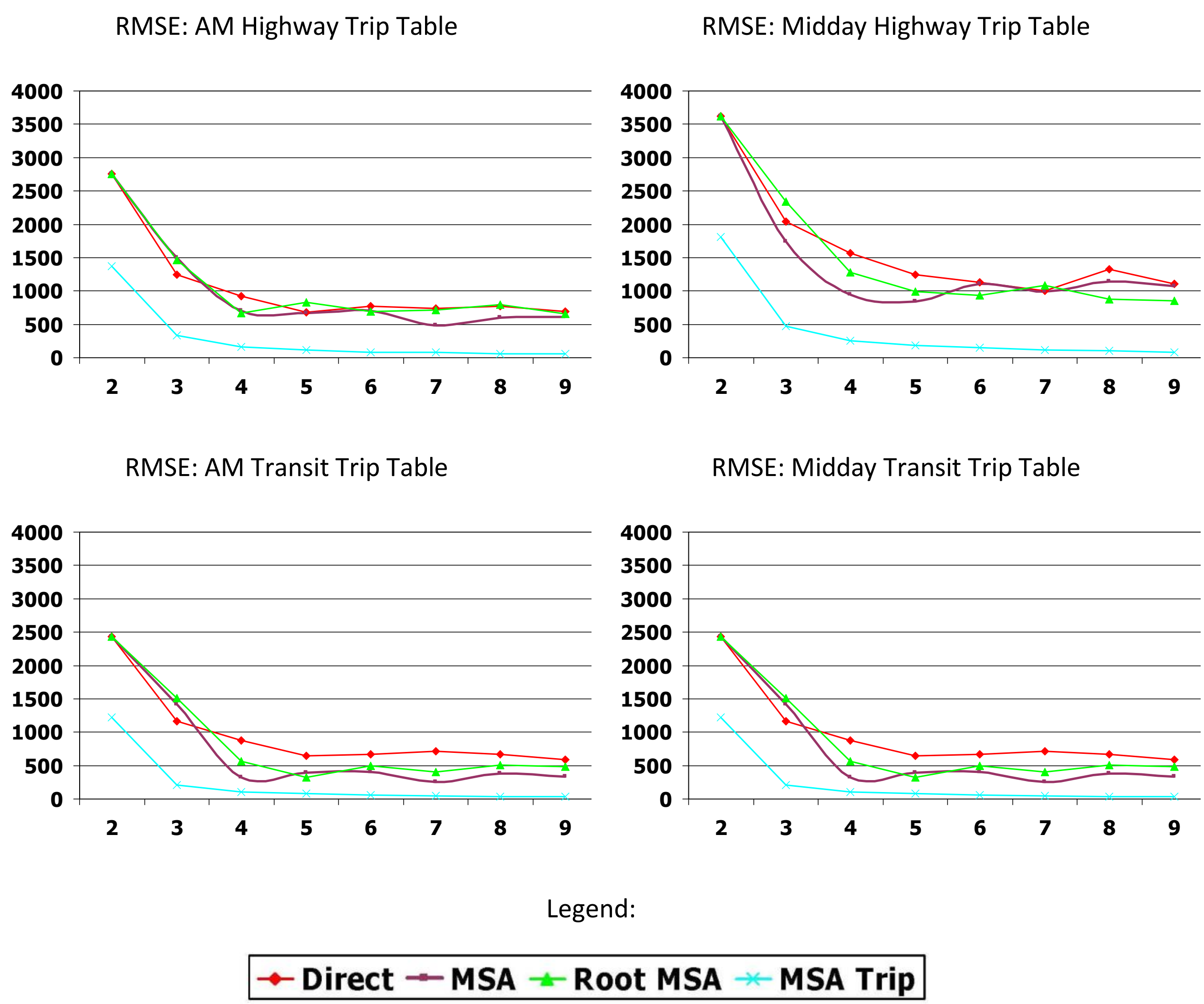
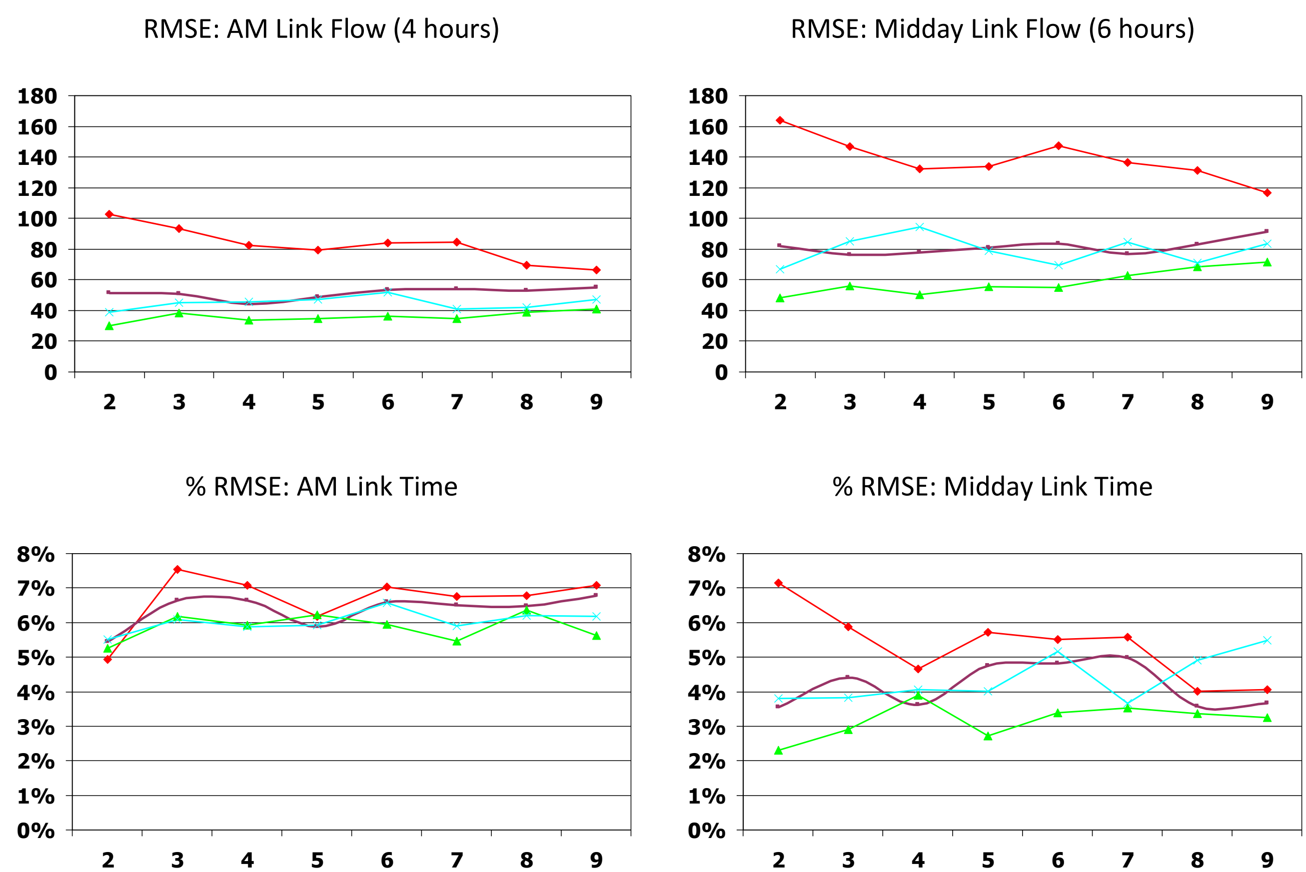


Figure 18: Convergence Statistics for Trip Tables (New York ABM)

Root of Mean Squared Error (RMSE) has been calculated for AM and Midday period highway and transit trip tables aggregated to the level of 28×28 county-to-county flows. The RMSE measure is for the difference between two successive global iterations. The results proved to be quite similar for both periods and for both modes. It can be seen that the strategy that includes a parallel MSA for trip tables and link volumes achieves the best, and practically absolute, convergences. All other strategies that do not include averaging of trip tables result in a certain level of non-convergence, though it is comparatively small (±500 trips for an aggregate county-to-county flow that is less than 2%).

Interestingly, the strategy of a parallel MSA for trip tables and link volumes even proved to ensure a reasonable level of convergence in terms of the detailed trip tables of 4000 × 4000 traffic zones. Taking into account the long run times associated with a significant increase in the number of global iterations (20-30) required, use of this strategy for planning forecasts, is probably impractical at this point in time, but could become feasible as model application code and computer technology continue to improve.



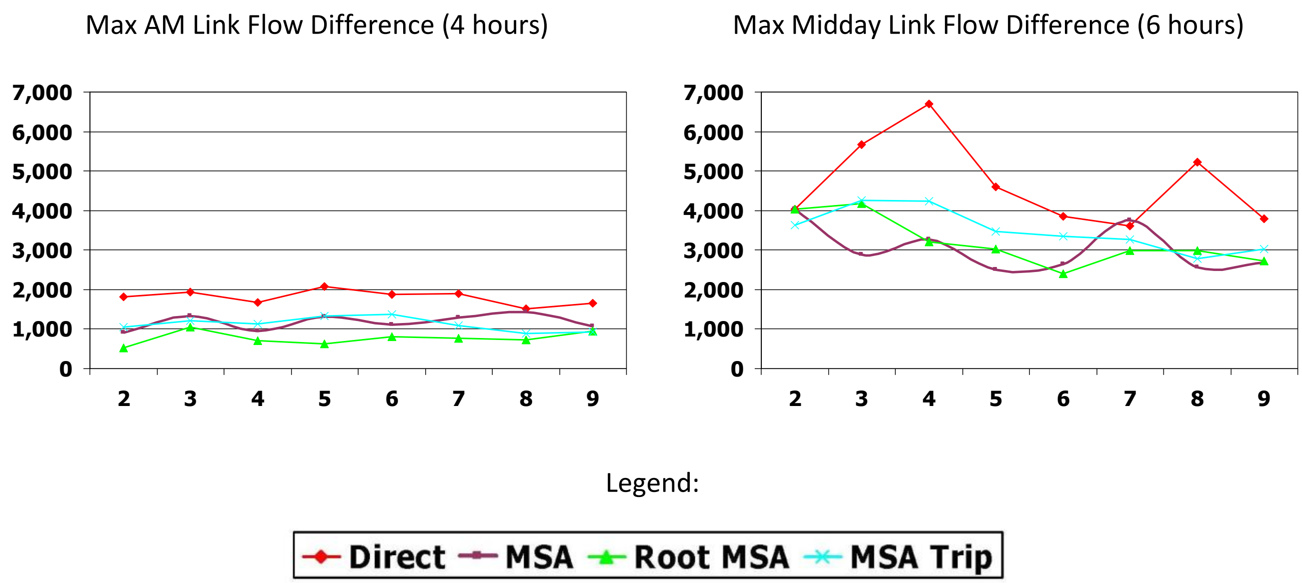


Figure 19: Convergence Statistics for Highway Link Flows and Times

The left hand side and right-hand side of **Figure 19** again relate to the AM and Midday periods consequently. The upper pair of graphs corresponds to the absolute value of RMSA calculated for link flows, i.e. the average variation of flows between two successive iterations at the link level. It can be stated that there is no significant difference between the strategies with only a relative disadvantage of direct feedback and specifically for earlier iterations.

In general, the average link flow variation with all averaging methods tends to be quite stable across iterations and does not improve from iteration to iteration. However, it should be pointed out that the actual value of the fluctuation is quite small. It is about only 50 vehicles for the 4-hour AM period (equivalent to 12 vehicles per hour), and about 80 for the 6-hour Midday period (equivalent to 13 vehicles per hour). These absolute values are much smaller than any conceivable level of accuracy in the traffic assignment (route choice) itself, or day-to-day variation in the observed traffic counts.

The pair of graphs in the middle tier of the figure relate to the % RMSE (relative to the average value) for highway link times. Again, it can be stated that there is no significant difference shown between the strategies, with only a relative disadvantage of direct feedback. There are two major reasons for unavoidable fluctuations at the level of 5-6% for any of these strategies: one stems from Monte Carlo error, and the other from the highway assignment results themselves which are unstable for the over-congested AM period (and significantly congested Midday period), when even a small fluctuation of trip table may result in a significant change in route choice. This is, however, typical for all types of models using conventional static link-based network loading methods, and can only be diminished by very large (and often impractical) increases in the number of iterations of the user equilibrium assignment procedures.

In contrast to the statistics averaged across all links in the highway network that were presented before, the bottom pair of graphs shows the largest variation for a specific directional link (i.e. the worst individual link case). The critical links with the maximum variation correspond to the major bridges and tunnels around Manhattan that are highly congested in both the AM and Midday periods. In particular, the George Washington Bridge, with more alternative route options, proved to be a frequent contributor to this statistic. Again, it can be stated that there is no significant difference between the averaging strategies with only direct feedback falling out of range. It should be noted, however, that in relative terms this fluctuation is not that significant. The absolute values of 1,000 vehicles for the 4-hour AM period and 3,000 vehicles for the 6-hour Midday period correspond to less than 5% of the average traffic flow on this facility. Thus, the largest absolute deviations do not match the largest relative deviations and *vice versa*.

It is also interesting that for the critical facility, the Midday period is characterized by the large absolute fluctuations than AM period. This could be explained by a significant share of “through” traffic, as well as trucks / commercial vehicles on the critical facilities, that do not follow the classic AM/PM commuting peak pattern and rather exhibit a high level of Midday circulation trips.

Another interesting observation can be made when comparing the performance of different averaging strategies in terms of trip table convergence in **Figure 18** with link flow / time convergence in **Figure 19**. While the parallel MSA method applied for both trip tables and link volumes has shown an obvious advantage in terms of stabilizing the trip table, it seems it has brought no visible improvement with respect to link flow / time convergence. A technical detail that can help explain this is that while the trip tables from different iterations are compared after averaging, the link flows and volumes correspond to the assignment output before averaging. Yet another possible explanation is that 9 global iterations is not enough to ensure a level of stability of the trip table that would already manifest itself at the individual link level.

Also, an important fact is that link volumes and especially times, are sensitive to small variations in trip table and to the parameters of traffic assignment itself. For example the number of internal iterations of highway assignment was set to 50 in the model test runs, and it frequently stopped by this maximum, before closing on one of the prescribed gaps. This indicates a certain level of non-convergence of the assignment itself. The convergence patterns could probably change if the convergence was set on achieving the gap tolerance rather than on reaching the maximum iteration limit. However, this is the point where the practice dictates certain rules. It is necessary to keep the entire model system run within a reasonable time frame and traffic assignments take a lion share of runtime. It would be possible to run assignments with say 100 iterations on an exceptional basis for research purposes. Our intention with this paper, however, was to present the result corresponding to the practically adopted model run framework with all the limitations.

All said, it seems that a reasonable level of convergence with respect to link volumes and times can be easily achieved with a comparatively crude trip table, and within a few global iterations; however, the “final tuning” would require a significant number of global iterations in combination with a high precision in the assignment procedure itself.

Equilibrium Techniques and Experience from the Sacramento ABM

In the overall system design of an AB system, as with a trip-based model system, there is a cyclical relationship between network performance and trip demand: DaySim and the auxiliary trip models use network performance measures to model person-trips, which are then loaded to the network, determining congestion and network performance for the next iteration. The model system is in equilibrium when the network performance used as input to DaySim and the other trip models matches the network performance resulting from assignment of the resulting trips. Network performance for this purpose is times, distances, and costs measured zone-to-zone along the paths of least generalized cost.

The equilibration procedure employs equilibrium assignment iteration loops (assignment iterations) nested within iterations between the demand and assignment models (global iterations). This is similar to the nested iteration in many trip-based model systems.

In DaySim, since the unit of analysis is households instead of Origin-Destination pairs, DaySim need not simulate the entire synthetic population in a global iteration; it is able to *run a selected sample of the population*. Since its runtimes are long but proportional to the number of households modeled, early global iterations can be sped up by simulating small samples. DaySim is parameterized to allow the user to select the size of sample for the simulation, so that a variety of time-saving iteration schemes can be used. Current implementations often use a scheme in which, during the first two global iterations, DaySim simulates schedules for one of every 256 households in the population. Thereafter it doubles the sample size for each global iteration, until it uses the full population in the tenth and final global iteration. In this way, the entire model system gains the advantage of ten global iterations to enhance the achievement of convergence between demand and assignment, while only incurring the runtime equivalent of two DaySim runs on the entire population [*Bradley, Bowman and Griesenbeck, 2010; Bowman, Bradley and Gibb, 2006*].

DaySim also contains facilities for random number synchronization between simulation runs, in order to reduce some of the more problematic aspects of simulation error due solely to the process of simulating discrete (stochastic) choices rather than multiplying choice probabilities (as is done in aggregate four-step models). This facility ensures that for the simulation of any given individual's choices, if no travel conditions have changed for that individual between two simulation runs, then the software will predict exactly the same simulated choices for that individual, even if conditions and choices have changed for many other persons in the population. If used across the latter iterations of a simulation run, this facility allows more rapid convergence toward global equilibrium across iterations. If used across the runs for two different policy scenarios, this facility ensures that the differences in simulated behavior between the two runs is due primarily to the policy scenario changes introduced, rather than to random "noise" in the simulation process.

Understanding and Managing Microsimulation Model Output

One of the major advantages of a microsimulation ABM is that it results in a *complete activity diary for all residents of the modeled region*. It includes a wealth of activity/travel results with practically any custom report/query/visual possible. The model user has a large degree of freedom to decide about the reasonable level of details and focuses for a particular study. In this sense, the situation is very different from working with a 4-step model that would dictate a limited set of output formats embedded in the model structure and segmentation. In many, cases, processing of the AB output can be effectively implemented with additional software managing data bases and/or GIS software.

Below are some examples how different policy questions and projects can be analyzed and visualized by processing the ABM output. Most of the output formats and types of analysis presented below are unavailable with aggregate 4-step models. First of all, it should be noted that the AB output still allows for all type of traditional analysis associated with 4-step models as shown in **Figure 20**. In particular, final trip tables by travel purpose, mode, and time of day can be produced and any statistics like trip-length distribution can be calculated. However, there are many more interesting and meaningful types of data transformation available with an AB model output that can be produced by the user.

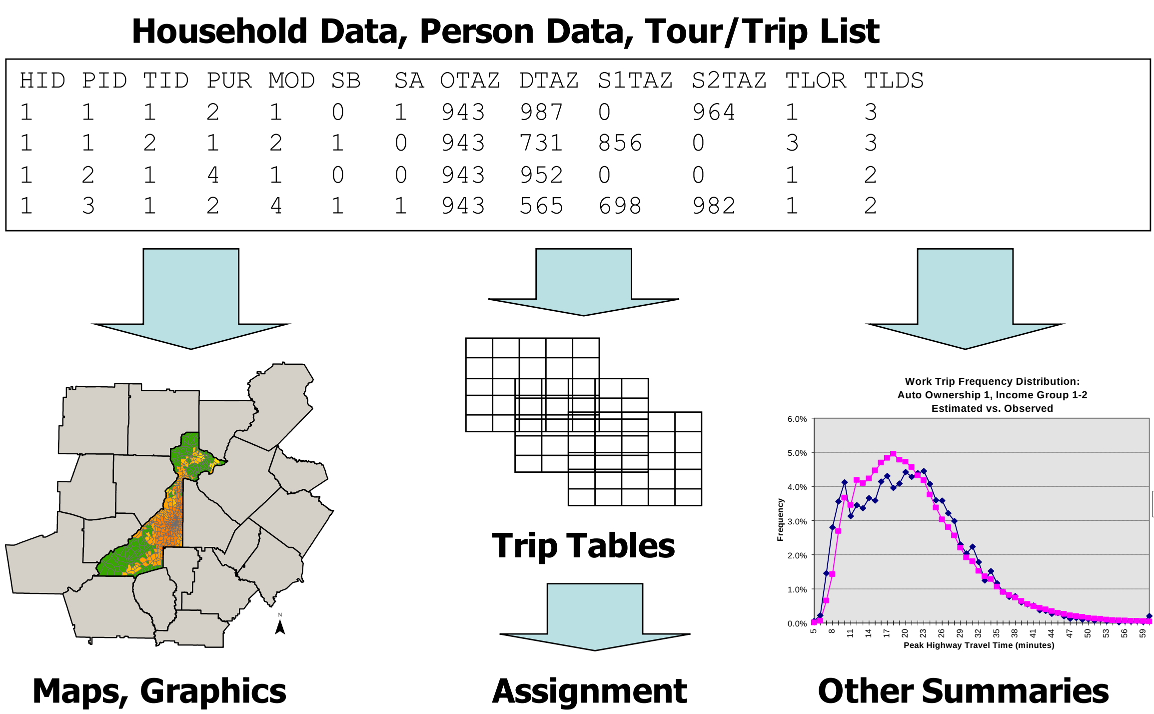


Figure 20: ABM Output Options

An interesting new type of analysis that can be implemented with an AB is individual travel time budgets as shown in **Figure 21**. It can be summarized by any household and/or person type and compared across different scenarios. This is an example of analysis that is unavailable with a 4-step model. This type of tabulation is possible with an ABM because all trips are linked to the corresponding individuals (person and households).

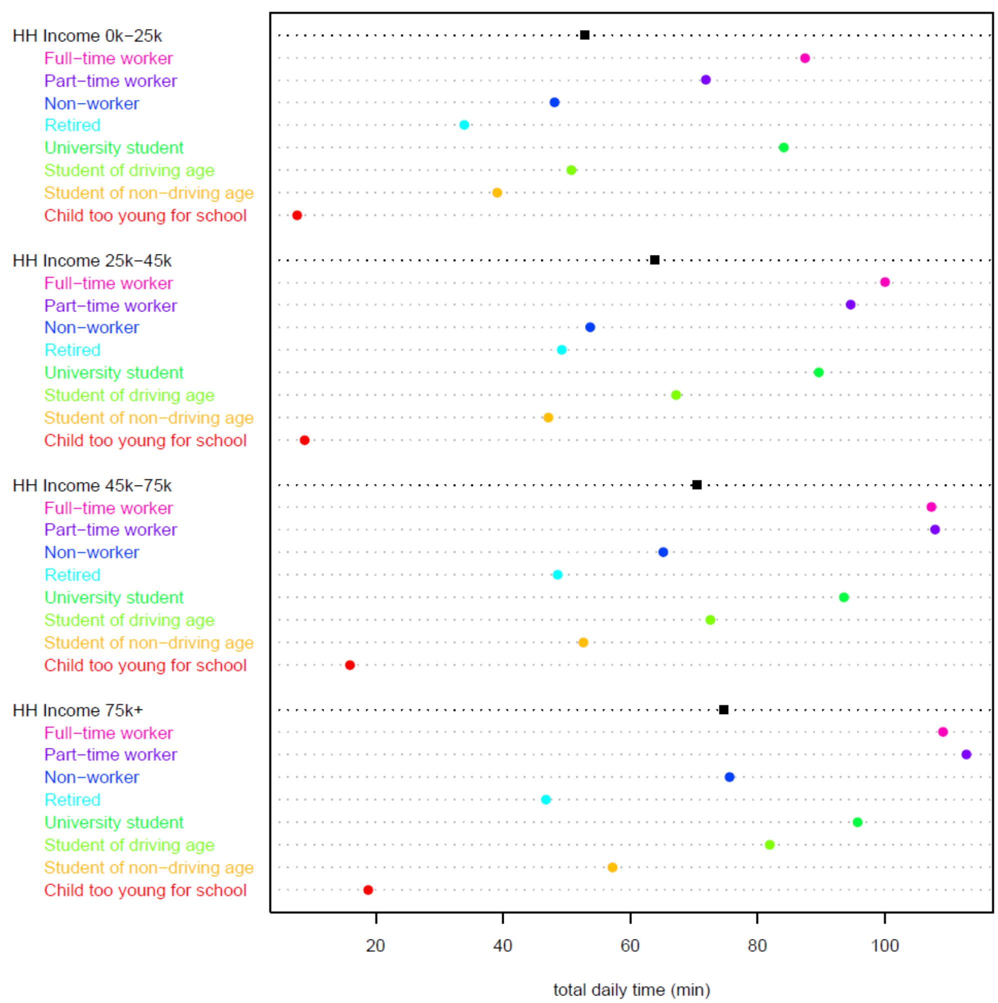


Figure 21: Time Spent Travelling by Household Income and Person Type

Planning interest in ABs was largely spurred by highway pricing studies where the limitations of 4-step models were exposed. For highway pricing studies, the following different aspects of analysis (and corresponding output formats from the AB) may be needed:

- *Toll Policy*: What is the best pricing form, toll rate, and toll schedule?
- *Revenue Generation*: How many transactions? Can test various discount policies (bridge toll-payers, residents of cordon, low-income, maximum tolls etc).
- *Equity Analysis*: Travel patterns of low-income households and zero-car households.
- *Congestion Mitigation*: Transit and Auto Speeds, Vehicles Hours of Delay, etc by corridor, area, or facility type.
- *Effective Revenue Reinvestment*: Responsiveness to various transit packages.

Another interesting new aspect that is beyond the reach of a 4-step model but has become possible with microsimulation AB modeling relates to evacuation / emergency planning. In the microsimulation process of ABM all individuals are tracked in space and time with the fine level of spatial resolution (TAZ or even smaller unit) and temporal resolution (30 min or even less). The AB model output can be easily processed to create population location maps during the day for each particular hour. The population locations can be segmented by main location types (home, workplace, school, other) if necessary. This is important information for planning evacuation measures by time-of-day.

The wide variety of options for processing AB output data has led to a reconsideration of the technical skills required from the model end-user at CMAP (i.e. person who applies the model and analyses the results). While for working with a 4-step model it is generally enough for a user to be proficient with the transportation package (EMME for CMAP) that is used to produce assignment and trip table summaries, processing of the AB microsimulation output requires proficiency with one of the general purpose data bases (Access, SQL, or Foxpro) and/or statistical packages like SAS or SPSS. The AB output contains millions of records with hundreds of data items organized by households, persons, tours, and trips that all have unique IDs used as the key fields in all output files. This type of data cannot be processed directly in basic desktop applications; more powerful tools are needed although eventually some processed data files can be customized in more traditional display applications. A powerful tool combination also includes a GIS platform that can be employed for mapping the output as well as using specific GIS engines for additional data processing. This means that in addition to proficiency with the transportation software package used for assignment and skimming, the end-user needs technical skills with data processing and GIS.

An alternative option that avoids the need in technical proficiency at the user end is to program a certain number of reports and summaries that would be routinely generated with each model run. It is interesting to note that early AB adopter MPOs did not choose this “easy life” option and preferred to have a full access to the lower-level data. The advantage of this approach is obvious: the user can customize any specific output format and extract statistics of interest without the programmer help. We strongly recommend CMAP to consider this approach as the major one.

Effective Software and Hardware Solutions

One of the key computational requirements in developing a successful AB system for use at an MPO is an acceptable run time. If the run time is too long, then the model's relevance may suffer because it simply cannot provide results fast enough to keep up with the policy and planning questions being asked of it. In terms of the region size and complexity comparable with the Chicago Metropolitan Region, two recently developed ABs provide good comparisons – the Atlanta AB developed for the Atlanta Regional Commission (ARC) and San-Francisco Bay Area Metropolitan Transportation Commission (MTC). The ARC modeled region has a 5-million population in 20 counties divided into about 4,000 TAZs. The MTC modeled region has 8-million population in 9 counties divided into about 3,000 TAZs. Both models, in term of their conceptual structure, belong to the CT-RAMP (**C**oordinated **T**ravel **R**egional **A**ctivity **M**odeling **P**latform) family of ABs that includes four members already in practice (Columbus, Lake Tahoe, Atlanta, and Bay Area) and three other members at different stages of development (San-Diego, Phoenix, and Jerusalem).

At the Atlanta Regional Commission (ARC) and the Metropolitan Transportation Commission (MTC), an acceptable run time is an overnight run time, i.e. 16 hours. In an effort to deliver acceptable run times, the ARC and MTC ABs were implemented with advanced software implementation strategies such as distributed/threaded processing, sampling, a shared software base, and others. Both ABs are based on the same software platform – Common Modeling Framework described below. In both cases, powerful computer clusters were used of 3-4 dual quad cores.

In addition to these two models that are presented with the maximum level of details, we also present software implementation issues, run times, and hardware specifications for some other models including SacSim and Columbus (MORPC) ABM. SacSim represents a different conceptual structure of AB compared to CT-RAMP and different software approach compared to the Common Modeling Framework. The Columbus (MORPC) ABM is a member of the CT-RAMP family but applied for a much smaller region of 1.4 million population compared to the metropolitan regions of Atlanta, Bay Area, and Chicago.

Example of Common Modeling Framework

Concept of Common Modeling Framework and Main Libraries

A flexible application programming structure is the Common Modeling Framework (CMF). CMF development for the recent years has been largely driven by the need to construct travel ABs although most of the CMF components are generic in nature and are frequently used to build 4-step models, freight models, and land-use models. The CMF is an object-oriented class library written in the Java programming language. Groups of classes with a similar functionality are grouped into "packages" that can be imported and used in any model development project.

The choice of Java as the programming language was based on the following considerations:

- Java is a fully Object-Oriented Programming (OOP) Language,
- Java is relatively easy to learn and use,

- Java encourages good software design,
- Java natively supports multi-threading,
- Java is computer-architecture-neutral (the compiled code is executable on many processors and platforms).

The CMF is not sold as commercial software, but is protected by the Apache License. This is an open-source software license that guarantees the CMF and all modifications to the CMF will remain open-source and free of charge. In other words, the CMF code cannot be sold and the source code is available in order to foster understanding of and improvement to the underlying algorithms.

A brief description of some of the packages that currently exist within the CMF follows:

- Matrix package: Most current tour-based models continue to rely on commercial transportation modeling software to develop mode-specific LOS matrices (skims) and for highway and transit assignment. The matrix package was developed specifically to allow the representation of skims in memory for rapid access and to allow read/write functionality to/from all major transportation planning packages. This package also includes classes for efficient memory management of sets of sparse matrices (such as transit) and for software-neutral disk representations of compressed matrices using zip format. Finally, an n-dimensional matrix class was developed to allow the representation and iterative proportional fitting of n-dimensional distributions of households, or any other data, primarily for use in the generation of synthetic populations.

- Model package: This package provides classes for the construction and application of discrete-choice models. It uses an interface class to represent alternatives within a discrete-choice model (i.e., any object can be an alternative in a model as long as it implements the alternative interface). It can be used to create and solve both multinomial and nested logit models. This feature is particularly useful for more advanced tour-based models, where the decision maker and the alternatives can vary from model to model (e.g., a fully joint tour can “choose” which household members will participate in it in one model, and a traveler can choose which TAZ to travel to in another). The package also contains a class to apply and solve “special event” models for use in both tour-based and advanced trip-based model systems.

- Calculator package: A flexible user interface has been created for the specification of utility equations used in discrete-choice models. The Utility Expression Calculator (UEC) class and supporting infrastructure relies on Microsoft Excel spreadsheets that are used to specify input data and utility expressions for any number of alternatives. A discrete-choice model is built using the model package, a UEC is constructed, and a solve method is called to return an array of utilities. The class has built-in intelligence on data formats and efficient storage, a built-in text parser with all standard mathematical formulas (e.g., if/then, trigonometric functions, etc.) and has been optimized for performance. The UEC can also be passed objects with data members that can change based on the outcomes of previous models, as opposed to relying on disk representations of all data required for utility calculations. The package was developed specifically to relieve programmers from coding a seemingly infinite number of utility equations, allowing them to focus on overall software design and avoid bugs.

- DAF package: A distributed application framework (DAF) has been developed to allow the rapid deployment of distributed applications with a straightforward interface. DAF uses Java sockets to send messages between Java virtual machines running on the same physical machine, and standard TCP/IP networking protocol to send messages between separate machines. Messages are sent between queues, queues are associated with tasks, and tasks are associated with the machine addresses on which they run. The messages are identified by a message header, and usually contain some data on which the tasks have been programmed to process. For example, an array of household objects might be sent to a task that has been programmed to run some set of models on the household objects it receives. The messages are completely flexible in terms of the data they contain and the actions that should be taken when they are received. A simple text properties file allows the user to set up the distributed application environment based on the number of machines available and the number of tasks to be initialized on each machine (e.g., usually one per processor). This package was developed specifically to take advantage of relatively inexpensive and widely available computing hardware to solve large-scale simulation problems, and has proven invaluable in implementing the TLUMIP (integrated Transportation & Land Use Model Improvement Program for the State of Oregon), MORPC, ARC, MTC, and SANDAG models.

Model Package

This package is specifically designed for effective implementation of discrete choice models of the logit type. Most of the components of an advanced travel model have a form of a choice model. In particular, an AB model represents a chain of 15-20 choice models (each of them being applied to multiple population and travel segments) of different levels of complexity that are applied at either household or person or tour/trip level. Effective programming implementation of a choice model that is going to be applied for millions of individuals and thousands of choice alternatives, include calculation of utilities for each alternative (that requires reading all variables and combining them with coefficients), calculation of choice probabilities (most frequently using logit model structure), and generation of a "crisp" choice (by Monte-Carlo or other technique). For each of these steps, recommendations for the most effective computing methods are made below. It is worth noting that utility calculation in majority of cases proves to be by far the most time-consuming component as well that a major reason for bugs in model implementation. An effective and user-friendly Utility Expression Calculator (UEC) applied in several ABs in practice is presented below in detail. An example of code used to implement a logit discrete choice model (mode choice) in CMF is shown in **Figure 22**.

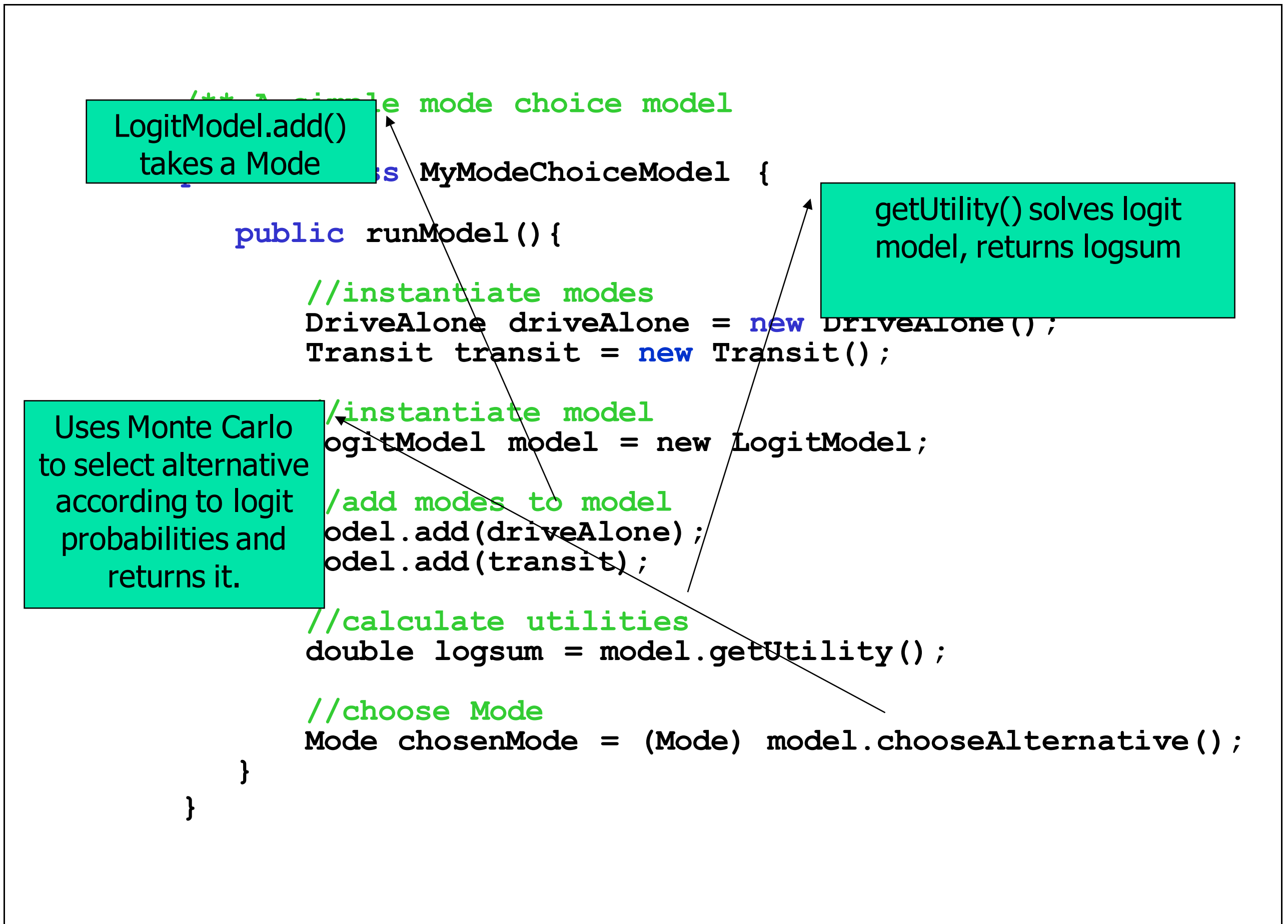


Figure 22: CMF Tools – Model Package Example

In the example, the required mode choice model functionality is encapsulated in the *MyModeChoiceModel* class that has *runModel* method. In this method, all choice alternatives (i.e. modes are instantiated first as objects. Then, an object of the *LogitModel* class is created and the modes are added to it using *LogitModel.add* method. After that the most time taking operation of utility calculation for each mode is implemented by calling *LogitModel.getUtility()* method that invokes the UEC package functionality described in the next sub-section. This method also returns the composite utility of the entire choice set, i.e. logsum. Finally, the mode is chosen by using *LogitModel.chooseAlternative* method. In the microsimulation framework, application of choice models is different from aggregate models and involves two steps as described in the section on microsimulation techniques above. At the first step, choice probabilities are calculated (as for an aggregate model). At the second step, that is unique to microsimulation, a single alternative is selected based on the choice probabilities using the Monte-Carlo method (independent random number draw) or some other method where a number between 0 and 1 is generated. In many instances, the number is generated for the given object once and then preserved for multiple model runs to ensure direct compatibility across compared scenarios.

Calculator Package

ABs typically utilize many logit choice models, some with a large number of alternatives. Traditional software relies on hard-coded utility equations that may take up to 70% of the actual code text. This is inefficient since the programmer becomes responsible for coding utility equations instead of focusing on the software architecture. This also proves to be extremely inflexible because it requires programmer to change equations and recompile the code each time if the model parameters change (for example, as the result of model calibration). This is also imperfect in a sense that only one person typically reviews equations, which increases probability of bugs. Utility Expression Calculator (UEC) was developed to overcome these limitations.

The UEC is a Java package that reads and interprets an Excel workbook containing a logit model specification and its inputs. The UEC solves the utility equations for a given decision-maker. The UEC "opens up" the model specification – anyone can edit the spreadsheets, change inputs and parameters, check that the model is properly specified, adjust model parameters during the calibration process without changing and recompiling the code, etc. The UEC also allows for user specification of all input data types and files as shown in **Figure 23**.

Table Data: CSV files of zonal, household, or person data

Table Data						
No	Type	Format	File			
1	zone	csv	%project.directory%/Inputs/retailAccessibility.csv			
Matrix Data						
No	Token	Format	File	Matrix	Group	Index
1	OP_SOV_TIME	http://localhost:6	%project.directory%/Outputs/SOVFFM05.SKM	time		
2	OP_SOV_DIST	http://localhost:6	%project.directory%/Outputs/SOVFFM05.SKM	distance		
3	OP_WLKPRE_FWT	http://localhost:6	%project.directory%/Outputs/offpre.skm	iwait		
4	OP_WLKPRE_XWT	http://localhost:6	%project.directory%/Outputs/offpre.skm	xwait		
5	OP_WLKPRE_WLK	http://localhost:6	%project.directory%/Outputs/offpre.skm	walkt		
6	OP_WLKPRE_LOCIVT	http://localhost:6	%project.directory%/Outputs/offpre.skm	locivt		
7	OP_WLKPRE_RAILIVT	http://localhost:6	%project.directory%/Outputs/offpre.skm	railivt		
8	OP_WLKPRE_XBUSIVT	http://localhost:6	%project.directory%/Outputs/offpre.skm	xbusivt		

Matrix Data: Trip tables or level-of-service skims in zone-zone format (TPPLUS, TRANSCAD, EMME2, and/or BINARY formats)

Figure 23: Calculator Package – Data Page (View 1)

Special features of UEC relate to treatment of LOS matrices (and associated I-O) that represent one of the most time-taking components of the ABM implementation – see **Figure 24**. All matrices must be consecutively numbered. Each matrix is given a meaningful token name by the programmer/user that is further used to refer to the matrix in the model specification page. The source file from which a matrix is read is specified by the path and filename with a convenient reference to the global project directory folder (set in the model properties file). The matrices of different formats used by different software packages (EMME, TransCAD, Cube, etc) can be read directly without time-taking exporting into the CSV or ASC format. In a case of several matrices stored in the same file (like EMME DataBank or TransCAD *.MTX files), the matrix number or name is used as the identifier.

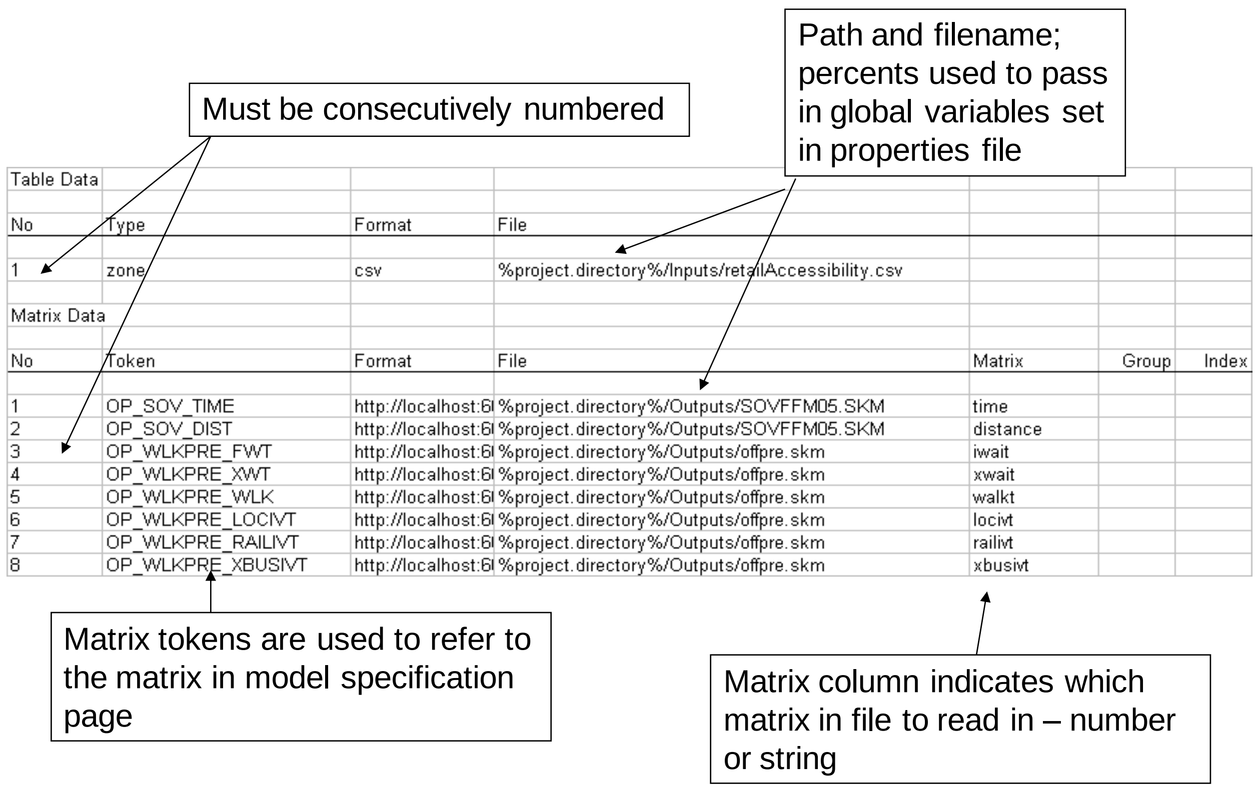


Table Data						
No	Type	Format	File			
1	zone	csv	%project.directory%/Inputs/retailAccessibility.csv			
Matrix Data						
No	Token	Format	File	Matrix	Group	Index
1	OP_SOV_TIME	http://localhost:6	%project.directory%/Outputs/SOVFFM05.SKM	time		
2	OP_SOV_DIST	http://localhost:6	%project.directory%/Outputs/SOVFFM05.SKM	distance		
3	OP_WLKPRE_FWT	http://localhost:6	%project.directory%/Outputs/offpre.skm	iwait		
4	OP_WLKPRE_XWT	http://localhost:6	%project.directory%/Outputs/offpre.skm	xwait		
5	OP_WLKPRE_WLK	http://localhost:6	%project.directory%/Outputs/offpre.skm	walkt		
6	OP_WLKPRE_LOCIVT	http://localhost:6	%project.directory%/Outputs/offpre.skm	locivt		
7	OP_WLKPRE_RAILIVT	http://localhost:6	%project.directory%/Outputs/offpre.skm	railivt		
8	OP_WLKPRE_XBUSIVT	http://localhost:6	%project.directory%/Outputs/offpre.skm	xbusivt		

Figure 24: Calculator Package – Data Page (View 2)

In the ABM application (like in any travel model application), specification of various runs is frequently bound to using a different set of LOS skims while all other model inputs (like population, employment, and land use) stay the same at least for a series of runs. In this case, the end-user task is to locate these different sets of skims in different folders and to properly refer each model run to the corresponding set of skims.

A significant improvement of runtime was achieved by special treatment of sparse matrices, i.e. matrices with a large number of randomly distributed zero cells that is a common characteristic of transit LOS skims as shown in **Figure 25**. The runtime savings are achieved by grouping the skims into sets (by transit modes and sub-modes) that share the same index matrix. The index matrix determines whether the Origin-Destination pair is connected or not for the given mode (typically by checking if in-vehicle time for this mode is positive). If the Origin-Destination pair is not connected by the given mode (in-vehicle time is equal to zero), the rest of skims becomes irrelevant since it is filled with zeros (or special characters depending on the transportation package convention). Thus, the rest of skims are not read which ensures a significant economy on run time.

Sparse matrices can be grouped for compression:

- Each group is a set of skims, such as “Peak walk-local”
- Each matrix group must have an index matrix, which determines whether the zone-pair is connected or not (typically in-vehicle time for the primary mode is used)

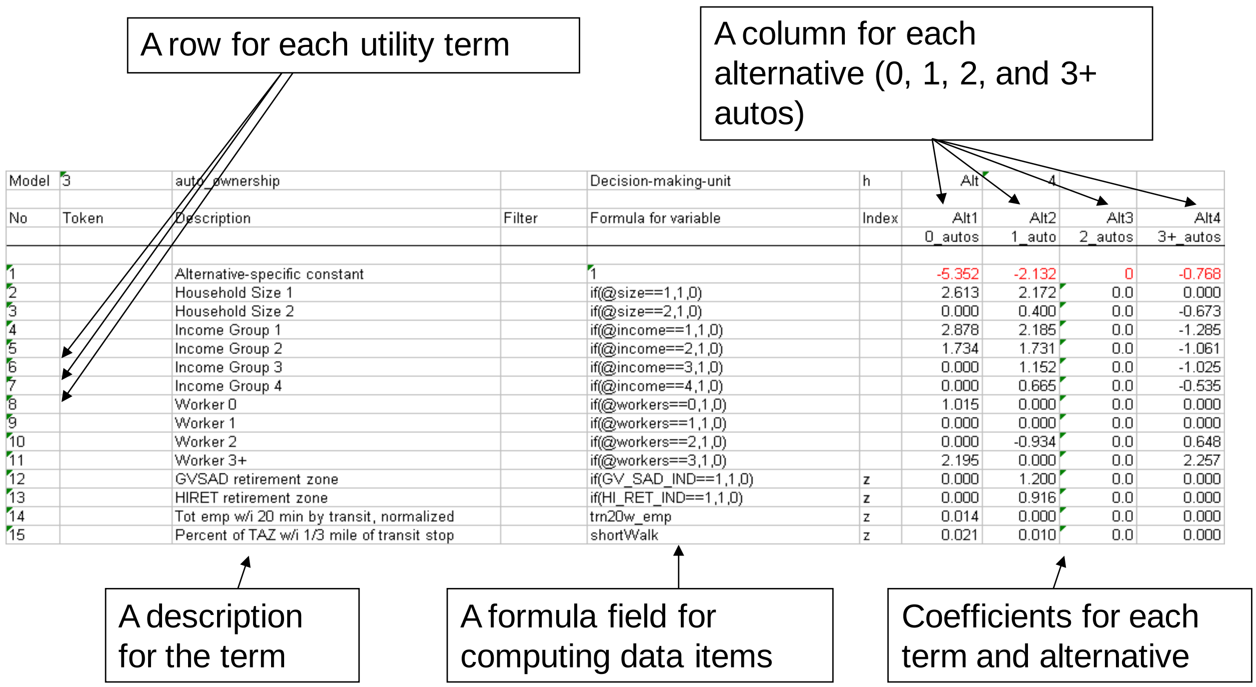
Matrix Data

No	Token	Format	File	Matrix	Group	Index
1	AUTO_TIME	BINARY	%project.directory%outputs/binary/skim_pm_3.binary	1		
2	AUTO_DIST	BINARY	%project.directory%outputs/binary/skim_pm_2.binary	1		
3	ZONE_DIST	BINARY	%project.directory%outputs/skims/StraightLineDistanceCentr	1		
4	WLKBUS_FWT	BINARY	%project.directory%outputs/binary/pmetrnwtw_1.binary	1	1	
5	WLKBUS_XWT	BINARY	%project.directory%outputs/binary/pmetrnwtw_2.binary	1	1	
6	WLKBUS_ACC	BINARY	%project.directory%outputs/binary/pmetrnwtw_3.binary	1	1	
7	WLKBUS_AUX	BINARY	%project.directory%outputs/binary/pmetrnwtw_4.binary	1	1	
8	WLKBUS_EGR	BINARY	%project.directory%outputs/binary/pmetrnwtw_5.binary	1	1	
9	WLKBUS_IVT	BINARY	%project.directory%outputs/binary/pmetrnwtw_6.binary	1	1	1
10	WLKBUS_XIVT	BINARY	%project.directory%outputs/binary/pmetrnwtw_8.binary	1	1	
11	WLKBUS_FAR	BINARY	%project.directory%outputs/binary/pmetrnwtw_11.binary	1	1	
12	WLKBUS_BRD	BINARY	%project.directory%outputs/binary/pmetrnwtw_12.binary	1	1	
13	PNRBUS_FWT	BINARY	%project.directory%outputs/binary/pmtrnwta_1.binary	1	2	
14	PNRBUS_XWT	BINARY	%project.directory%outputs/binary/pmtrnwta_2.binary	1	2	
15	PNRBUS_DRV	BINARY	%project.directory%outputs/binary/pmtrnwta_3.binary	1	2	
16	PNRBUS_AUX	BINARY	%project.directory%outputs/binary/pmtrnwta_4.binary	1	2	
17	PNRBUS_EGR	BINARY	%project.directory%outputs/binary/pmtrnwta_5.binary	1	2	
18	PNRBUS_IVT	BINARY	%project.directory%outputs/binary/pmtrnwta_6.binary	1	2	2
19	PNRBUS_XIVT	BINARY	%project.directory%outputs/binary/pmtrnwta_8.binary	1	2	
20	PNRBUS_FAR	BINARY	%project.directory%outputs/binary/pmtrnwta_11.binary	1	2	
21	PNRBUS_BRD	BINARY	%project.directory%outputs/binary/pmtrnwta_13.binary	1	2	
22	KNRBUS_FWT	BINARY	%project.directory%outputs/binary/pmtrnwtk_1.binary	1	3	
23	KNRBUS_XWT	BINARY	%project.directory%outputs/binary/pmtrnwtk_2.binary	1	3	
24	KNRBUS_DRV	BINARY	%project.directory%outputs/binary/pmtrnwtk_3.binary	1	3	
25	KNRBUS_AUX	BINARY	%project.directory%outputs/binary/pmtrnwtk_4.binary	1	3	
26	KNRBUS_EGR	BINARY	%project.directory%outputs/binary/pmtrnwtk_5.binary	1	3	
27	KNRBUS_IVT	BINARY	%project.directory%outputs/binary/pmtrnwtk_6.binary	1	3	3
28	KNRBUS_XIVT	BINARY	%project.directory%outputs/binary/pmtrnwtk_8.binary	1	3	
29	KNRBUS_FAR	BINARY	%project.directory%outputs/binary/pmtrnwtk_11.binary	1	3	
30	KNRBUS_BRD	BINARY	%project.directory%outputs/binary/pmtrnwtk_13.binary	1	3	

Figure 25: Calculator Package – Matrix data

In addition to sparse transit modes and sub-modes this technique is also beneficial for building specific highway skims for toll road users or HOV/HOT lane users (if the corresponding facilities constitute a small percentage of links in the regional network).

The main UEC page that contains the utility specification for each choice alternative is shown in **Figure 26**. Car ownership choice is an example of a simple choice structure with a limited number of alternatives that is a good illustration tool for different functions of UEC. Each row represents a utility term that is either a variable or some transformation of variables. Each choice alternative has a column where the utility coefficients are specified.



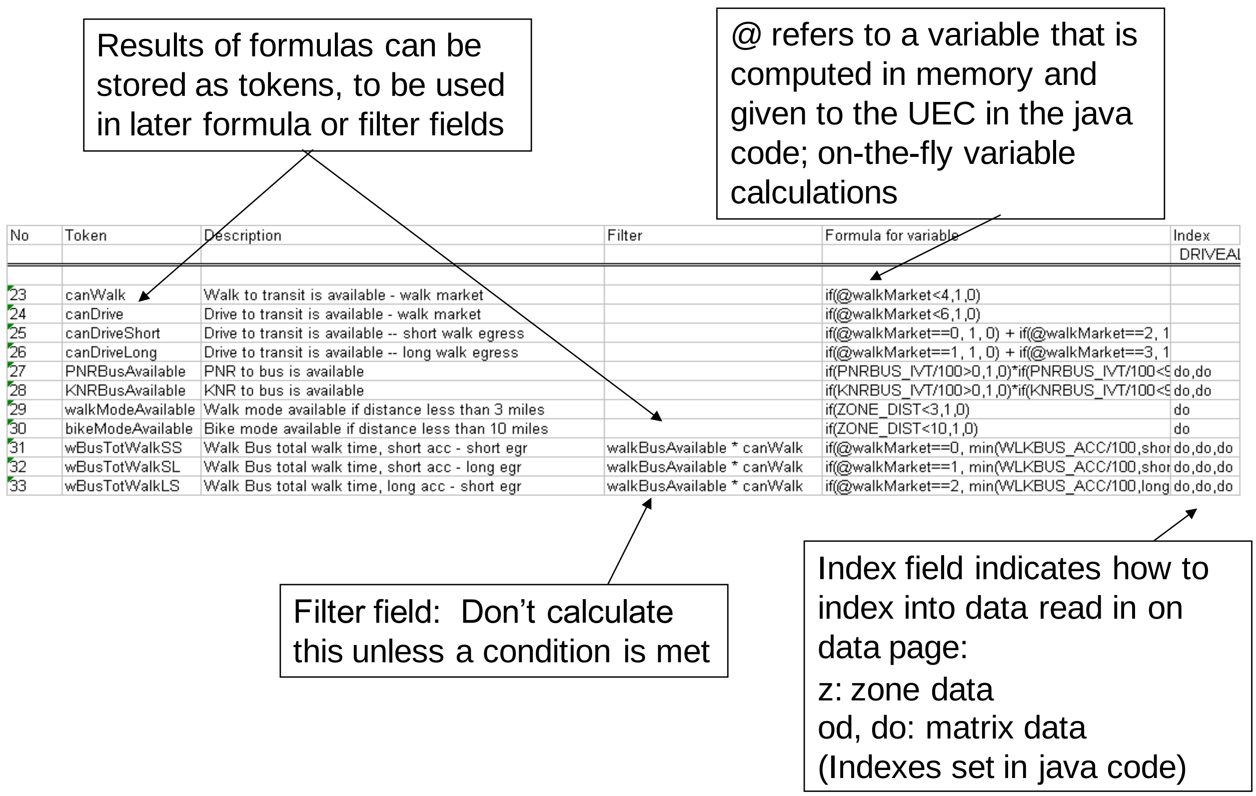
Model	3	auto_ownership		Decision-making-unit	h	Alt	4		
No	Token	Description	Filter	Formula for variable	Index	Alt1	Alt2	Alt3	Alt4
						0_autos	1_auto	2_autos	3+_autos
1		Alternative-specific constant		1		-5.352	-2.132	0	-0.768
2		Household Size 1		if(@size==1,1,0)		2.613	2.172	0.0	0.000
3		Household Size 2		if(@size==2,1,0)		0.000	0.400	0.0	-0.673
4		Income Group 1		if(@income==1,1,0)		2.878	2.185	0.0	-1.285
5		Income Group 2		if(@income==2,1,0)		1.734	1.731	0.0	-1.061
6		Income Group 3		if(@income==3,1,0)		0.000	1.152	0.0	-1.025
7		Income Group 4		if(@income==4,1,0)		0.000	0.665	0.0	-0.535
8		Worker 0		if(@workers==0,1,0)		1.015	0.000	0.0	0.000
9		Worker 1		if(@workers==1,1,0)		0.000	0.000	0.0	0.000
10		Worker 2		if(@workers==2,1,0)		0.000	-0.934	0.0	0.648
11		Worker 3+		if(@workers==3,1,0)		2.195	0.000	0.0	2.257
12		GVSAD retirement zone		if(GV_SAD_IND==1,1,0)	z	0.000	1.200	0.0	0.000
13		HIRET retirement zone		if(HI_RET_IND==1,1,0)	z	0.000	0.916	0.0	0.000
14		Tot emp w/i 20 min by transit, normalized		trn20w_emp	z	0.014	0.000	0.0	0.000
15		Percent of TAZ w/i 1/3 mile of transit stop		shortWalk	z	0.021	0.010	0.0	0.000

Figure 26: Calculator Package – Car Ownership Model Example

Each term (row) has a meaningful description that makes the entire model specification transparent and easy to manipulate. Token (i.e. an internal variable name) can be given to some terms that can be reused in the subsequent rows. A formula is frequently specified for each term rather than a single variable in order to avoid proliferation of variables that have to be prepared in the code or data page. For example, a single original data item – household income category that takes values of 1, 2, 3 and 4 can be used instead of four dummy indicators on each particular income category. Then, this data item is converted into four utility terms in the UEC formula field using *IF()* function. In the model application run, each UEC is read consecutively row after row and utility expressions are evaluated for each alternative as a sum of all row expressions multiplied by the coefficients.

Manipulating the UEC is a routine part of model calibration and application when local adjustments are to be made by the user to better match the sub-area or corridor targets. Most frequently, these adjustments relate to alternative-specific constants but can also include distribution K-factors and other variables.

The UEC package provides very useful methods for managing initial and derived variables that is essential for complicated models like tour mode choice as shown in **Figure 27**. In particular, once calculated variables can be tokenized and reused multiple times for different models.



No	Token	Description	Filter	Formula for variable	Index
					DRIVEAL
23	canWalk	Walk to transit is available - walk market		if(@walkMarket<4,1,0)	
24	canDrive	Drive to transit is available - walk market		if(@walkMarket<6,1,0)	
25	canDriveShort	Drive to transit is available -- short walk egress		if(@walkMarket==0, 1, 0) + if(@walkMarket==2, 1	
26	canDriveLong	Drive to transit is available -- long walk egress		if(@walkMarket==1, 1, 0) + if(@walkMarket==3, 1	
27	PNRBusAvailable	PNR to bus is available		if(PNRBUS_IVT/100>0,1,0)*if(PNRBUS_IVT/100<9	do,do
28	KNRBusAvailable	KNR to bus is available		if(KNRBUS_IVT/100>0,1,0)*if(KNRBUS_IVT/100<9	do,do
29	walkModeAvailable	Walk mode available if distance less than 3 miles		if(ZONE_DIST<3,1,0)	do
30	bikeModeAvailable	Bike mode available if distance less than 10 miles		if(ZONE_DIST<10,1,0)	do
31	wBusTotWalkSS	Walk Bus total walk time, short acc - short egr	walkBusAvailable * canWalk	if(@walkMarket==0, min(WLKBUS_ACC/100,shor	do,do,do
32	wBusTotWalkSL	Walk Bus total walk time, short acc - long egr	walkBusAvailable * canWalk	if(@walkMarket==1, min(WLKBUS_ACC/100,shor	do,do,do
33	wBusTotWalkLS	Walk Bus total walk time, long acc - short egr	walkBusAvailable * canWalk	if(@walkMarket==2, min(WLKBUS_ACC/100,long	do,do,do

Figure 27: Calculator Package – Mode Choice Model Example

Another important run-time savings relates to using the filter field. For each row, the filter field is evaluated first and if it is equal to zero the entire term is set to zero without further calculations. In addition to static variables read from the input files, the UEC program can use any situational variable that is computed in the microsimulation process and is ready in the memory. These variables are referred with a prefix @. The index field indicates how to find the appropriate element in the indexed data array. The most frequent indexed items relate to the zonal data and Origin-Destination matrix data. The corresponding indexes are set in the Java code and the user can use them to specify such data items as zonal densities / accessibilities and Origin-Destination travel times and cost. The matrix data can be transposed if necessary by using a reversed index that is especially convenient for coding tour-level models with bi-directional skims.

General Software Architecture

This section itemizes some key computer programming elements in the general software architecture associated with an advanced AB model. It is mostly based on the available documentation for the CT-RAMP family of ABs of which the Atlanta (ARC) and Bay Area (MTC) ABs are two recently completed model systems. In addition to that, we discuss certain features of the SacSim model system that represents a different approach.

CT-RAMP Example

As shown in **Figure 28**, the CT-RAMP software for the microsimulation components of the model has been co-developed for MTC, and relies on the Common Modeling Framework (CMF), a collection of Java libraries specifically designed for the implementation of disaggregate travel demand models. Both the ARC and MTC models utilize the CT-RAMP Java package, which contains model logic, choice model structure, and model flow, while utility equations and model inputs and outputs are specific to the ARC or MTC implementation and are contained in Utility Expression Calculator (UEC) files. These Excel-based files open up the models so the parameters, input filenames, etc can be easily accessed which helps prevent errors and makes the model equations fully accessible.

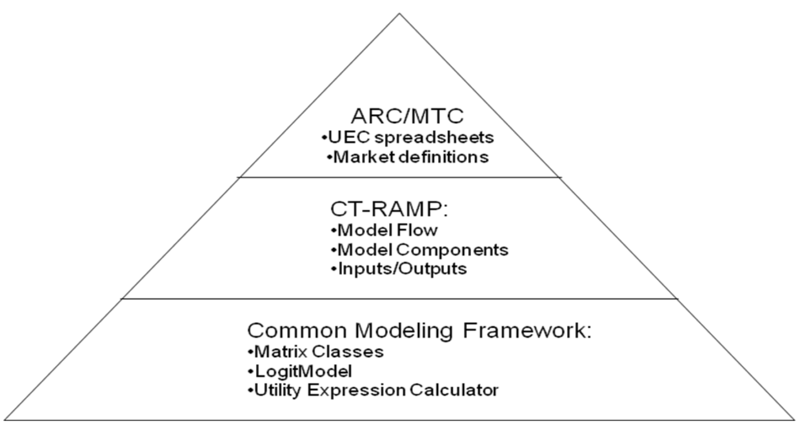


Figure 28: Disaggregate Travel Demand Model Software Components

As was discussed above, the CMF package provides a library of general classes and methods that are used to construct operational AB models. This part represents an open source and typically requires professional programming support. It is used by software architects and programmers in the software development process, but is not normally by the end user. The CMF is roughly comparable to the matrix constructors embedded in the software packages like EMME, TransCAD, or Cube.

A typical CT-RAMP model flow is comparable to the macro script of EMME or GISDK script of TransCAD that essentially defines the model system structure. This layer of the pyramid is also created by the software architects and programmers and may change or be tailored for a specific regional conditions, but is normally transparent to the median end user.

Finally each regional AB model calls a specific set of UECs and market segment definitions. The least transferable components are normally tailored for each region and include mode choice (an exact subset of relevant highway, transit modes, and non-motorized modes), sub-models for special events, and population synthesizer (list of controlled variables). In all cases, the actual parameters are specified in the UECs. This part is analogous to control files with parameters (or relevant parts of macro scripts) used in trip-based software packages. The top of the pyramid is completely open to the end-user will be maintained and updated by the agency modeling staff.

In the sections below we will specifically address some critical issues associated with AB model implementation, including general multi-threading and distributing tools as part of the CMF layer and more specific software-hardware setups as part of the CT-RAMP layer (or specific AB layer).

DaySim Example Version 3—Under Development

Following experience gained in the initial adaptations of DaySim for different users starting from the Sacramento, 2006 version, programming has begun on the newest version, which is designed to be more flexible and configurable for different projects, different levels of data availability, and different methods of network assignment/simulation. Major objectives include faster execution, application to different regions, a wider variety of temporal and spatial aggregation schemes, and easier adjustment of the model components and parameters using the application software itself. The impetus for several features of the new design is the use of DaySim as the AB component for the SHRP 2 (Strategic Highway Research Program 2) C10 project, which is aimed at creating an open-source, state-of-the-art ABM software system that can be integrated with advanced DTA and simulation methods [*Resource Systems Group, et al, 2010b*]. Several major aspects are identified in this effort including multi-threading and distributing (described below in the corresponding sub-section), flexible temporal and spatial disaggregation (described above in the corresponding sub-section), and integrated model specification and estimation (a somewhat unique approach that is described in the current sub-section).

DaySim includes many models, and each one has model-specific routines designed to implement logit choice models. These typically change somewhat from project to project, and thus need to be flexible and user-configurable. In the original versions of DaySim, the code to implement models was copied over from code used to generate model estimation control files (for estimation using ALOGIT). This required a combination of cut-and-paste, global search and replace editing, and Delphi functions that mimic the most commonly used ALOGIT data transformation functions. While this is an excellent approach in terms of avoiding errors and creating efficient code, it is less than ideal for a general-purpose program that can be easily adapted from project to project. DaySim Version 3 will include a series of model specification routines that make it easy for a non-expert programmer to easily change a model's specification for application or (re)estimation, while preserving the advantage of avoiding errors in transcribing model specifications into the application program.

For model application only, the list of required routines is relatively short. They include routines that construct a new model object, read in previously estimated model coefficients, set the model nesting structure, set the values of model variables and parameters, set alternative availability and simulate choices. These commands are the basis of a model "scripting" language that also includes syntax for various looping, logical and data access operations to specify a choice model. This model scripting approach will be familiar to users of network software packages such as CUBE, TransCAD and VISUM.

These and additional routines are used to automate model re-estimation. The code currently generates files to automate model estimation with the ALOGIT package, although it could easily be adapted to work with other model estimation packages as well. Re-estimation adds overhead to the model application routines, because some logic is relevant only for re-estimation, and not even needed if one is only simulating choices. All the routines are programmed so that DaySim incurs the extra overhead associated with model estimation only when it is run in "estimate" mode. Model estimation routines include those that label alternatives and parameters for user readability, identify the chosen alternative, and write data and control records needed by the program used to estimate the model coefficients.

Critical Issues and Time Taking Operations

Two components of an AB model take the lion share of the runtime: 1) location choices (primary tour destinations and intermediate stop locations) and 2) multi-class assignment and network skimming procedures. Both components are primarily sensitive to the region network size in terms of number of TAZs (or smaller spatial units like MGRA or parcel, if applied). In particular, each location choice model has to be applied for millions of objects. Usual workplace location is needed for each worker. Usual school location is needed for each student. The primary tour destination is needed for each tour. Intermediate stop locations are needed for each half-tour according to the stop frequency on this half-tour. Each particular choice involves an evaluation of thousands of alternatives at the TAZ level or dozens of thousands of alternatives at the MGRA level or hundreds of thousands of alternatives at the parcel level. In addition to that, a constrained location choice requires multiple iterations for calculation of shadow prices or some other techniques that also require additional run time. Highway assignment run times are exponentially proportional to the number of zones that, in general, prohibits any other level of spatial resolution except TAZ for these model components.

Different solutions are discussed in the subsequent sub-sections including such general solutions as parallel processing (multi-threading and distribution) that can be applied for both location choices in ABM and assignments, as well as some specific methods for location choices like sampling of zones, packeting, smart pre-calculations, etc.

Location Choices

An example of implementation of usual work and school location choices and associated time-taking components is shown in **Figure 29**. In this case, a large number of alternatives is combined with the necessity to implement multiple shadow pricing iterations that represents a computational challenge.

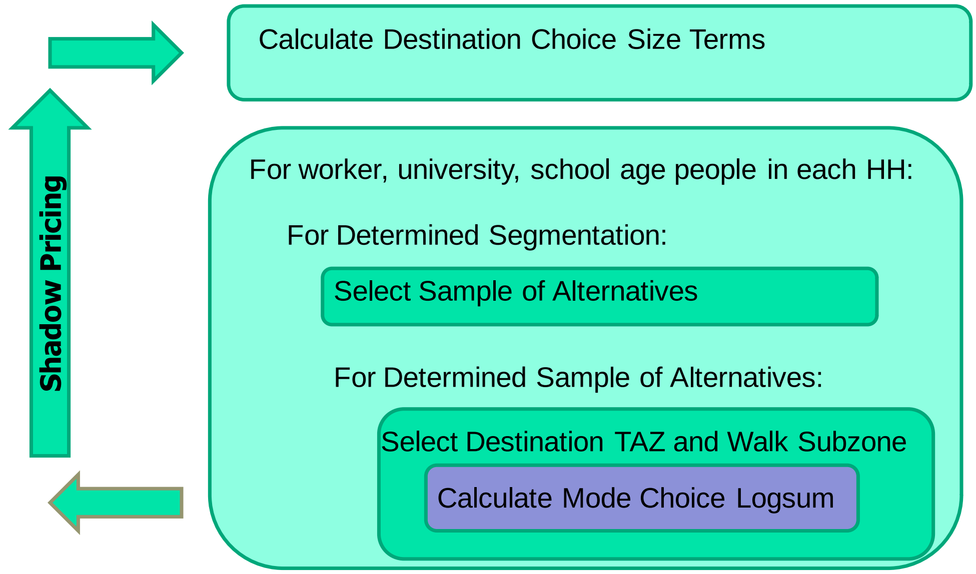


Figure 29: Location Choice and Shadow Pricing

Shadow price adjustments are needed to ensure the following features:

- Constraining mechanism to get total tour origins and destinations to match for long term segments,
- Size variables adjusted to reflect more/less attractiveness by segment to influence destination choice,
- Iterate procedure in flowchart until sufficiently constrained,

- Calibration determines required number of iterations; normally 3-4 shadow pricing iteration per each global iteration of the entire model system is enough to reach a very good level of convergence (match to the size variables).

The only way to implement this procedure for millions of workers / students and thousands of location alternatives in a reasonable amount of time is to sample alternatives effectively by importance. Sampling by importance allows us to reduce the sample size significantly (to 30-40 locations instead of thousands) while preserving statistical features of the choice model. This, in general, cannot be achieved by a simple random sampling.

Sampling represents a simplified location choice model that is kept as close as possible to the final model but can be run fast (since it has to be applied for each zone!). In a typical location model, zonal size variables (like number of jobs for workplace location choice or school enrollment for school location choice) can be pre-calculated in advance and do not add much computational burden. Thus, the zonal size variables in the sampling model are equal to the size variables applied in the actual choice model. However, travel impedance measures applied in the actual choice model are quite time-taking and involve logsum measures calculated over multiple modes and time-of-day periods where each mode & time-of-day utility is a function of many LOS skims and person/household characteristics. This part in the sampling procedure has to be simplified. In most ABs, distance is used as a proxy for travel impedance for sampling. Another approach (applied for example in the New York Model to avoid sampling at all as discussed in the next section) is to pre-calculate and store a limited number of logsum-type matrices for the most important segments.

Packeting and Smart Pre-calculations

An example of packeting applied in the New York ABM for several choice dimensions is shown in **Figure 30**. With this strategy, the most time-taking operations of utility and probability calculations are implemented for a packet of households rather than for each particular household.

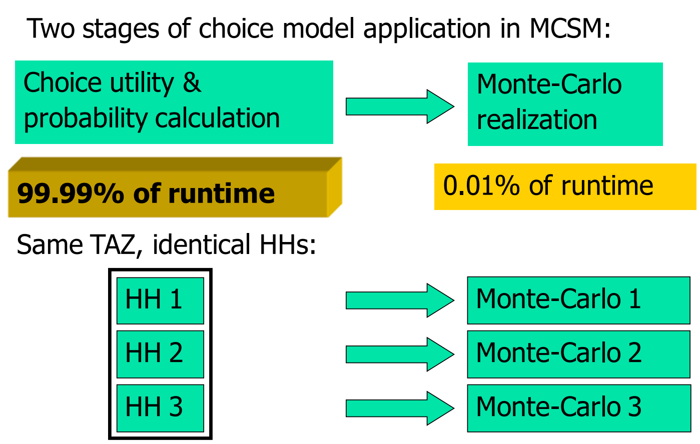


Figure 30: Packeting Method for Microsimulation

Packeting might be an effective method if the groups of households or persons with identical attributes are sufficiently large, i.e. number of segments is relatively small. For example, for a usual workplace location choice this method can be effectively applied if workers are segmented only by 3-4 income groups. In this case, for example, all low-income workers residing in the same zone would have the same workplace location choice probabilities. These probabilities can be pre-calculated once and then to each worker a unique zone will be assigned by Monte-Carlo. However, this method quickly reaches its limits if workers are also segmented by occupation, full-time/part-time status, age, etc.

An example of smart pre-calculations applied for destination choice in the New York AB is illustrated in **Figure 31**. Smart pre-calculation is a general programming principle that requires that calculations should be organized in such a way where the intermediate results could be reused multiple times. While this principle is of great help to make the program more efficient it makes it more difficult to comprehend the flow of calculations since it becomes less similar to the conceptual model structure.

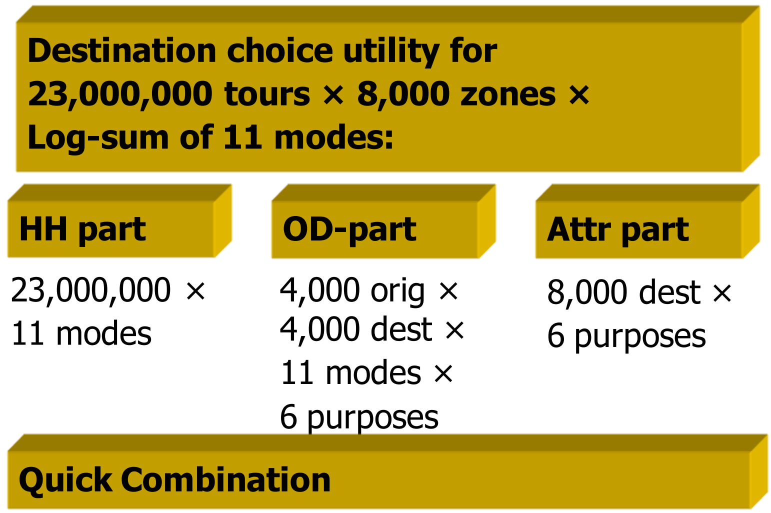


Figure 31: Example of Smart Pre-calculations

In this example, the calculation of a large number of destination choice utilities is broken into three parts:

- The household part encapsulates all household and person variables and takes a form of segmented mode-specific constants. These components are not dependent on the destination. Thus they can be pre-calculated for each tour once and applied to all destinations.
- The origin-destination part encapsulates LOS skims. This part, in order to be pre-calculated, has to have a limited number of segments. For this reason, in the New York AB model, time and cost coefficients in the mode utility functions are only segmented by travel purpose (and not segmented by any household or person attribute).
- The attraction part represents zonal size variables. These variables can be pre-calculated for each zone and sub-zone and travel segment.

After all components have been pre-calculated a quick combination is implemented on the fly for each tour. This way, sampling was completely avoided in the New York AB. However, the required simplifications, and primarily a limited segmentation for LOS variables are not behaviorally appealing. Thus, this example is more useful for improved sampling models than for the final destination choice.

Distributing and Threading

In general, parallel processing by means of multi-threading of calculations across CPUs and distributing across multiple computers (“boxes”) is a universal solution that can reduce run time almost inversely proportional to the number of CPUs in the cluster. Until recently, however, inter-linkage between model components and associated I/O time was coming into play strongly with an efficiency limit for a transportation model beyond which multi-processing did not bring an additional fruit. Today, with the progress achieved on the hardware side, powerful computer clusters can be built with a very small I/O interactional overhead. Below are several examples of distributing and threading in applied models.

Distributing and Threading Trip-based model components

The starting point for the ARC AB was the ARC trip-based model. The trip-based model's internal-internal (II) demand model was replaced by CT-RAMP, while the other components were used as is or updated as needed. The other components include network processing, a commercial vehicle model, an airport model, an external model, a time-of-day model, transit network building, and highway and transit assignment and skimming. Many of these components were already threaded in the trip-based model to take advantage of ARC's dual core machine. However, additional optimization was possible now that ARC had more computing power.

The first type of threading/distribution threaded all calculations in an origin zone loop using Cube Cluster's DistributeINTRAStep command, which distributes the calculations in blocks of origin zones to waiting Cube Cluster processes. When Cube Cluster completes, it writes an end text file and the main Cube Script reads this file and continues to the next step. This type of threading is independent of the number of processes available and is flexible for adding/removing processes. Distribution by origin zone loop allowed for threading two types of calculations: highway assignments and matrix processing.

There are currently seven highway assignments; four A.M., Midday, P.M., and Night. Threading the assignments therefore had substantial impacts on run time. As shown in **Figure 34** below (and discussed in more detail in the subsequent sections), the distributed highway run times are approximately 4 times faster than without distribution. The second primary process to be threaded with DistributeINTRAStep was matrix processing. This meant that much of the matrix processing, such as creating time-of-day matrices, was substantially sped up as well.

A second type of threading/distribution was done using Cube Cluster's DistributeMULTIStep command, which essentially distributes code blocks across multiple processes and then waits for all of them to complete before continuing. Unlike the first type of threading/distribution, this type requires explicitly assigning the tasks to specific processes, thereby being less flexible to adding/removing processes. This was implemented for highway assignments by time-of-day, conversion of trip lists to time-of-day matrices, and transit assignments by access mode, local/premium, and time-of-day. This type of threading improved model run times as well. For example, the amount of time to create demand matrices from the trip lists with distribution was 28 minutes compared to 112 minutes without distribution, or an improvement of almost 3 times.

Distributing and Threading the Activity Based Components

CT-RAMP, the microsimulation component of the model, is also distributed. The approach allows the utilization of one or more computers, each with multiple computing cores, and efficiently balances the computation among the computers. In addition, the configuration of the cluster of computers is relatively flexible, allowing computers to be easily added or removed. The model results, whether solved entirely on one computer or cooperatively with many computers, are identical.

Two essential design decisions were made at the start: 1) how to decompose the AB into independent sequences of computation that could be performed in parallel, and 2) how to efficiently compute thousands of tasks in parallel regardless of the number of computers available. In CT-RAMP, activity-based choice models are applied to households and to persons independent of those applied to other households. Since households are independent, groups of households are likewise independent. Thus, the application of activity-based choice models to groups of households was distributed. Given that CT-RAMP is implemented in Java, it was decided to use a robust open source library called the Java Parallel Processing Framework (JPPF) to manage the distribution of tasks. JPPF is an application written in Java that makes it easier to run applications in parallel.

As illustrated in **Figure 32** below, the JPPF framework consists of 3 main parts: a driver, a set of one or more nodes, and a client. The client is in this case CT-RAMP. The nodes are also additional separate processes, typically one per computer. The driver is a separate server process that is run on one of the cluster machines. The driver is a facilitator that receives tasks from the client application, sends them to node processes, receives results from nodes, and returns those back to the client.

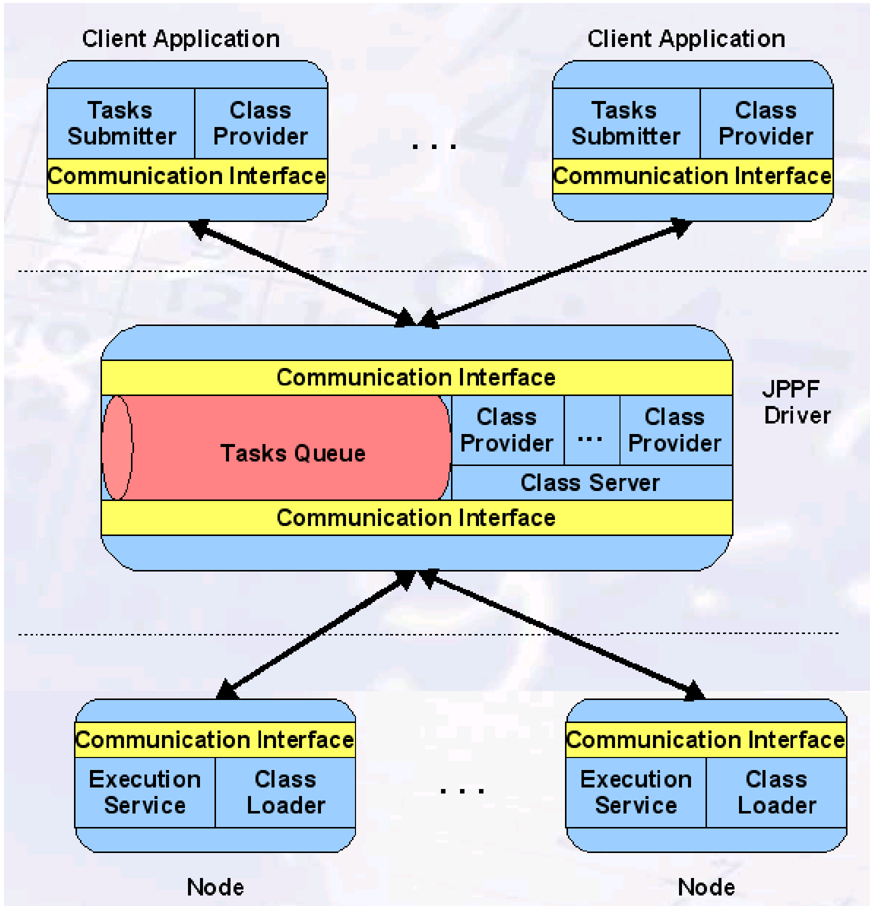


Figure 32: Java Parallel Processing Framework

Node processes receive tasks that perform a set of calculations, perform those calculations, and return results. Nodes are configured through a properties file to communicate with the driver process upon their start-up. A typical configuration might be to set memory equal to 32 GB and threads equal to 8 (for an 8 core machine). The majority of parallel computation in the CT-RAMP implementation occurs through tasks executed in parallel on nodes.

The driver process uses logic contained in the JPPF framework to balance computational loads across Java Virtual Machines running on the nodes in the cluster. The driver receives tasks from the client application and submits them in bundles to the nodes. The driver also retrieves class files from the client application and passes those to the nodes, as needed by the nodes. Additional nodes can therefore be added by simply editing two properties files and running a Java command.

The client application, which is called by the main Cube model script and configured through a properties file dynamically written by the model GUI, communicates with the driver as described above. The client application is responsible for creating task objects that can be run in parallel and submitting those to the driver. In the CT-RAMP implementation for ARC, 1.76 million households are split into 880 tasks of 2000 households each. The tasks are submitted to the driver and the driver assembles the 880 tasks into bundles and submits them to nodes that have notified the driver that they are part of the cluster. As the nodes complete the tasks, the driver receives their results and submits new bundles, while balancing the submission of bundles to keep the nodes uniformly busy.

In addition to the JPPF components, CT-RAMP has a Household Manager and a Matrix Manager process. The Household Manager's purpose is to manage the household and person synthesized populations in memory and to provide the JPPF nodes with all household and person related data. The Matrix Manager's purpose is to read the skims into memory and to provide the JPPF nodes with all requested skim values. These Java processes run on the main computer and have substantial memory footprints. To help reduce run time, the synthetic population is only created once in the Household Manager and then left in memory between feedback loops.

A sample of households is modeled during each feedback loop in order to conserve run time. For example, for the three feedback loop runs, the sampling shares are 33%, 50%, and 100%. The demand matrices created from the trip lists are scaled by the sampling percents before being loaded onto the network. In order to ensure that results are identical regardless of the order of processing, random number seeds are stored with each household and person so that random number cycles never go out of sequence.

The computational gains achieved by JPPF based applications depend very much on how the tasks are programmed and to what extent they can be solved in parallel. The speedup observed with CT-RAMP was a factor of approximately 9, or 587 minutes versus 7700 minutes using the hardware system described below.

It is relatively easy to divide the DaySim workload among multiple parallel processes, because the computations for one household do not interact with those of other households (this is a common feature of all ABs that is also discussed above for CT-RAMP examples). However, because the model reads a lot of data, parallel processes that separately read data from the same external source spend a lot of I/O time and experience I/O conflicts that can severely degrade performance. In order to overcome this, DaySim is being converted to a multi-threaded application that allows the parallel threads to access data from a shared memory location. This feature is especially important for implementations that specify a large number of zones, which is common in large regions, because the amount of I/O increases with the square of the number of zones. It is also important in regions with many households, because the amount of basic computation increases in proportion to the size of the population.

Hardware Configurations

In general, there is no need for CMAP to make an early decision on this since these are some of the most dynamic industries and new / better options become available every year. However, it is useful to identify hardware and configuration options as applied today for advanced ABMs, assess strengths and weakness of each option as well as costs and benefits in terms of runtime improvement. Below is a series of relevant examples.

DaySim Version 1 — Sacramento, 2006

The first version of DaySim was implemented in 2005 and 2006 for the Sacramento Area Council of Governments (SACOG). This version is single-threaded, without the ability to split the computational workload among multiple processors. This version of DaySim performs as follows in Sacramento:

- Hardware: Quad-Core AMD Opteron Processor, 2.49 GHz, 12 GB of RAM.
- Software: Microsoft Windows Server 2003 Standard, 64bit; Citilabs CUBE Voyager software.
- Scenario: 3.3 million people, 1.25 million households, 1533 zones, 665,000 parcels.
- Run Setup: Single process stream; no multithreading or parallel computing; 10 global iterations starting from a 1 in 256 household sample, doubling sample each iteration to a 100% sample on 10^{th} iteration.
- Run time: 32.5 hours.

DaySim Version 2 — Sacramento, 2008

A second version of DaySim was implemented in 2008 for SACOG. This version offered a simple way of distributing the computational load among multiple parallel processors in order to decrease total run time. In order to do this, the method of doubly constraining work location choice was changed to an iterative shadow-pricing method. Version 2 of DaySim performs as follows in Sacramento:

- Hardware: Quad-Core computer, 3.2 GB RAM.
- Software: Microsoft Windows XP 32 bit; Citilabs Cube Voyager software, with Cube Cluster.
- Scenario: 2.06 million people, 0.7 million households, 1533 zones, 612,000 parcels.
- Run Setup: 2 parallel processes for DaySim; 4 parallel processes for network models; 10 global iterations starting from a 1 in 256 household sample, doubling sample with each iteration to a 100% sample on 10^{th} iteration.
- Run time: 16 hours.

DaySim version 2 is also the basis for two other DaySim implementations. The first, mentioned earlier, is an integration of DaySim with the Transims Router for the Sacramento region, with Transims used instead of CUBE Voyager for highway assignment [*Resource Systems Group, et al, 2010a*]. The DaySim code was very easily adaptable to run with Transims, requiring only some changes in input and output file formats to meet the specifications of Transims. A minor adaptation was made to DaySim to use highway travel time matrices for 22 different periods of the day, taking more advantage of the dynamic, minute-by-minute nature of Transims network output.

The second adaptation of DaySim version 2 is for the PSRC Activity Generator, implemented as Phase 1 of AB model development for the Seattle Region [*Bradley, et al, 2008*]. This implementation includes a few important variations on the SACOG version of DaySim. The code was adapted to read and write to EMME binary databanks instead of CUBE Voyager binary formats. Specific models were re-estimated based on Seattle region survey data.

CT-RAMP Atlanta (ARC) Example

The first requirement for a distributed/threaded modeling system is a cluster of computers. ARC purchased three computers to run the ABM. The specifications of these computers are:

- Windows Server 2003 64bit,
- Dual Quad Core Intel Xeon X570 2.93 GHz Processors (8 total),
- 32 GB of RAM,
- Cube Voyager + 8 seat Cube Cluster license (16 total seats with hyper-threading).

While the hardware and general IT challenges were not insurmountable, it required some trade-offs in ARC's IT strategic plan for supporting its modeling department, with the key challenge consisting of "right-sizing" its hardware and software balance for the AB. The total cost of the new hardware and software licenses was around $30,000.

In addition, the modeling system requires Java (32bit and 64bit versions), Cube Base for the GUI, and the CT-RAMP Java package. The model requires a 64 bit OS in order to take advantage of larger (64-bit) memory addresses. Since Cube is a 32bit application, a 32bit version of Java is required in order to natively read binary Cube matrices.

CT-RAMP Bay Area (MTC) Example

The MTC cluster consists of four machines: a master and three satellite machines. Each machine has an identical configuration, except that the master has a larger hard drive. The basic specifications for each are:

- 16 64-bit processors (2 hyper-threaded quad-core chips);
- 48 GB RAM;
- 2 TB hard drive on master; 1 TB hard drive on slaves;
- Microsoft Windows 2008 Server (64-bit) operating system.

Vendor price quotes were approximately $35,000.

The machines are on a local network. They are accessible for use by MTC staff via remote desktop, and accessible to PB via a GotoMyPC account.

The MTC model utilizes software that must be installed on each machine prior to a model run. The software needed is listed below.

- 64-bit version of the Java Development Kit (JDK) installed on each machine. This will allow the java software, used for implementation of the model, to execute.
- 32-bit version of the Java Runtime Environment (JRE) installed on each machine. This is required in order to read and write matrices in TP+ format.
- Cube Base, Cube Voyager, and Cube Cluster installed on the main machine. This provides the full functionality of Cube on one machine, for use in the skimming and assignment steps.
- Cube Voyager Node installed on the satellite machines. Voyager Node allows instructions to be processed as slave nodes, but not initiated. Therefore all Cube processes must be initiated from the main machine, but can be executed on the satellite machines to save runtime.

System Setup and Design

ARC's AB system is implemented with Java and Cube, with JPPF and Cube Cluster for distributed/threaded computing respectively. MTC's setup is slightly different, with DOS BAT files running the overall model, but the distribution/threading approach is the same. The ARC system design consists of:

1) A main computer which:
    a. runs the main Cube work, including some of the Cube Cluster processes,
    b. runs the main CT-RAMP (ArcTourBasedModel) Java process,
    c. runs the Household Manager Java process,
    d. runs the Matrix Manager Java process (including the slave 32bit process to natively read, Cube binary matrices in Java),
    e. runs the JPPF Driver Launcher processes which is called by ArcTourBasedModel, and which manages communication with the nodes,
    f. optionally runs a JPPF node process which listens for tasks from the JPPF driver:
2) Two additional node computers which:
    a. Listen for tasks from the JPPF driver,
    b. Listen for Cube Cluster tasks.

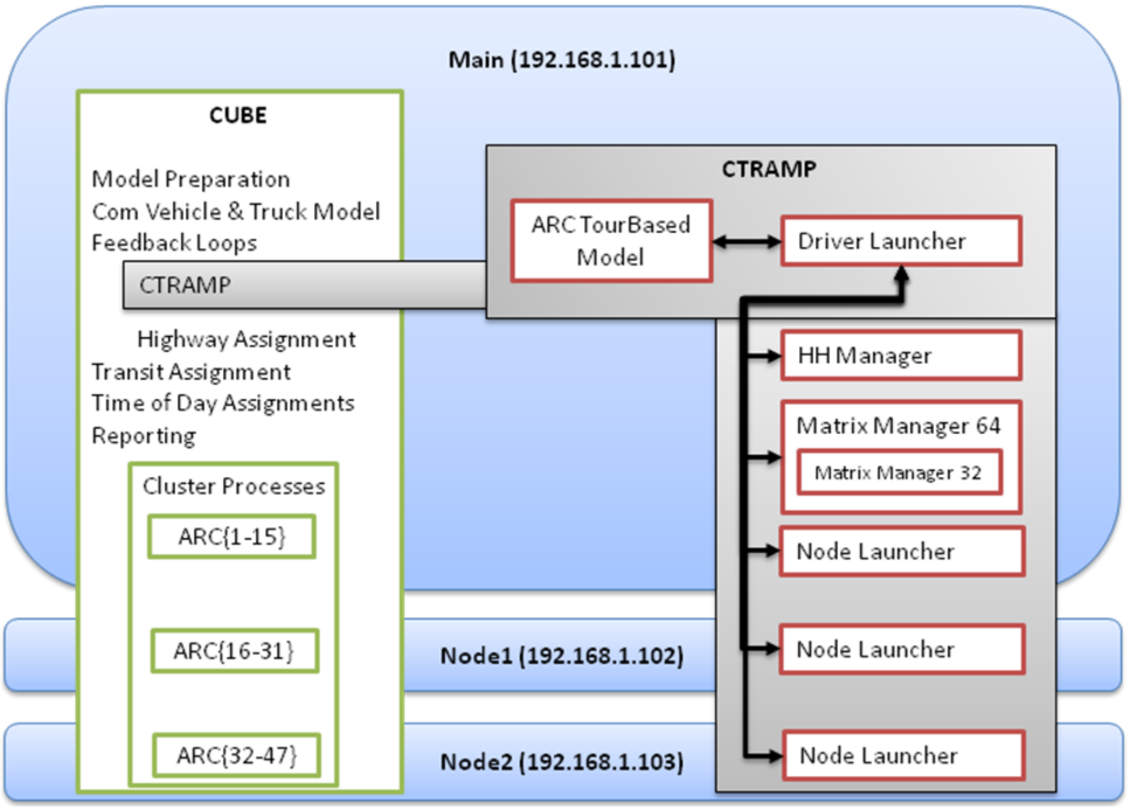


Figure 33: ARC ABM System Design

Configuring and running the model is straightforward. The Java and Cube Cluster slave processes must be started on the two additional computers. There is a startup script for each computer that starts all required Java processes on the machine. Next the user starts the Cube Cluster processes in wait mode on each machine. The final step is to open and execute the main run script in Cube. All the model components on each machine talk to one mapped network directory. All inputs are read, and all outputs written to this folder, thereby simplifying model setup, inspection, and error detection as if the model was run on a single computer.

Runtime Statistics

Atlanta (ARC) AB Runtime Statistics

ARC decided a few years ago to invest in additional “fire-power” for its trip-based travel demand model, especially to properly address tolling and managed lanes alternatives. This led to the purchase of a dedicated computer with a dual core processor, 8 GB of RAM, and a two seat Cube Cluster license for distributed processing of Cube Scripts. This machine was used to run the trip-based model in approximately 12 hours.

Over the past several years, ARC, along with its consulting team, was developing a new AB based on the CT-RAMP (Coordinated Travel – Regional Activity Based Modeling Platform) family of models, which is based on models implemented in New York, Columbus, OH (MORPC), and Lake Tahoe. The plan for the ARC AB was to develop a modeling system that was distributed across multiple computers/threads and had an overnight run time.

As currently implemented, the ARC AB runs a complete base year run in approximately 16 hours and 10 minutes. As shown in **Figure 34**, this includes microsimulating travel patterns for 1.76 million households, running multiple highway and transit assignments, running the entire modeling system (including network building, external demand, etc), and running three feedback loops. Without any distribution/threading, the model runs in approximately 146 hours (or approximately 6 days).

Two other features of the ABM implementation are important. First, highway assignment procedures that are implemented for four periods of a day take about 50% of the entire model system run. Thus, it is imperative to take a full advantage of multi-processing with respect to these procedures. The core demand ABM is the second time-taking component that has to be parallelized with great potential savings. It also can be seen that the runtime of the microsimulation procedure is almost exactly proportional to the number of modeled agents (households). Thus, a substantial economy can be achieved if only a subset of agents is modeled instead of the entire population. In this case, each agent obtains a weight (in other words represents 2, 3 or more households that are assumed identical with respect to their structure, place of residence, and all travel choices). This of course, results in a certain “lumpiness” of the results and small aggregation biases but this is a reasonable compromise at first global iterations when high precision is not essential (and LOS skims are not final yet).

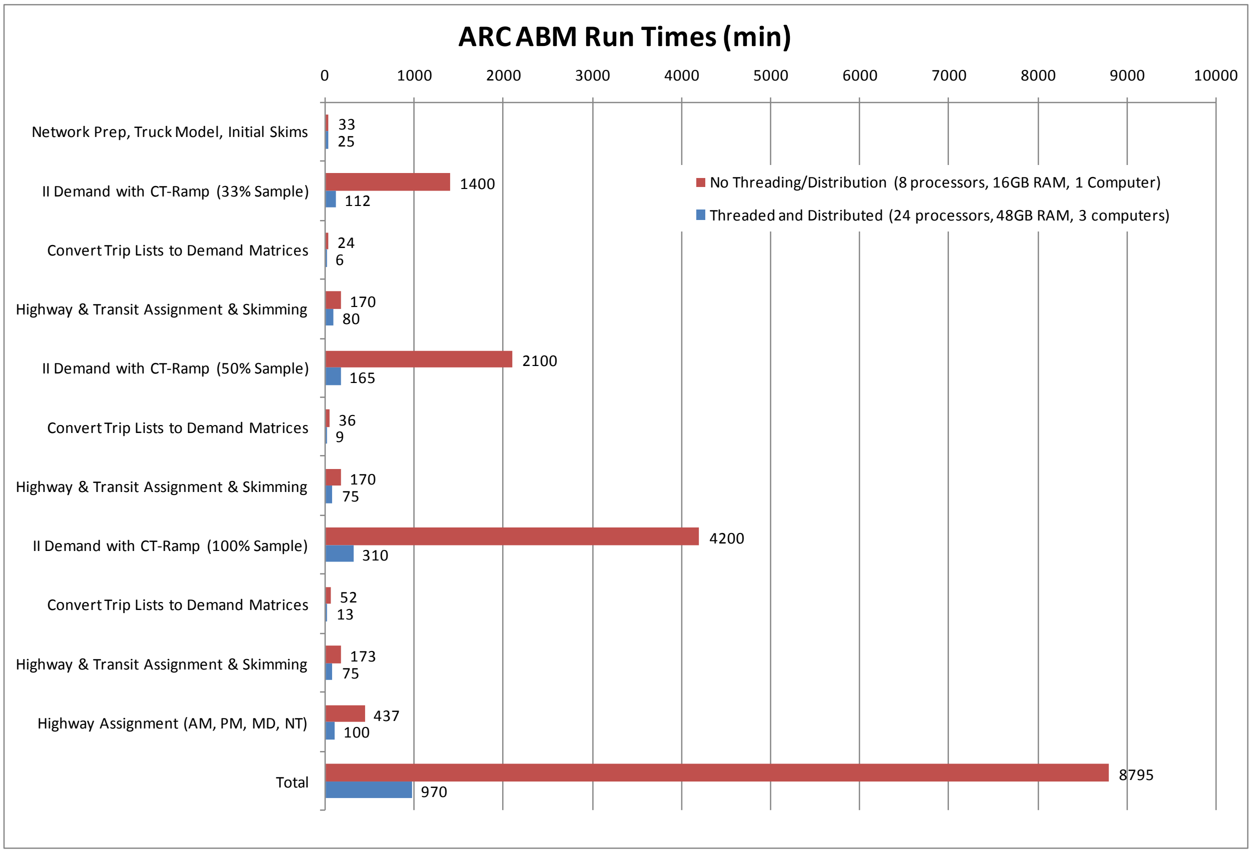


Figure 34: ARC ABM Run Times

Columbus (MORPC) ABM Runtime Statistics

Runtime statistics for the Columbus, Ohio, ABM developed for the Mid-Ohio Regional Planning Commission (MORPC) are summarized in **Table 16**. The MORPC model is another representative of the CT-RAMP family of ABMs although the region size is much smaller than the CMAP region.

Table 16: Runtime Statistics for MORPC model

Model components	2000			2000	2030		
	MORPC 3	MORPC 4	COTA	ODOT	MORPC 3	MORPC 4	COTA
Households	610,774	610,774	610,774		872,919	872,919	872,919
Population	1,435,389	1,435,389	1,435,389		1,956,660	1,956,660	1,956,660
Tours	2,074,618	2,073,659	2,075,797		2,997,507	2,997,214	2,996,117
Core Model Total (3 iterations)	35:43	31:20	20:55	10:25	48:35	41:23	26:43
Iteration 1	11:27	10:08	6:51	3:29	16:18	13:28	8:30
Iteration 2	11:26	9:55	6:28	3:28	14:59	12:48	8:06
Iteration 3	12:49	11:16	7:36	3:27	17:17	15:06	10:06
Iter 1 - Population Synthesis	0:02	0:02	0:01	0:01	0:02	0:02	0:01
Iter 1 - Sending Files to Workers	0:20	0:20	0:12	0:39	0:19	0:20	0:14
Iter 1 - Auto Ownership	0:01	0:01	0:00	0:00	0:02	0:02	0:01
Iter 1 - Mandatory Tour Generation	0:53	0:53	0:39	0:29	1:15	1:15	0:39
Iter 1 - Mandatory DTM	4:01	3:14	1:59	0:55	6:07	4:48	2:50
Iter 1 - Joint Tour Generation	0:12	0:12	0:08	0:07	0:14	0:14	0:08
Iter 1 - Joint Tour DTM	0:08	0:06	0:04	0:05	0:08	0:07	0:05
Iter 1 - Individual Tour Generation	0:05	0:05	0:05	0:03	0:07	0:07	0:05
Iter 1 - Individual Tour DTM	0:54	0:41	0:23	0:11	1:15	0:56	0:30
Iter 1 - At-Work Sub-Tour DTM	0:08	0:07	0:06	0:03	0:12	0:10	0:07
Iter 1 - Mandatory Stops Model	0:49	0:38	0:21	0:11	1:14	0:59	0:32
Iter 1 - Joint Stops Model	0:07	0:06	0:04	0:07	0:08	0:07	0:05
Iter 1 - Individual Stops Model	0:54	0:43	0:24	0:14	1:11	0:54	0:31
Iter 1 - At-Work Stops Model	0:06	0:05	0:04	0:05	0:09	0:08	0:05
Iter 1 - Writing Files and Trip Tables	0:13	0:13	0:10	0:12	0:35	0:34	0:26
Iter 1 - External Model +	0:00	0:00	0:00		0:01	0:01	0:01
Iter 1 - Commercial Vehicle +	0:02	0:02	0:01		0:02	0:02	0:01
Iter 1 - IE Trips +	0:00	0:00	0:00		0:00	0:00	0:00
Iter 1 - Highway Assignment - 2 period +	1:08	1:14	1:07		2:03	1:31	1:16
Iter 1 - Highway and Transit Network Skimming +	1:17	1:17	0:53		1:04	1:03	0:44

Iter 3 - Highway Assignment - 4 period +	2:14	2:18	1:51		3:11	3:07	2:19
Iter 3 - Transit Assignment - 2 period +	0:16	0:16	0:10		0:12	0:12	0:07
Core Model	2:59	2:30	1:36	1:09			
		16%	36%	27%			

What makes these statistics unique is that multiple trials were implemented with computer clusters of different power and configurations including 3 workers, 4 workers, 6 workers (for COTA – Columbus Transit Authority), and 8 workers (for ODOT – Ohio DOT). It can be seen that across the range of computer powers there is a clear tendency of runtime reduction inversely proportional to the computer power (number of CPUs).

Sacramento (SACOG) AB Runtime Statistics

A summary of hardware features and corresponding runtimes for two generations of the Sacramento AB is presented in **Table 17**. It can be seen again that parallel processing allowed for cutting the runtime to a half.

Table 17: Runtime Statistics for SACOG model

Feature / characteristic	**Version 1 (2006)**	**Version 2 (2008)**
Hardware	Quad-Core AMD Opteron Processor, 2.49 GHz, 12 GB of RAM	Quad-Core computer, 3.2 GB RAM
Software	Microsoft Windows Server 2003 Standard, 64bit; Citilabs CUBE Voyager software	Microsoft Windows XP 32 bit; Citilabs Cube Voyager software, with Cube Cluster
Population size	3.3 million people, 1.25 million households	2.06 million people, 0.7 million households
Network size and spatial units	1533 zones, 665,000 parcels.	1533 zones, 612,000 parcels.

Feature / characteristic	Version 1 (2006)	Version 2 (2008)
Run Setup	Single process stream; no multithreading or parallel computing; 10 global iterations starting from a 1 in 256 household sample, doubling sample each iteration to a 100% sample on 10th iteration.	2 parallel processes for DaySim; 4 parallel processes for network models; 10 global iterations starting from a 1 in 256 household sample, doubling sample each iteration to a 100% sample on 10th iteration.
Runtime	32.5 hours	16 hours

Open-Source Software vs. Proprietary Applications

The AB modeling approach includes many software components developed by different parties with a differential degree of "openness". From the CMAP perspective it is important to ensure continuity in terms of the software use, maintenance, as well as possible modifications for the entire lifecycle of the model (normally 10-15 years). Ideally, for a public agency having the entire model system as an open (and readily modifiable) source would be desirable. However, certain components of the model will always be proprietary pieces of software where the agency will have to rely on the software vendor and will be subject to the maintaining/upgrading policies of this vendor that may not necessarily coincide with the agency plans for model upgrade and/or modification.

In general, the following strategy can be outlined with respect to the main components of the ABM system:

- *Core AB demand model*. This software is normally developed by transportation consultants (that was the case with the absolute majority of ABs in practice). It is reasonable for CMAP to request a well-documented source code for the entire core AB (including population synthesizer and all travel sub-models) and request ownership for the final product contractually. This way, CMAP will be able to modify the code if necessary and share it with the other public agencies or contractors working for CMAP.

- *Network simulation software*. CMAP will have to rely on one of the software vendors with respect to highway and transit assignments, LOS skimming and other network procedures. CMAP is currently using EMME as part of the existing 4-step model. In this part, requesting an open source will be unrealistic in most cases although certain types of software, for example DTA models might be available through FHWA programs (University of Arizona made Dynus-T available as an open source for the SHRP 2 C10 project that is going to be disseminated after completion of the project). It also has to be mentioned that comparing to the core AB model it is less likely that CMAP would be able to take advantage of open source for network simulation procedures. Highway and transit assignment algorithms are quite complicated and even if some bugs are revealed or modifications are necessary the most realistic approach would be to report the problem to the vendor and request a new version of the software rather than try to fix the problem in-house.

- *Multi-threading and distribution software*. Taking into account the size and complexity of the Chicago Metropolitan Region, it is quite probable that the CMAP AB will require a modern powerful computer cluster to ensure realistic (overnight) runtimes. In order to take advantage of this hardware set-up, certain software components (multi-threading of the flow of calculations and distribution of the threads across multiple computers and CPUs) must be developed and integrated with the core AB and network simulation software. In general, a distributing of the core model has to be programmed in accordance with the core AB software algorithm (processing packets of households) and implemented by the same developer. It is reasonable to request this piece of software as an open source in the same way as the core AB itself. With respect to network procedures, there are two possible options to consider. One option is to distribute network procedures by time-of-day where each node (worker) in the cluster would process an entire set of assignment/skimming operations for a particular time of day. In this case, the distributing would be programmed as part of the open-source custom software developed by the consultant. Another option that has recently become available with majority of the transportation packages (EMME, Cube, TransCAD) is to effectively distribute each assignment. In this case, the distributing software piece would be a part of the proprietary software. In any case, it is essential to ensure a modifiable distribution system since adding computers to the existing cluster (or moving the ABM system to a large cluster) is a frequent case.

Staffing and Qualification Requirements

This section identifies the basic knowledge and skill level needed to develop, maintain, and apply an advanced AB (or its particular components) at CMAP. These requirements are formulated for both the consultant staff and CMAP staff with some recommendations regarding the split of responsibilities between the consultant and CMAP staff. Development of an advanced ABM and its successful "absorption" in the agency requires a significant and consolidated effort from both sides. The key factor for success is to work as a single project team where different participants (model developers, model users, software engineers, IT/hardware experts) complement each other and yet are able to communicate properly and understand each other's needs and constraints.

It should be mentioned that while the required qualifications for the combined team can be well outlined, there is a great deal of variation with respect to the split of responsibilities between the consultant and agency staff. All existing ABs in practice have somewhat a different story with successful examples of a significant involvement of the agency staff in some stages of model development. For example, MTC and SANDAG staff was significantly involved in the model estimation effort and the consultant delivered special hands-on training courses for logit model estimation for the agency as part of the model development contract. DRCOG has largely coordinated and implemented the entire software development cycle in-house. MORPC staff was responsible for the entire set of network procedures. Experience has shown that this is largely a function of the staff qualification. In general, it is beneficial for the CMAP staff to be involved in the ABM development and application process as far as possible given the staff qualifications and availability.

In the entire AB development, application, and maintaining process, six different levels of proficiency required from different groups of developers and users:

- *CMAP Project Manager / coordinator*. Development of an advanced AB is a multi-stage inter-disciplinary project of a high level of technical and management complexity. Despite the fact that the technology and the most complicated pieces of software are provided by consultants and software vendors, the ultimate "assembly" of the system and various inputs (socio-economic and land-use data, highway and transit networks, specifications for projects, policies, and scenarios) are in the CMAP domain. The proper organization of the entire project requires an experienced PM from the agency with a good technical knowledge of models, substantial local experience, and administrative power to ensure all necessary inputs from the CMAP side. In some agencies like SANDAG, a decision was made to create a new full time position of PM for AB development.

- *Model system designer / architect (Principal Investigator and/or Project Manager from the consultant side)*. It is normally a single individual with substantial credentials as a researcher and proven experience in delivery of ABs in practice. This person manages the consultant side and is responsible for the model design, structure, and specifications of all sub-models. This person would normally supervise the model estimation and calibration effort as well. This person has also to possess very good communication skills and, in particular, ability to communicate effectively with a wide audience of transportation planners. The CMAP PM and consultant PM form a project management unit that have to be effective at all stages of the project. The consultant PM will play a role of an interface between CMAP (and wider audience of potential model users) and the technical group of model developers and programmers.

- *System analyst / architect*. An AB model represents a complicated piece of software that includes several different components and involves multi-threading and distributing. The software design has to be implemented by a professional system analyst who is responsible for the entire-system architecture and fitting different software pieces together also taking into account the hardware configuration. This person should be involved each time if a significant modification of the model system structure, for example adding an interface between AB and DTA is planned (if DTA is applied instead of conventional static highway assignment). A system analyst of the required level should be a part of the model development team (normally, on the Consultant side).

- *Programmer*. A person on this position is responsible for actual code for custom pieces of software. This person is involved each time if any modification of the code is required, for example, adding new transit mode in the mode choice model. This position can be split between several persons with responsibilities for different pieces of software that are relatively independent from each other and have only I-O type of relation. Examples of such pieces that can be implemented by different programmers are population synthesizer, long-term location models, mid-term models for mobility attributes, freight model, and core day/tour/trip-level model. However, splitting the programming task further (for example, between day level, tour-level, and trip level sub-models given to different programmers) is not recommended because these sub-models are very closely intertwined. This position is normally filled by 1-2 members of the Consultant team but there are examples of MPOs like DRCOG and SANDAG that employed their own programmers for some tasks.

- *Modeler*. Throughout the entire cycle of model estimation, development, calibration, application, and long-term support there is a need in major modeling skills. In the modular system design presented in the report and strongly advocated by the authors there is a clear distinction between software architecture (main code) and sub-model specifications fully encapsulated in the UEC. Each UEC corresponds to a particular choice sub-model and contain the entire specification including all variables and coefficients in the utility expressions. The modeler task is to estimate each model and code it in the UEC. Manipulating UECs is a routine part of model calibration as well as a frequent part of model application. This does not require any change in the code and consequently does not require programming skills but rather requires understanding of the model structure and parameters. It is essential to have at least one person from CMAP who qualifies for this position and would take over the model UEC setup at least at the model application stage. It is highly recommended also for the CMAP staff to be involved in model estimation and calibration stages as much as possible although in most ABM development projects these stages were staffed by modelers from the Consultant team.

- *End user / data processing person*. The ultimate users of the model will be applying it for evaluation of different projects and policies. This position does not require programming skills per se but some familiarity with the modeling side is important since in reality model application and calibration aspects are closely intertwined. Model users have to be able to manipulate input data (for example, population and employment by zone), networks (for example, code new transit lines or highway tolls), and outputs (for example, prepare input files for evaluation of User Benefits for a New Starts mass rapid transit project). This position should be staffed by CMAP (as well as some other regional agencies that will be using the model) and Consultant's responsibility is to train the CMAP staff with respect to model application. As was explained above in the sub-section on processing AB output, an end-user of AB is required to possess an additional set of skills compared to 4-step users. It includes proficiency with data-processing software that can handle large files with millions of records and hundreds of data fields (Access, SQL, FoxPro, SAS, or SPSS) in combination with one of GIS packages.

There can be different configurations of the project team as well as different distribution of responsibilities between the agency staff and consultant team, especially with respect to the modeler and programmer positions. It is largely a function of available personal qualifications. It also should be mentioned that the required skills for the modeler, programmer, and user positions are very much subject to the educational and training process organized in parallel with (or in advance of) the model development process . In general, our general recommendation to CMAP (and any other agency that might envision development of an advanced transportation model) is to consider a maximum realistic in-house commitment with respect to the modeling and programming tasks and positions. This approach ensures a reasonable level of independence in subsequent model maintenance and/or modifications compared to a total outsourcing of all tasks to a Consultant.

It should be stressed that the proposed software implementation paradigm clearly distinguishes between programmers (who see the main code but do not see model specification details and parameters) and modelers (who do not have access to the main code but can manipulate model specifications and parameters). This makes it realistic to organize a focused training / professional development process at CMAP with staff designated to one of these tasks.

As an example, in **Table 18**, is shown a list of possible general tasks where the CMAP staff support is welcome with an indication on the possible degree of involvement.

Table 18: Possible CMAP Staff Participation

Task	**CMAP staff involvement**	**Consultant role**
Preparation and processing of the input population and land-use data for the base year and future years	Provides the final data files and implements all data processing steps	Specification of the input data and help in data processing if needed
Development of highway and transit networks	Provides final coded highway and transit networks in the EMME format for all specified years, scenarios, and time-of-day periods	Help in network coding specifications if needed
Development of network procedures (assignment, skimming, and others) using EMME macro scripts	Participation in development depending on CMAP staff availability; full participation in testing and calibration	Specification of all procedures and responsibility for their development, testing and calibration; guidance for CMAP staff
Processing of the household survey in formats for the model estimation	Participation in data processing and statistical analysis depending on CMAP staff availability	Full specification and primary responsibility for data processing; guidance for CMAP staff

Task	CMAP staff involvement	Consultant role
Disaggregate estimation of behavioral models	Participation in estimation of certain models depending on CMAP staff availability and qualification (1-week hands-on training course might be needed); full participation in statistical analysis and discussions on model estimation results	Full specification of all models and primary responsibility for the statistical analysis and estimation; guidance for CMAP staff and delivery of the training course if needed
Development of the core demand model software	Participation in programming of certain models and procedures, primarily Utility Expression Calculators for choice models	Primary responsibility for the software architecture and programming; guidance for CMAP staff
Model validation, testing, and calibration	Full participation in model validation and testing including specification of the validation targets and scenarios for model testing	Guidance and training courses for CMAP staff regarding the model application; primary responsibility for model calibration and compliance with the adopted calibration targets

Conclusions and Next Steps

This section summarizes the main aspects of AB implementation and computing environment that are relevant to the CMAP plan to move towards an advanced travel model. The following general conclusions can be made at this stage:

- Activity-based models are characterized by three main features that make them very different from the 4-step models: 1) derivation of travel (tours and trips) from individual activity patterns; 2) using tour as the main unit for travel modeling instead of elemental trip, 3) individual microsimulation of households and persons instead of aggregate zone-to-zone trip flows. These features, especially individual microsimulation change the flow of calculations in a principal way compared to 4-step models. While in a 4-step model most of the calculation sets are implemented for each Origin-Destination pair (and the main computational loop is over Origin-Destination pairs), in a microsimulation AB, calculations are organized by individual records – households, persons, tours, and trips (and the main computational loop is over households).

- Microsimulation ABs of travel demand currently cannot be implemented using a script language of one of the transportation software packages like EMME, TransCAD, Cube, or Visum. The required functionality and flexibility in treatment of individual objects is not trivial and is far beyond simple matrix and vector manipulations embedded in these packages. So far, all ABMs in practice have been implemented as a custom piece of software developed by a Consultant and tailored to the particular region. There are first examples of software libraries developed by Consultants to construct ABMs like CMF developed by PB and used to construct 7 ABMs of the CT-RAMP family as well as several other models. The DaySim software is being used for several ABMs (including two SHRP 2 C10 projects).

- Contrary to demand ABs developed by Consultants, network assignment and skimming procedures continue to rely on commercial software packages. In the short term and taking into account the size and complexity of the CMAP network, an ABM will probably have to be combined with a conventional static assignment (like in the existing EMME-based 4-step model) that has to be implemented for multiple periods of a day. This has been the case so far with all ABMs in practice. The long-term tendency is to integrate ABM and DTA. For this reason, it is recommended to CMAP to start preparations for a regional DTA procedure in parallel with the ABM development. In particular, DTA setting needs certain additional network details that have to be coded and calibrated that is in itself a tme-taking effort.

- Microsimulation has certain specifics and major dimensions that affect ABM algorithmic structure and eventually, run time. In general, the core demand model runtime is roughly proportional to population size since each set of calculations has to be implemented for each household. In this regard, there are several ABMs developed for a region comparable to the CMAP in terms of population like Atlanta or Bay Area ABMs that provide good examples and indications on what CMAP might reasonably expect.

- Run times for network assignment and skimming procedures are roughly proportional to the squared number of TAZs and they normally take 50% or more of total model system runtime taking into account that multiple multiclass highway assignments and mode-specific transit assignments have to be implemented for different time-of-day periods. All major vendors of transportation planning packages (INRO, Caliper, Citilabs, PTV) provide today various multiprocessing options that can be effectively utilized with the corresponding hardware configurations.

- Overnight model runs for large regions comparable to CMAP are possible with threading and distribution. For example, the current ARC trip-based model runs in 12 hours. Consistent (16 hour) run times are achieved for the ABM, allowing staff to set up a scenario before leaving for the day, and results analyzed the next morning. The substantial improvements in run times compared to an AB without distribution/threading were made possible by a strong supply of computing power, common software, co-development with ARC / MTC, and a distributed/threaded implementation. The ability to turn around modeling results quickly is very important to the model remaining relevant to ARC's planning work.

- More hardware can reduce runtime. The modeling system is built to take advantage of adding additional computers/processes to reduce run times even more than described in the current report that describes experience with 24-32 CPUs configurations. However, the law of diminishing returns applies to adding additional threads, and improvements beyond a few additional hours are not expected although more practical experimentation with large clusters is needed to assess this.

- In longer term, some other computing technology solutions might prove effective, including possibly cloud computing. Also, with the current tendency of combining more and more CPUs on the same motherboard, multi-threading on a single super-computer may eventually replace distributing amongst multiple computers in a cluster. Indeed, already today, engagement of 32 CPUs can be implemented either through multi-threading on a single computer with 32 CPUs or on a cluster of 4 computers with dual quad cores at a comparable cost. Since multi-threading is essentially a simpler operation compared to distributing that requires moving gigabytes through the wire, the future probably belongs to supercomputers rather than large clusters. At this point it is premature to make an exact recommendation to CMAP regarding the hardware configuration. However, in terms of budget planning for the next 5 years, it is reasonable to expect that either a computer cluster or single super-computer with 32+ CPUs will be needed to implement an advanced microsimulation model system. In both cases, cost of $50,000 would be a reasonable estimate.

References

- Activity-Based Travel Model Specifications: Coordinated Travel – Regional Activity Based Modeling Platform (CT-RAMP) for the Atlanta Region. (2009). Parsons Brinckerhoff.

- Activity-Based Travel Model Specifications: Coordinated Travel – Regional Activity Based Modeling Platform (CT-RAMP) for the San Francisco Bay Area. (2009). Parsons Brinckerhoff.

- Bowman, J. (2009) Historical Development of Activity-Based Model Theory and Practice, Traffic Engineering and Control, Vol. 50, No.2, 59-62.

- Bowman, John L. and Mark A. Bradley (2006) Upward Integration of Hierarchical Activity-based Models, working paper. Available at http://www.JBowman.net.

- Bowman, John L. and Mark A. Bradley (2008) Activity-Based Models: Approaches used to achieve integration among trips and tours throughout the day, presented at the 2008 European Transport Conference, Leeuwenhorst, The Netherlands, October, 2008. Available at http://www.JBowman.net.

- Bowman, John L., Mark A. Bradley and John Gibb (2006) The Sacramento Activity-Based Travel Demand Model: Estimation And Validation Results, presented at the European Transport Conference, September 18-20, 2006, Strasbourg, France. Available at http://www.JBowman.net.

- Bradley, Mark A., John L. Bowman and Joe Castiglione (2008) PSRC Activity-Based Model Phase I Final Report: Activity Generator Model Estimation, Calibration & User Guide, November 30, 2008—Draft 1. Available at http://www.JBowman.net.

- Bradley, Mark, John L. Bowman and Bruce Griesenbeck (2010) SACSIM: An applied activity-based model system with fine-level spatial and temporal resolution, Journal of Choice Modeling Vol, 3(1) 2010. Available at http://www.JBowman.net or http://www.jocm.org.uk/

- Castiglione J, Freedman J, Bradley M. (2003) Systematic Investigation of Variability due to Random Simulation Error in an Activity-Based Microsimulation Forecasting Model. In Transportation Research Record 1831, Washington, D.C., 76-88.

- Cube Cluster, http://www.citilabs.com/cube_cluster.html

- Davidson, W, Donnelly R, Vovsha P, Freedman J, Ruegg S, Hicks J, Castiglione J, Picado R. (2007) Synthesis of First Practices and Operational Research Approaches in Activity-Based Travel Demand Modeling, Transportation Research A, Vol. 41A, Issue 5, 464-488.

- Freedman, J., P. Vovsha, W. Sun, S. Gupta, Z. Ouyang, R. Curry, C. Daniels (2010) Space and Accessibility in the SANDAG Activity-Based Model. Presented at the *Innovations in Travel Modeling* Conference, Phoenix, AZ.

- NYMTC Transportation Model and Data Initiative: New York Best Practice Model. General Final Report. (2004) Prepared by PB Consult, Inc for the New York Metropolitan Transportation Council.

- Resource Systems Group, Mark A. Bradley and John L. Bowman (2010a) TRANSIMS Deployment Case Study: SACOG DaySim-TRANSIMS Integrated Model Development, Final Report, March 2010.

- Resource Systems Group, AECOM, Mark A. Bradley, John L. Bowman, Mohammed Hadi, Ram Pendyala, Chandra Bhat and Travis Waller (2010b). SHRP 2 C10A: Partnership to Develop an Integrated, Advanced Travel Demand Model and a Fine-Grained Time-Sensitive Network, Task 2 Report: System Architecture (DRAFT), April 2010.

- SACOG (2008) Sacramento Activity-Based Travel Simulation Model (SACSIM07): Model Reference Report. Available at http://www.JBowman.net.Vovsha, P. (2005). Activity-Based Models in Practice: The Mid-Ohio Regional Planning Commission (MORPC) Model. 84th TRB Meeting, http://www.trb-forecasting.org/TRBWorkshop-158_Vovsha.pdf.

- The MORPC Travel Demand Model. Validation and Final Report. (2005) Prepared by PB Consult, Inc for the Mid-Ohio Regional Planning Commission.Vovsha, P. Peterson, E., and Donnelly, R. (2002), Microsimulation in Travel Demand Modeling: Lessons Learned from the New York Best Practice Model. Transportation Research Record 1805. Pages 68-77.

- Vovsha, P., Bradley M, Bowman J. (2005) Activity-Based Travel Forecasting Models in the United States: Progress since 1995 and Prospects for the Future. In Progress in Activity-Based Analysis, H. Timmermans (ed.), Elsevier Science Ltd, Oxford, UK, 389-414.

- Vovsha, P., R. Donnelly, and S. Gupta (2008) Network Equilibrium with Activity-Based Microsimulation Models: The New York Experience. In Transportation Research Record 2054, Washington, D.C., 102-109.

- Willison, C. (2007). Development of an Activity-Based Model in the Lake Tahoe Region. 11th TRB Planning Applications Conference. http://www.trb-appcon.org/TRB-Presentations/Session%208/1_2007TRBPresentation_ChristiWillisonNew.ppt

Appendix:
Integration of Model Components in Activity-Based Models

Prepared by

John L. Bowman, Ph. D.
Bowman Research and Consulting
28 Beals Street, Brookline, MA 02446 USA
+1-617-232-8189   John_L_Bowman@alum.mit.edu   http://JBowman.net

Introduction—Integration Principles and Terminology [1]

This appendix examines model integration in activity-based models of passenger transport demand. Effective model integration is important for forecasting the demand effects of congestion, pricing, unreliability and other changes in transport conditions. After introducing integration principles and terminology in this section, we examine integration features of two AB models.

In simulating the itinerary of one person, many dimensions of choice are modeled, including activity participation, timing and location, as well as the mode of associated travel. In addition, if intra-household interactions are explicitly modeled, it is necessary to identify joint activity and travel participation in various ways. These various aspects of the activity and travel outcome are related to such an extent that it seems appropriate to treat them as a single complex outcome, modeling all dimensions simultaneously. However, the outcome is so complex that it is impractical to capture all the needed detail in a single set of mathematical expression. Therefore, in practice, activity-based models have been implemented as a large number of carefully integrated components. The objective of these models, and hence the objective that guides the selection of integration techniques, is to realistically model travel behavior that can be affected by changes in activity opportunities and travel conditions, such as congestion and reliability, especially those that are affected by public policies and programs, such as pricing. For example, a new peak-period toll on a major facility might affect route choices, departure time choices, mode choices, destination choices, trip chaining decisions, non-travel activity substitution, joint travel arrangements, activity participation, car ownership, work location and residential location. Effective model integration is essential for realistically modeling these effects.

What principles should serve as the foundation for model integration? Consider two different phenomena that affect the activities and travel a person ends up completing in any given day. The first one is the passage of time. Every new action emerges from the situation in the present created by events of the past. It is tempting to thus organize an AB model in the same way, with outcomes modeled in strict temporal sequence. Indeed, models of activity and travel have been developed according to this organizing principle. In such a framework, newly modeled activities can take as given all outcomes that occurred earlier, and must treat as unknown all outcomes that may (or may not) occur later. This does not necessarily preclude taking into consideration the various possibilities for the future; but to do so requires some complex techniques to quantify those future possibilities so that they can be considered in the model's prediction of the present activity.

A second phenomenon is purposeful human planning. People often think ahead, schedule important future activities, and arrange other activities around them in order to achieve their objectives more effectively. It is also tempting to organize an AB model in this second way, with outcomes also modeled sequentially, but according to a plan in which more important activities are modeled first, and less important activities fill the remaining time. This does not preclude taking into consideration, at the time that a more important activity is modeled, the various possibilities for the less important activities. But

[1] Much of this introduction comes from Bowman and Bradley (2008).

to do so requires complex techniques to quantify the possibilities so that they can be considered in the model's prediction of the more important activity.

In reality, the actual activities of most people are probably the result of both planning and the passage of time. Furthermore, most data that are available for developing these models provide an observed itinerary, but little or no information about how planning and the passage of time combined to cause it. Because of this, it is tempting to organize an AB model in the same way, representing a person's (or even an entire household's) day as a single complex outcome, simultaneously representing many components that are highly correlated because of the forces of planning and passage of time that together shaped the person's day. Indeed, because of the practical data limitation, and the importance of both planning and the passage of time, it seems better to formulate a simultaneous model than to formulate a sequential model organized only on planning or only on the passage of time.

However, a person's day (and even more so an entire household's day) is so complex that a simultaneous representation is not feasible. It is too complex to understand all at once, put into a mathematical form and carry out the needed computation. Model developers have been forced to break the outcome into pieces that can be implemented sequentially, specify a model of each of the components, specify important relationships among the components, and integrate them in an attempt to preserve the important relationships.

Without explicitly explaining the need to implement the models in a sequence, which is what gives rise to the concern about adequate integration, Vovsha et al (2005) have used the terms "downward integrity" and "upward integrity" to describe effective integration of sequentially applied models:

"Downward integrity means that all lower-level decisions in the choice model hierarchy are properly conditioned upon the upper-level decisions and take into account a gradually narrowed scope of lower-level choice alternatives as the upper-level choices progress....Downward integrity is ensured by properly sequencing the models, tracking the important variables from choice to choice that accurately describe the feasible scope left for each subsequent choice, and preventing conflicting choices for the same individual.

"Upward integrity means that when modeling upper-level choices the composite measure of quality of the lower-level choices available for each upper-level alternative is properly taken into account."

These terms associate a direction of movement for the model sequence, with the beginning of the sequence at the 'top level' and the end of the sequence at the 'bottom level'. In this appendix we adopt this commonly used—and entirely arbitrary—directional reference. Accordingly, we call downward and upward integrity two aspects of vertical integration.

However, in some cases, it may be important and feasible to retain the simultaneous modeling approach for a portion of the overall outcome consisting of two or more components. This would be the case where there is important complex correlation among component outcomes that can be correctly represented by a known and practical model structure. We call this horizontal integration. Defined in this way, horizontal integration is superior to vertical integration. A sequence of vertically integrated model components is a second best approach, to be used only when horizontal integration is infeasible. The efforts of most academic researchers in this field are often limited to the domain of horizontal integration. In the United States, vertical integration has primarily been the work of consultants carrying out projects with the clear objective of implementing complete model systems that can be used by a public agency for travel forecasting and policy analysis.

Since reality is so complex that full horizontal integration is infeasible, and it is necessary to use a sequential approach with vertical integration, then what organizing principle(s) might be used in choosing the sequence? In the context of a one-day itinerary, where choices are made for only one day, but are often heavily dominated by prior choices with longer term consequences—such as where to live and work, and whether to have a car for every driver in the household—the time horizon of decisions can serve as an organizing principle, with longer term choices coming before shorter term choices in the sequence. The principles of temporal sequence and human planning, described above, are also both good candidates. All three principles have been used in all or most of the existing US AB model systems. Increasingly, modelers have recognized the impact of long-term choices and habits on within-day behavior, specifying more of these models and placing them first in the sequence. For within-day choices, the existing model systems rely primarily on a human planning sequence, with temporal sequence used in some cases for what can be viewed as minor decisions.

Unfortunately, it has not even been feasible for modelers to capture, via vertical integration, all the apparent correlations among the components of a person's one-day itinerary. Nor has it been feasible to test enough variously specified vertically integrated models to state with confidence that the most important correlations have been correctly captured by the selected specifications. As a result, a great deal of modeler judgment undergirds the existing model systems, and will probably continue to do so for the foreseeable future. Judgment guides the modeler's choice of the components to include (and exclude), the components to keep together via horizontal integration, the specific sequence to use for the separate components, and the techniques of vertical integration to employ.

In the next sections, we will look at specific model systems, examining how they break the modeled outcome into component models, and emphasizing especially how they integrate them to capture the effects of transport conditions in the various model components.

DaySim[2]

DaySim is an AB model that uses the day activity schedule approach developed by Bowman and Ben-Akiva (2001). DaySim is the AB component of the travel demand model system called SACSIM that has been used by the Sacramento Area Council of Governments (SACOG) since 2007 (Bradley, Bowman and Griesenbeck, 2010). A version of DaySim is being implemented for the Puget Sound Regional Council (PSRC) and elsewhere with enhanced integration features.

[2] Much of this section comes from Bradley, et al (2010), Bowman and Bradley (2008), Bradley, Bowman and Castiglione (2008) and Bowman, et al (2009).

Figure 1: SACOG DaySim Flow Diagram

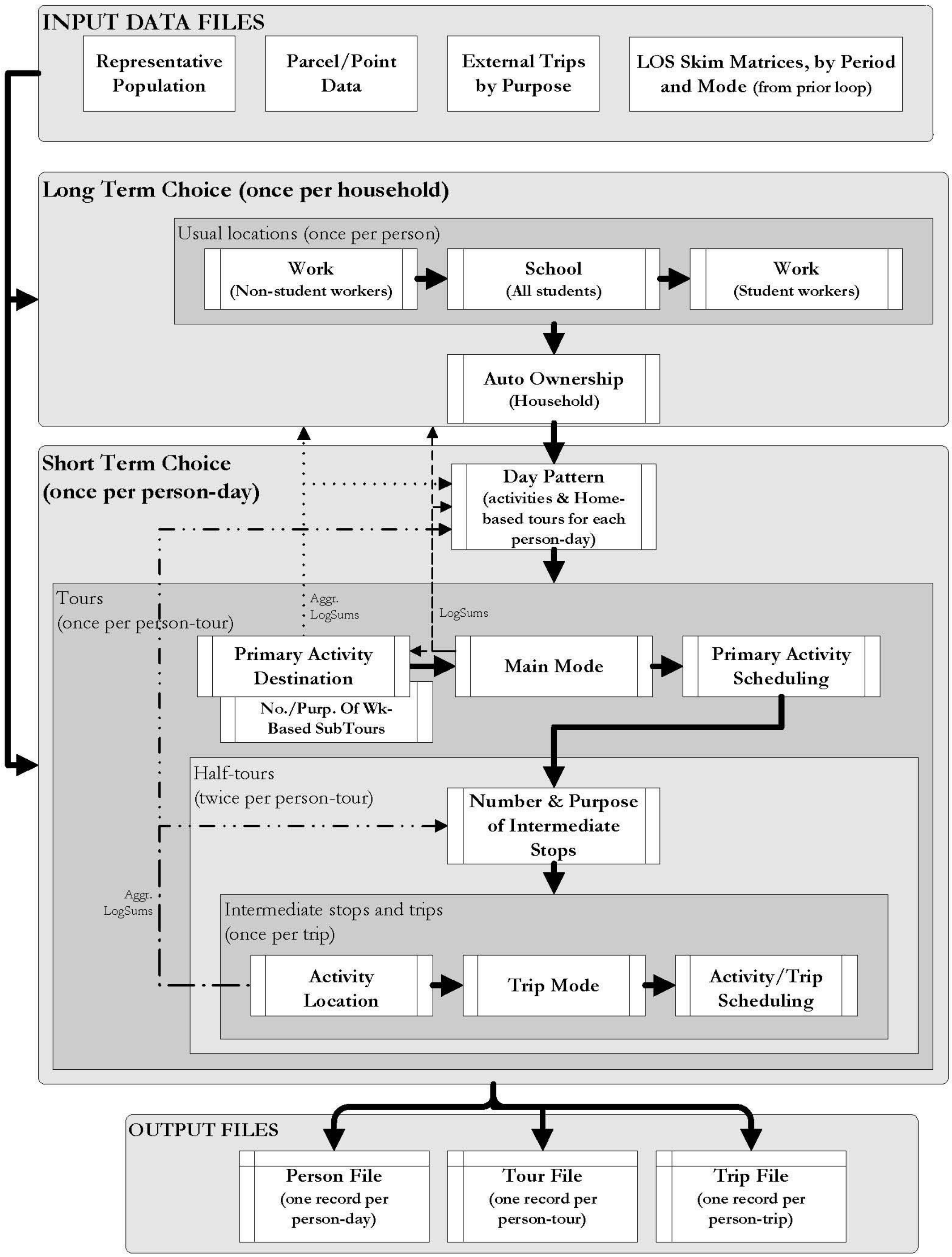
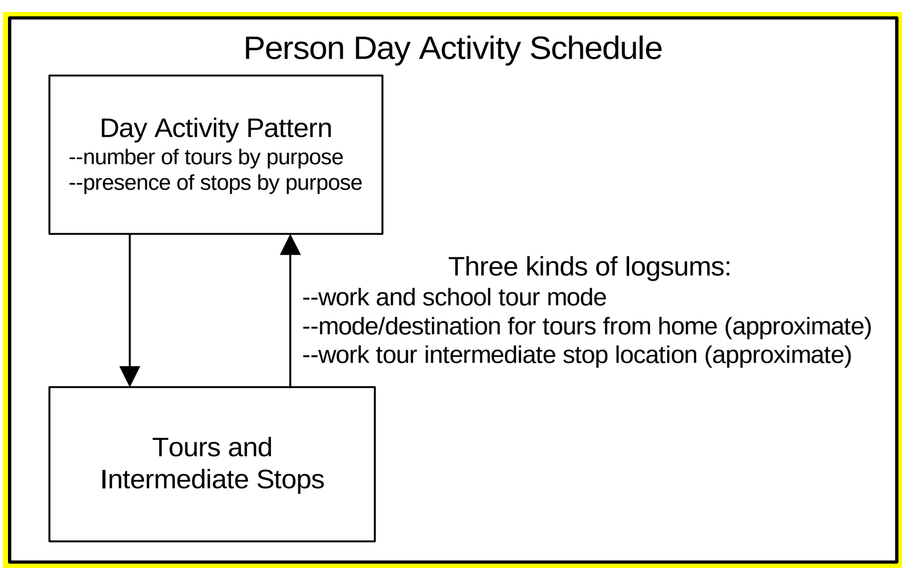


Figure 1 is a flow diagram showing the relationships among DaySim's component models as implemented for SACOG. The hierarchy embodies assumptions about the relationships among simultaneous real world outcomes. In particular, outcomes from models higher in the hierarchy are treated as known in lower level models. It places at a higher level those outcomes that are thought to be higher priority to the decision maker. The model structure also embodies priority assumptions that are hidden in the hierarchy, namely the relative priority of outcomes within a given level of the hierarchy. The most notable of these are the relative priority of tours in a pattern, and the relative priority of stops on a tour. The formal hierarchical structure provides what has been referred to above as downward vertical integrity.

DaySim has a day activity pattern model (see Figure 2) spanning an entire day (Bowman and Bradley, 2006), which provides important horizontal integration to the model system. It identifies the participation in one or more tours for each of seven purposes, and the participation in one or more additional stops (throughout the day) for each of the same seven purposes. This single model provides a complete inventory of the activity purposes in a day and, for each purpose, whether it is the primary objective of a tour and/or a supplemental objective on a tour. This enables the model to realistically capture the mix of tours and stops, by purpose, throughout the day. The utility of a tour or stop for one purpose directly affects its probability, as well as the probability of tours and stops for all other purposes. However, the model does not provide information on the exact number and purpose of the stops on a specific tour, or the positions of the stops on the tour; this is left for models later in the sequence.

The SACOG model includes a horizontally integrated day activity pattern that encompasses tours and trip-chaining for seven purposes. Its upward integration uses three kinds of logsums that account for differences among persons and their available travel destinations and modes, but not times of day.

Figure 2: SACSIM Person Day Activity Schedule



The subsequently simulated tour and stop models are conditioned by the activity pattern, providing downward integration that makes the tour models consistent with the modeled aspects of the day activity pattern.

DaySim uses upward integration to capture the effect of tour and intermediate stop accessibility on long term choices, on the overall pattern of activities and tours in the day, and on tour choices. It does this with logsum accessibility variables for tours and intermediate stops, capturing several types of composite accessibility effects on the pattern, including mode choice logsums for tours to the usual work and school locations, approximate mode/destination logsums for tours from home, and approximate location choice logsums for intermediate stops on work tours.

The approximate, or aggregate, logsum is calculated in the same basic way as a true logsum, by calculating the utility of multiple alternatives, and then taking expectation across the alternatives by calculating the log of the sum of the exponentiated utilities. However, the amount of computation is reduced, either by ignoring some differences among decisionmakers, or by calculating utility for a carefully chosen subset or aggregation of the available alternatives. The approximate logsum is pre-calculated and used by several of the model components, and can be re-used for many persons. This makes it computationally feasible to use logsums at the upper levels of the model system. The categories of decisionmakers and the aggregation of alternatives are chosen so that in all choice cases an approximate logsum is available that closely approximates the true logsum. In essence, this is a sophisticated ad hoc measure that is intended to achieve most of the realism of the true logsum at a small fraction of the cost.

The approximate **tour mode-destination choice logsum** is used in situations where information is needed about accessibility to activity opportunities in all surrounding locations by all available transport modes at all times of day. Because of the large amount of computation required for calculating a true logsum for all feasible combinations in these three dimensions, an approximate logsum is used with several simplifications. First, it ignores socio-demographic characteristics, except for age and car availability. Second, it uses aggregate distance bands for transit walk access. Third, sometimes it uses a logsum for a composite or most likely purpose instead of calculating it across a full set of specific purposes. Finally, instead of basing the logsum on the exact available time window of the choice situation, and calculating it across all of the available time period combinations within the window, it uses a particular available time window size and time period combination. With these simplifications, it is possible to pre-calculate 84 logsums for each zone, and use them when needed at any point in the simulation of any person’s day activity schedule. The 84 logsums represent the combinations of car availability, transit accessibility and purpose as follows:

Car availability:

(1) Child age under 16
(2) Adult in HH with no cars
(3) Adult in HH with cars, but fewer cars than drivers
(4) Adult in HH with 1+ cars per driver

Transit accessibility segments:

(1) Origin is within ¼ mile of transit stop
(2) Origin is more than ¼ mile from transit stop, but walk to transit is available
(3) Walk to transit not available

Non-mandatory tour purpose:

(1) Home-based personal business
(2) Home-based shopping
(3) Home-based meal
(4) Home-based social/recreation
(5) Home-based escort
(6) All home-based purposes combined
(7) Work-based

The logsums are calculated using simplified mode and destination choice models that are estimated using only the above explanatory variables.

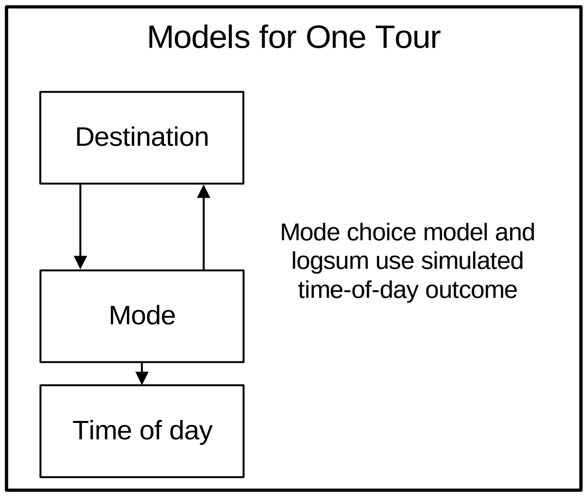
The approximate **intermediate stop location choice logsum** is used in the activity pattern models, making intermediate stops more likely between zone pairs where useful stop locations can be conveniently reached by auto. Peak and off-peak logsums are calculated for each OD zone pair. Each logsum is calculated as the log of the sum over all intermediate stop zones of the utility formula *Size* exp (- 2 * extra time / 6.0 minutes)*, where *Size* is a weighted function of various attraction variables (the size variable function estimated for the composite non-mandatory tour purpose in the aggregate destination choice model), and *extra time* is the auto travel time from the origin zone to the stop zone plus the auto travel time from the stop zone to the destination zone, minus the direct auto travel time from the origin zone to the destination zone (i.e. the detour time required to make the stop on the way from the origin to the destination).

Although the upward integration of the activity pattern is better than in other existing AB models, it is still limited. The approximate logsums are limited as described above. In addition, most of the non-work purposes do not have mode/dest logsums, so some of the benefit of the purpose-specific specification is lost in the upward integration.

For each tour there is a vertically integrated sequence of destination, mode and time-of-day models (see Figure 3). Downward integration prevents unrealistic destination, mode and time-of-day combinations. There is also upward integration from mode to destination, but it does not use time-of-day logsums. Rather, a simulation technique was implemented to make the mode and destination choice models sensitive to policy effects that may vary by time of day. The basic idea is to avoid the use of a logsum (and its associated computational costs) when applying an upper level model by treating as given a conditional outcome that is not known, and would otherwise require the calculation of a logsum from all possible conditional outcomes. In this case the assumed conditional outcome is the tour time-of-day. It is selected by a Monte Carlo draw using approximate probabilities for the conditional outcome. Rather than making every simulated outcome sensitive to variability in the conditional outcome, sensitivity is achieved across the population through the variability of outcome in the Monte Carlo draws. In this way, the mode and destination choice models are sensitive to variations in transport level of service and spatial attributes across all possible combinations of time-of-day, with the affects approximately weighted by the time-of-day choice probabilities.

For each tour, DaySim uses a vertically integrated sequence of destination, mode choice, and time of day (TOD). Mode choice upwardly integrates with destination choice; it uses a simulated TOD that makes the destination and mode choice models sensitive to changes in transport conditions that vary by time of day.

Figure 3: Models for One Tour

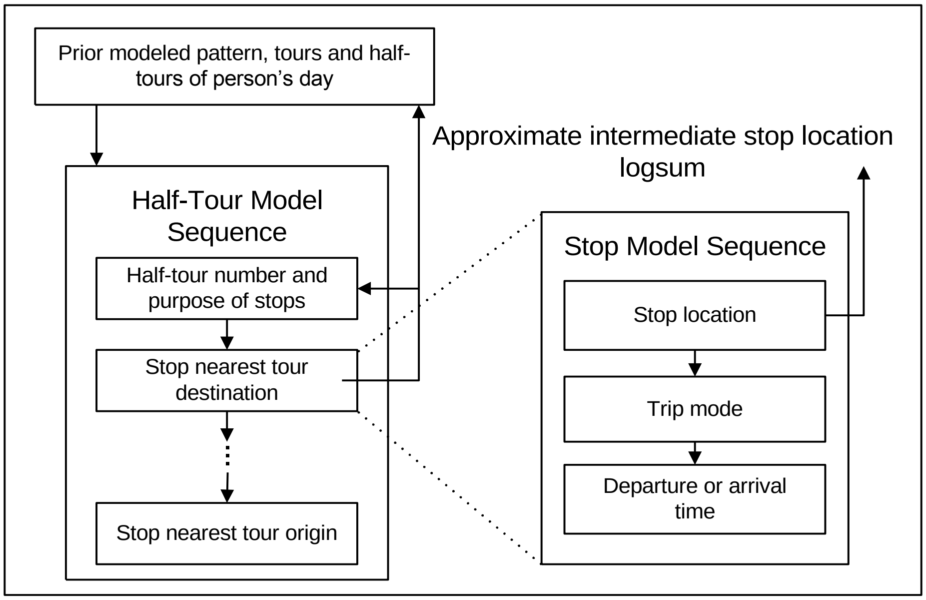


DaySim conditions half-tour stop participation and purpose upon the modeled aspects of the day activity pattern, prior modeled tours, current tour, and—in the case of the second half-tour of the tour—the first half-tour (see Figure 4). This enables consistency of stop participation by purpose among models at all levels, and complements the horizontal and upward integrity of stop purpose enabled by the DaySim specification of the activity pattern model.

The DaySim intermediate stop models of location, mode and timing are conditioned by the same models as the half-tour model, by the half-tour model itself, and by all prior-modeled stop outcomes on the half-tour. The intermediate stops are simulated in sequence emanating from the tour's primary destination, in reverse chronological order for stops before the tour destination, and in chronological order for stops after the tour destination. This is based on the assumed importance of arriving at and departing from the primary destination at a previously modeled time. Accordingly the timing of each intermediate stop is conditioned by the timing of all prior-modeled stops. Similarly, the intermediate stop location choice is conditioned by all previously modeled locations, and the trip mode is conditioned by the mode set allowed by the modeled tour mode and the modes used on all prior modeled trips on the tour. This downward integration allows a consistent and feasible representation of each tour's entire travel itinerary, with regard to timing, mode and location. This is complemented by upward integrity provided by the use of intermediate stop location choice logsums in the half-tour and activity pattern models. These capture the effect of accessibility on the stop participation choices modeled in those two models.

The figure below shows downward integration of half-tour and intermediate stop models conditioned by the pattern, tour, half-tour and prior-modeled stop models, with accompanying upward integration via approximate stop location logsums.

Figure 4: Downward integration example



DaySim includes long-term choices of usual work location for each worker and usual school location for each student. For young student workers, the usual school location model conditions the usual work location model, and for old worker students the sequence is reversed. These condition a downwardly integrated household vehicle ownership model, and all of them are downwardly integrated with the day activity schedule of all household members. All three long-term models are upwardly integrated via work and/or school tour mode choice logsums, which measure accessibility to the usual work and school locations. They also use approximate mode-destination logsums. For the usual location models the mode/dest logsums measure accessibility for tours from the usual location. For auto ownership, they measure accessibility for tours from home.

In summary, DaySim implements downward integrity of the model system with regard to the modeling and accounting of participation, time-of-day, mode and location of intermediate (chained) stops on all tours of an individual in a day. Just as importantly, it emphasizes and implements techniques for achieving upward integration to help the model system capture sensitivity to travel conditions at all levels. Table 1 summarizes, for each DaySim model, the explanatory variables that capture the effects of transport conditions on the model's predictions, including those that measure impedance directly as well as those that do it with vertical integration via logsum measures.

Table 1: Effect of transport conditions in the DaySim component models

	Residential location	Usual work location	Usual school location	Auto ownership	Day activity pattern	Work-based tour generation	Tour destination choice	Tour mode choice	Tour time of day choice	Stop frequency and purpose	Intermediate stop location	Trip mode choice	Trip time of day choice
Parcel-level accessibility variables													
Parking density		X					X				X		
Average parking price				X				X				X	
Street intersection density		X			X		X	X		X	X	X	
Distance to nearest transit stop				X	X			X				X	
Zone-level accessibiliy variables													
Auto and transit costs								X	X		X	X	X
Auto, transit and non-motorized times								X	X		X	X	X
Transit connectivity/availability								X	X		X	X	X
Auto time on very congested links									X				X
Driving distance		X	X				X	X			X		
Mode choice accessibility logsum		X	X	X	X		X						
Mode/dest logsums (accessibility at destination)		X	X				X						
Mode/dest logsums (accessibility at home)				X	X								
Intermediate stop accessibility logsums					X					X			

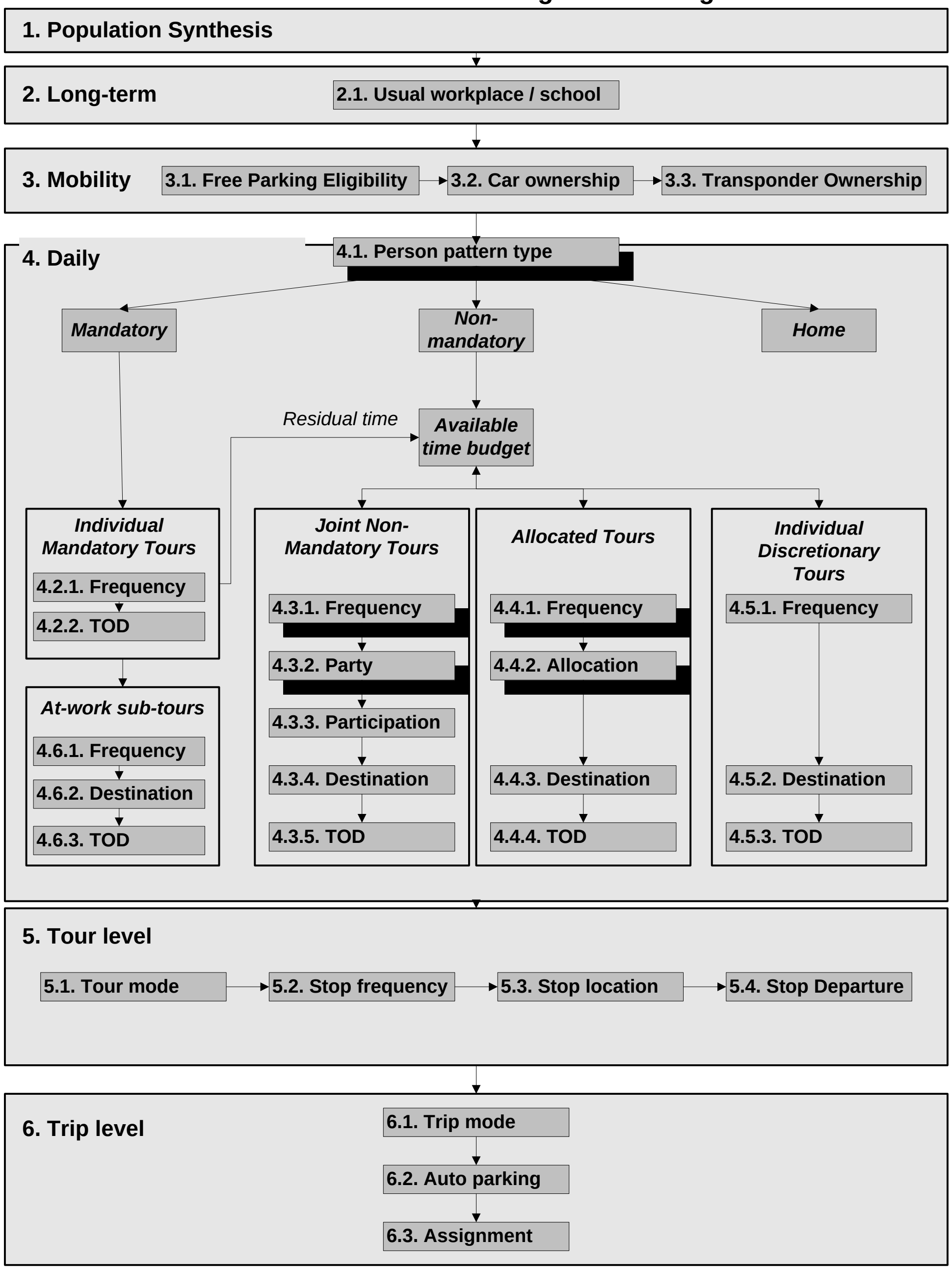
The AB model system being implemented for PSRC is an enhanced version of DaySim, to be integrated with PSRC's UrbanSim land use model within the Opus software framework. The PSRC DaySim model will have the integration features described above for SACOG DaySim, with some important enhancements. The introduction of explicit models of joint household outcomes will provide horizontal integration across household members that SACOG DaySim lacks. At the same time, DaySim's careful use of upward integration to make the upper level models sensitive to transport conditions, including congestion and tolls, will be extended upward to include the new household models.

The PSRC design also includes expanded and consistent use of time-dependent logsums. The benefit of this is that the models that use logsums will be more realistically sensitive to time-of-day-specific changes of travel time and cost. There are three basic types of logsums proposed for the PSRC model system (as used in the SACOG model): disaggregate tour mode choice, aggregate tour mode-dest choice, and aggregate intermediate stop location choice. Since, in the model structure, the time-of-day models are conditioned by the destination and mode choice models, the logsum calculations use an assumed time of day. In the SACOG model, in a few cases the assumed time of day is drawn from a time-of-day distribution for tours of the type for which the logsum is needed. But in most cases, an arbitrary "most likely" time-of-day is assumed. In the PSRC implementation, the plan is to use simulated times of day in all mode and mode-destination logsums, drawing the time-of-day in all cases, from appropriate time-of-day distributions. This will make the logsums sensitive to time-of-day-specific changes in travel time and cost.

SANDAG CT-RAMP[3]

An AB model system called SANDAG CT-RAMP is currently under development for the San Diego Association of Governments (SANDAG). It belongs to the family of models that also includes Columbus (MORPC), Atlanta (ARC) and the San Francisco Bay Area (MTC). The general design of the SANDAG CT-RAMP model is presented in **Figure 5** below.

Figure 5: SANDAG CT-RAMP Basic Model Design and Linkage between Sub-Models



3 Much of this section comes from Freedman et al (2009) and an internal project document of the SANDAG AB model development project written by Parsons Brinckerhoff.

Choices that relate to the entire household or a group of household members and assume explicit modeling of intra-household interactions (sub-models 3.2, 4.1, 4.3.1, 4.3.2, 4.4.1, and 4.4.3) are shadowed in Figure 5. The other models are assumed to be individual-based for the basic design.

The planned SANDAG AB model system includes accessibility measures in upper level models that play the role of simplified tour-level logsums, used instead of full logsums. They are created in order to reflect the opportunities to implement a travel tour for a certain purpose from a certain origin (residential or workplace). Since they only represent a certain proxy for full logsums there is no automatic guarantee that the entire AB would exhibit a completely consistent sensitivity to any transportation LOS improvements.

Each accessibility measure is stored as a vector of values (one value per zone). In the AB application, accessibility measures are recalculated at each global iteration as part of the overall equilibration process. For this reason, it is important to make this calculation computationally efficient, especially if a large number of zones is involved (as is the case of SANDAG with 32,000 MGRAs (Master Geographic Reference Areas)).

In the SANDAG AB model, these accessibility measures will explain cross-sectional differences in travel generation across area types that are quite significant. Additionally, through these measures, a better linkage between the ABM and the land use model, PECAS, will be ensured.

General form of the accessibility measures

The SANDAG accessibility measures have the following general form:

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

Where:

$i, j \in I$ = origin and destination zones,

A_i = accessibility measure calculated for each origin zone,

S_j = attraction size variable for each potential destination,

c_{ij} = cost of travel between origins and destinations,

γ = dispersion coefficient.

In this form, an accessibility measure is essentially a sum of all attractions in the region discounted by the travel impedance. It can be equivalently viewed as a destination choice logsum. The dispersion coefficient expresses a sensitivity of the given type of activity to travel cost, i.e. travelers' tolerance to longer travel times in order to participate in the given activity. Larger dispersion coefficients reflect a greater sensitivity to travel times and costs, reflecting more localized activity types.

In the form of equation 1, an accessibility measure can be linearly included in a utility function of an upper-level model. To preserve consistency with the random-utility choice theory, the coefficient for any accessibility measure should be between 0 and 1; though it is not as restrictive as in a case of a proper nested logit model.

Accessibility measures are operationalized in four steps:

- Specification of **segmentation** (i.e. the number of different measures for different activity types that are going to be used in the AB). In the SANDAG AB we need these measures primarily to support the car ownership model and DAP / tour-generation models for non-mandatory purposes. This leads to at least 3 mode-specific measures (for modeling car-ownership assuming that only 3 major modes are considered: Auto/SOV, Walk-to Transit, and Non-motorized) and 6 purpose-specific measures (for modeling tour generation for escorting, shopping, maintenance, eating out, visiting, and discretionary purposes).

- Specification of **size variables,** i.e. zone attractors for the given segment that are normally modeled as a combination of certain LU variables (like for tour and trip destination choice).

- Specification of **transportation cost**. Two major approaches can be taken for calculation of transportation cost: 1) mode-specific, where an accessibility measure is calculated separately for each main mode; 2) mode-choice-logsum-based where all modes are converted into a single measure. The second method is more complicated computationally and requires an additional step for estimation or calibration of a simplified mode-choice model. However, it has certain theoretical and practical advantages for tour-generation models, as explained elsewhere. In both cases, a one-directional cost is calculated instead of a round-trip cost.

- Specification of **time-of-day (TOD)** period for each measure. Transportation cost is also dependent on time-of-day (TOD) period. To avoid the complexity associated with calculating logsums over TOD periods, normally a certain representative TOD period is assumed for each accessibility measure. Alternatively, several accessibility measures for each TOD period can be statistically tested in the model estimation where the best one is chosen. However, it is normally difficult to include several accessibility measures of the same type calculated for different TOD periods in the same model because of correlation between them.

The dispersion coefficient, as well as other parameters embedded in the size variable and transportation cost function, can be either estimated (which means an estimation of a simplified mode choice model and/or destination choice model) or asserted based on the normative values of the parameters known from the previously estimated mode and destination choice models.

The SANDAG accessibility measures play a similar role as, and can be compared to, the approximate mode-destination choice logsums used in the SACOG DaySim model system, and the accessibility measures used in the prior CT-RAMP model systems (MORPC, ARC and MTC). SACOG DaySim uses 84 distinct mode-destination logsums for each zone in the region. The MORPC / ARC / MTC accessibility measures are extremely simple (only 10 types are used per zone). The planned SANDAG ABM structure falls between these two examples.

The set of accessibility measures planned for the SANDAG ABM model is summarized in Table 4.

Table 4: Accessibility Measures for the SANDAG AB

No	Description	Model where it is used	Attraction size variable S_j	Travel cost c_{ij}	Dispersion coefficient $-\gamma$
1	Access to non-mandatory attractions by SOV in off-peak	Car ownership	Total weighted employment for all purposes	Generalized SOV time including tolls	-0.05
2	Access to non-mandatory attractions by transit in off peak	Car ownership	Total weighted employment for all purposes	Generalized best path walk-to-transit time including fares	-0.05
3	Access to non-mandatory attractions by walk	Car ownership	Total weighted employment for all purposes	SOV off-peak distance (set to 999 if >3)	-1.00
4-6	Access to non-mandatory attractions by all modes	CDAP	Total weighted employment for all purposes	Off-peak mode choice logsums (SOV skims for ipersons, HOV skims for interaction) segmented by 3 car-availability groups	+1.00
7-9	Access to shopping attractions by all modes except SOV	Joint tour frequency	Weighted employment for shopping	Off-peak mode choice logsum (HOV skims) segmented by 3 HH adult car-availability groups	+1.00
10-12	Access to maintenance attractions by all modes except SOV	Joint tour frequency	Weighted employment for maintenance	Off-peak mode choice logsum (HOV skims) segmented by 3 adult car-availability groups	+1.00
13-15	Access to eating-out attractions by all modes except SOV	Joint tour frequency	Weighted employment for eating out	Off-peak mode choice logsum (HOV skims) segmented by 3 adult HH car-availability groups	+1.00
16-18	Access to visiting attractions by all modes except SOV	Joint tour frequency	Total households	Off-peak mode choice logsum (HOV skims) segmented by 3 adult car-availability groups	+1.00
19-21	Access to discretionary attractions by all modes except SOV	Joint tour frequency	Weighted employment for discretionary	Off-peak mode choice logsum (HOV skims) segmented by 3 adult car-availability groups	+1.00
22-24	Access to escorting attractions by all modes except SOV	Allocated tour frequency	Total households	AM mode choice logsum (HOV skims) segmented by 3 adult car-availability groups	+1.00
25-27	Access to shopping attractions by all modes except HOV	Allocated tour frequency	Weighted employment for shopping	Off-peak mode choice logsum (SOV skims) segmented by 3 adult car-availability groups	+1.00
28-30	Access to maintenance attractions by all modes except HOV	Allocated tour frequency	Weighted employment for maintenance	Off-peak mode choice logsum (SOV skims) segmented by 3 adult car-availability groups	+1.00
31-33	Access to eating-out attractions by all modes except HOV	Individual tour frequency	Weighted employment for eating out	Off-peak mode choice logsum (SOV skims) segmented by 3 car-availability groups	+1.00
34-36	Access to visiting attractions by all modes except HOV	Individual tour frequency	Total households	Off-peak mode choice logsum (SOV skims) segmented by 3 car-availability groups	+1.00
36-39	Access to discretionary attractions by all modes except HOV	Individual tour frequency	Weighted employment for discretionary	Off-peak mode choice logsum (SOV skims) segmented by 3 car-availability groups	+1.00
40-41	Access to at-work attractions by all modes except HOV	Individual sub-tour frequency	Weighted employment for at work	Off-peak mode choice logsum (SOV skims) segmented by adult 2 car-availability groups (0 cars and cars equal or graeter that workers)	+1.00

Concluding Discussion

The purpose for examining the AB model systems at SACOG, PSRC and SANDAG has been to bring to light important model integration features, existing and needed, that could improve the effectiveness of model systems in forecasting the effects of transport conditions, especially congestion, pricing and reliability. This section identifies some possible conclusions to draw from this examination.

Since the existing AB models consist of many model components applied in sequence, few of which can directly include measures of travel conditions (time, cost, reliabilty, availability), upward integration is of great importance. These "accessibility" measures need to accurately incorporate the effects of changes in conditions (including tolls, congestion and reliability) that are specific by mode, location and time of day, and they need to accurately differentiate the effects depending on the characteristics of the traveler (such as age, VOT and auto availability) and the travel situation (such as purpose and location relative to work, school and home). The methods being implemented for SANDAG and PSRC need to be evaluated and can serve as a starting point for further improvements of upward integration.

Downward integration is also important. The methods employed in DaySim and CT-RAMP to condition lower level models on upper level outcomes—the restriction of available times, modes and locations—are important, and can serve as a starting point for further improvements of downward integration.

Horizontal integration is also important. The use of individual day activity patterns provides an important aspect of integration that enables direct trade-offs between tour generation and trip-chaining. The use of coordinated activity patterns among household members enables the models to generate schedules that are consistent among household members, and the modeling of joint activities and travel enables the modeling of sensitivities to travel conditions that may be different than for household members traveling alone. Further improvements are also needed in the horizontal integration of the models of destination, mode and schedule, so that correlations among these three important aspects of travel choice (and the associated sensitivity to travel conditions) can be more realistically represented.

Bibliography

Bowman, John L and Moshe Ben-Akiva (2001) Activity-based disaggregate travel demand model system with activity schedules, Transportation Research Part A, 35 1-28. Pre-publication draft available at http://jbowman.net.

Bowman, John L. and Mark Bradley (2006), Activity-Based Travel Forecasting Model for SACOG: Technical Memo Number 6: Day Pattern Activity Generation Models, July 31, 2006 – Draft 2, 21 pages. Available at http://jbowman.net.

Bowman, John L and Mark A Bradley (2008) Activity-Based Models: Approaches to Achieve Integration Among Trips and Tours Throughout the Day. Presented at the 2008 European Transport Conference, Leeuwenhorst, The Netherlands, October, 2008. Available at http://jbowman.net.

Bowman, John L, Mark Bradley, Joe Castiglione, Urban Analytics and Resource Systems Group (2009) Development and Implementation of an Activity-Based Transport Model System: Report 1: Design Issues and Their Resolution, Working Draft. Project document available at http://jbowman.net.

Bradley, Mark A, John L Bowman and Joe Castiglione (2008) PSRC Activity Model Work Plan & Activity Generation Model: Work Plan Report—Draft 2, project document available at http://jbowman.net

Bradley, Mark, John L Bowman and Bruce Griesenbeck (2010) SACSIM: An applied activity-based model system with fine-level spatial and temporal resolution. Presented at the 11th World Conference on Transport Research, Berkeley, CA, June 2007. Publication pending in Journal of Choice Modelling (www.jocm.org.uk). Pre-publication draft available at http://jbowman.net.

Freedman, Joel, Peter Vovsha, Wu Sun, Surabhi Gupta, Ziying Ouyang, Rick Curry and Clint Daniels (2009) Space and Accessibility in the SANDAG Activity-Based Model. Submitted for presentation at the 3rd TRB Conference on Innovations in Travel Modeling, Tempe, Arizona; May 9-12, 2010.

Vovsha, Peter., Mark A. Bradley and John L Bowman (2005) Activity-based travel forecasting models in the United States: Progress since 1995 and Prospects for the Future, in Progress in Activity-Based Analysis, H Timmermans, ed., Elsevier, pp 389-414. Pre-publication draft available at http://jbowman.net.

Appendix: Conceptual and Technical Issues of ABM-DTA Integration

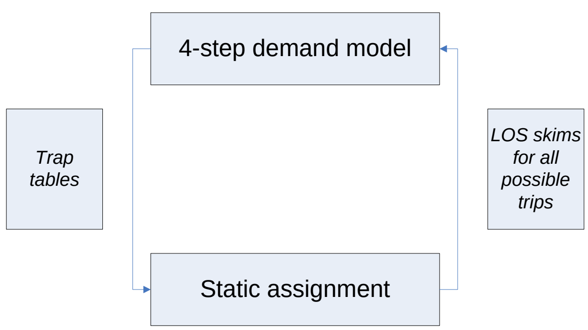
Prepared by

Peter Vovsha
Parsons Brinckerhoff
1 Penn Plaza, 2nd Floor, New York, NY 10019
212-465-5511

Vovsha@PBworld.com

Since both technologies of microsimulation on the demand side (ABM) and supply (network) side (DTA) have been brought to a certain level of maturity, the perspective of AB-DTA integration has become one of the most promising avenues in transportation modeling. Seemingly, the integration between too models should have been natural and straightforward as was the integration concept between a 4-step model and static traffic assignment (STA) shown in **Figure 35**. The integration is based on the fact that both I/O entities involved in the process have the same matrix structure. The demand model produces trip tables needed for assignment and the assignment procedures produces full Level-of-Service (LOS) skims in a matrix format needed for the 4-step model. Note that the LOS variables are provided for all possible trips (not only for the trips generated by the demand model at the current iteration). In this case we can say that the network model provides a full feedback to the demand model. The theory of global demand-network equilibrium is well developed for this case and guarantees a unique solution for the problem as well as effective practical algorithms.

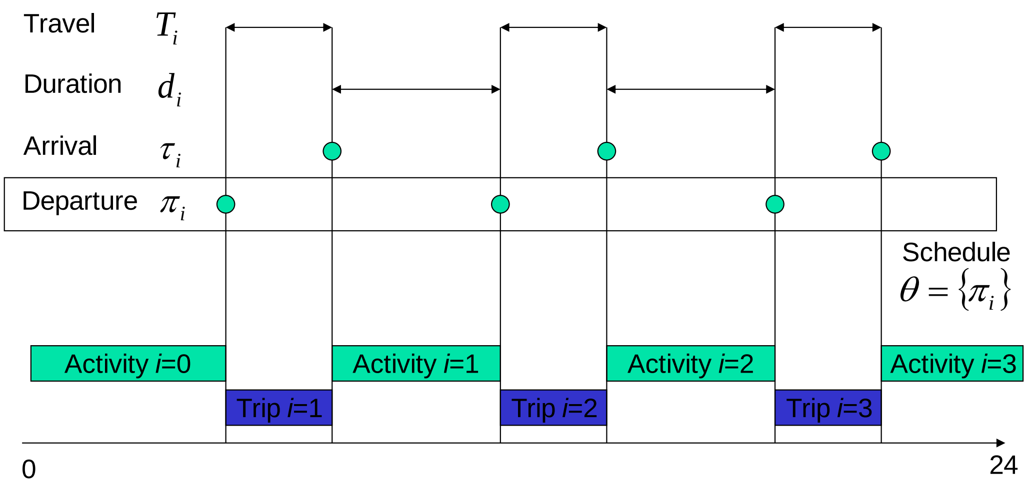
Figure 35: Integration of 4-Step Model and Static Assignment



Both ABM and DTA operate with individual particles as modeled units (individual tours and trips) and have compatible levels of spatial and temporal resolution. It may look that exactly the same integration concept as applied for 4-step models should be just adjusted to account for a list of individual trips instead of fractional-number trip tables.

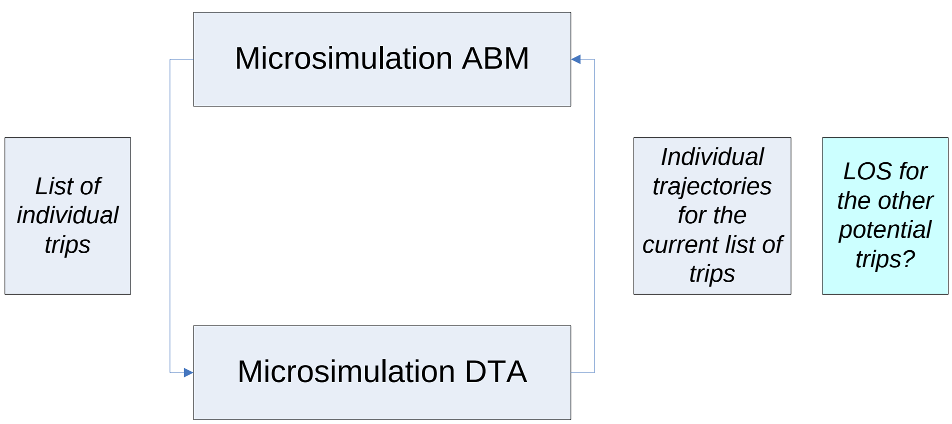
Moreover, the advanced individual ABM-DTA framework would provide a new beneficial dimension for integration in a form of individual schedule consistency (that can never be incorporated in an aggregate framework). Individual schedule consistency means that for each person, the daily schedule (i.e. a sequence of trips and activities) is formed without gaps or overlaps as shown in **Figure 36**. This way, any change in travel time would affect activity durations and vice versa.

Figure 36: Individual Schedule Consistency



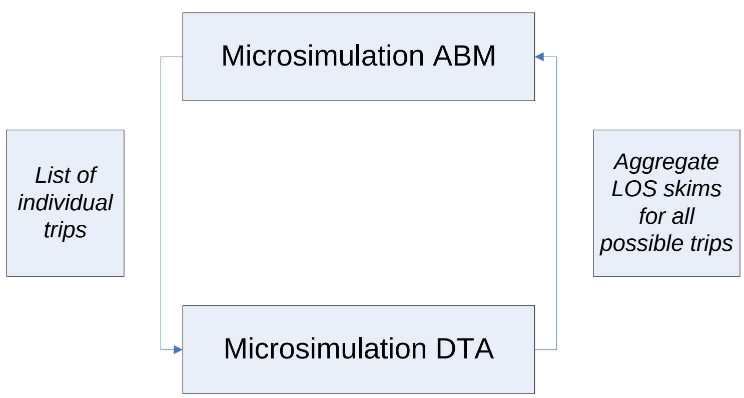
However, a closer look at the ABM-DTA framework and actual technical aspects of implementation reveals some non-trivial issues that have yet to be resolved before the advantages could be taken of the overall microsimulation framework. The problem is specifically with the feedback provided by the DTA procedure that does not cover all needs of the ABM as shown in **Figure 37**.

Figure 37: Integration of ABM and DTA (Direct)



The crux of the problem is that contrary to the 4-Step-STA integration, the microsimulation DTA can only produce an individual trajectory (path in time and space) for the list of actually simulated trips. It does not automatically produce trajectories for all (potential) trips to other destinations and at other departure times. Thus, it does not provide the necessary feedback for ABM at the disaggregate level. Any attempt to resolve this issue by brute force would result in infeasible number of calculations since all possible trips cannot be processed by DTA at the disaggregate level. In fact, the list of trips for which the individual trajectories can be produced is a very small share of the all possible trips to consider. One of the possible solutions is to employ DTA to produce crude LOS matrices (the way they are produced by STA) and use these LOS variables to feed the demand model as shown in **Figure 38**. This approach however, would lose most of the details associated with DTA and advantages of individual microsimulation (for example, individual variation in Values of Time or other person characteristics) in the aggregation of individual trajectories into crude LOS skims. Essentially with this approach, the individual schedule consistency concept would make a very limited value because travel times referred to in **Figure 36** will be very crude.

Figure 38: Integration of ABM and DTA (Aggregate Feedback)



Several other ideas are currently being considered and tested in SHRP 2 C10 and L04 projects. Two of them are presented in **Figure 39**. The first idea is based on the fact that a direct integration at the disaggregate level is possible along the temporal dimension of the other dimensions (number of trips, order of trips, and trip destinations) are foxed for each individual. Then, a full advantage can be taken of the individual schedule constraints and corresponding effects as shown in **Figure 36**. The second idea is based on the fact that trip origins, destinations, and departure times can be pre-sampled and the DTA process would only be required to produce trajectories for a subset of origins, destinations, and departure times. In this case, the schedule consolidation is implemented though corrections of the departure and arrival times (based on the individually simulated travel times) and is employed as an inner loop. The outer loop includes a full regeneration of daily activity patterns and schedules but with a sub-sample of locations for which trajectories are available (it also can be interpreted as a learning and adaptation process with limited information).

Figure 39: Integration of ABM and DTA (Split Feedback)