INTRODUCTION

1.1 The Need for Determining Travel Demand: Existing and Future

The basic purpose of transportation planning and management is to match transportation supply with travel demand, which represents ‘need’. A thorough understanding of existing travel pattern is necessary for identifying and analyzing existing traffic related problems. Detailed data on current travel pattern and traffic volumes are needed also for developing travel forecasting/prediction models. The prediction of future travel demand is an essential task of the long-range transportation planning process for determining strategies for accommodating future needs. These strategies may include land use policies, pricing programs, and expansion of transportation supply – highways and transit service.

1.2 Scope of Analysis and Levels of Planning

There are different levels of planning directed to different types of problems. The terminology for these levels of planning and analysis varies according to the context. For example, the expressions ‘micro’, ‘meso’, and ‘macro’ are sometimes used to describe the level of detail or the size of an area used for an analysis. Similarly, the expressions ‘site specific’, ‘corridor’, and ‘areawide’ or ‘metropolitan’ are used to describe variations in the scope of a problem.
The approach and techniques for analyzing and forecasting travel would vary according to the level of analysis. Even for a particular level of analysis, the techniques may have to be adjusted to match the constraints of available data and manpower.

An example of a micro-level or site-specific analysis is the case of a congested road intersection. In this case traffic engineers would be interested in detailed traffic flow characteristics including turning movements of vehicles along each approach, and pedestrian volumes across each approach. Management strategies in this case would involve traffic operation and roadway design oriented techniques. A corridor level analysis on the other hand would cover a larger area, say, ten miles long and two miles wide. A major highway with severe congestion problem may require a corridor analysis. The origin and destination of trips, and modal choice of travelers would be of interest in this case. Station-to-station movements of passengers may have to be estimated in the case of a rapid transit service along the corridor. At the macro level the concern may be total energy consumption by the transportation sector or the total emission of an air pollutant; for these cases, information on total vehicle-miles traveled (VMT) on each functional class of roads will be needed.

It is important to recognize that the nature of problems to be examined dictates the level of planning to be used as well as the technique for travel demand analysis. The discussion of this chapter will be oriented mostly to ‘meso’ scale or areawide travel demand analysis that is commonly performed in urban transportation planning studies. Even for this type of analysis for an urban area at the ‘meso’ scale, the approach and details of techniques and models to be used would depend on the size of the area as well as the resources available for carrying out the work. For example, a small urban area may not have the manpower or funding needed for carrying out large-scale surveys and developing advanced mathematical
models. The need for customizing the planning and modeling approaches based on specific situations was discussed in detail in the National Cooperative Highway Research Program (NCHRP) Report 167 by Grecco, W. L., et al.

2 CHARACTERISTICS OF TRAVEL

There are certain special characteristics of travel demand that require recognition for planning and design purposes, and these are discussed below:

2.1 Spatial and Temporal Variations

The total magnitude of travel demand alone is not sufficient for detailed planning and management purposes. The spatial and temporal distributions of travel also are important items of information to be considered in determining supply strategies. The peaking of travel at certain time periods requires a level of transportation supply that is not needed at other times. However, due to the nature of supply which cannot be adjusted easily, large investments have to be made to provide roadway or transit service capacities to accommodate peak period travel, and this capacity is not utilized efficiently at other times. An imbalance in the directional distribution of travel also creates similar inefficiencies.

The spatial orientation of trips has important influence on supply requirements and costs. A few typical spatial distribution patterns of trips in urban areas are listed below:

- Travel along dense corridors, which are usually radial connecting suburbs to central business district (CBD)
- Diffused travel pattern caused by urban sprawl
• Suburb to suburb or circumferential travel

• Travel within large activity centers in CBD and suburbs

Different modes of transportation may be needed to serve these different travel patterns. For example, fixed-route public transit service usually is efficient for concentrated travel along a dense corridor, but it is not ideally suited to serve a diffused travel pattern in a cost-effective manner.

Choice of domicile and workplace, lifestyles and different travel needs of individuals and families make the comprehension of trip making characteristics of a large metro area very complex. These complexities may be illustrated through trips made by a typical suburban US household on a given weekday (Figure 1). Assume that this household has four members, including two kids who go to a grade school, and two cars. It can be seen that there are at least 11 trips made by this household at different times of day. Most of the trips are auto trips and two trips are taken in the “walk” mode. Travel demand modeling attempts to capture such spatial and temporal variations in travel at an aggregate level, such as a zone, in which a number of households, businesses and offices exist.

2.2 Classification of Travel by Trip Purpose and Market Segments

In addition to the spatial and temporal characteristics of travel demand, there are several other aspects of travel demand which must be recognized. ‘Trip purposes’ such as work, shopping, and social-recreation; and trip maker’s characteristics such as income and car ownership, are important factors influencing the elasticity of demand reflecting its sensitivity with respect to travel time and cost. For example, ‘work’ trips may be more likely to use
public transit for a given level of service than trips of other trip purposes.

For a metropolitan study, it is useful to classify travel according to spatial orientation and trip purpose as shown in Figure 2. The concept of “market segmentation” is applicable to the classification of travel based on trip purpose, trip makers’ characteristics, and spatial-temporal concentration. This concept is used in the field of ‘marketing’ for developing different types of consumer products targeted to match different tastes and preferences of potential users/buyers of these products. The concept of market segmentation is applicable to public transportation planning. A single type of transit service is not suitable for all transit market segments. For example, express buses may be needed for a commuter market segment. Taxicabs serve a different market segment. NCHRP Report 212 by Woodruff, et al. examined this subject in depth.

3 UNITS FOR MEASURING TRAVEL DEMAND

Travel demand is measured and expressed in different ways for different types of analysis. Examples of different units of measurement are:

a. Trip (between two areas)

b. Trip end (in a given area)

c. Traffic volume (on a road segment)

d. Person trip and vehicle trip

e. Passenger vehicle and freight vehicle

f. Person-mile traveled and vehicle-mile traveled
The definition of each of these units should be understood clearly, and an appropriate unit of measurement should be used to match the case being analyzed. For example, for a parking study, “trip end” is the appropriate unit for expressing parking demand. For estimating the number of lanes to be provided in a road segment, the demand should be expressed in terms of “traffic volume”. As it was pointed out earlier, the appropriate unit of travel for estimating fuel consumption and/or air pollution attributable to transportation is vehicle miles traveled (VMT).

4 MEASURING EXISTING TRAVEL

Detailed information on existing travel is needed for two purposes – 1) analyzing existing problems, and 2) developing mathematical models for forecasting travel. A variety of surveys can be performed for gathering information related to existing travel demand. However, travel surveys are expensive, and, therefore, care must be taken to identify the types of information that really would be useful for specific purposes, and then the most suitable procedures should be selected for gathering the information. Sampling techniques are useful, and adequate time and care must be devoted to developing sampling procedures. There are several different types of survey techniques, some of which are suitable for automobile travel, some for transit travel, and some for general passenger movement. Survey procedures for freight vehicles and commodity movements may be very different in certain respects from those of passenger travel. A few good references for travel demand related survey techniques are included in the list of references: Dresser and Pearson, Stopher and Metcalf, and Travel Survey Manual.
4.1.1 Time Frame for Travel Surveys.

Since travel demand varies during a given year according to the season (or month of year) and day of week, a decision must be made carefully to select a specific time frame or reference for surveys. For urban transportation studies it is a common practice to develop travel demand information for an average weekday in the ‘fall’ season. However, the time can be different based on the nature of the problem to be analyzed. For example, in the case of a tourist oriented urban area the major concern may be traffic problems during weekend days or holidays, and surveys may be done to capture information for one of those days. A few major types of surveys are discussed in the following sections.

4.1.2 Origin-Destination (O-D) Surveys

The classification of trips into the three classes of internal, external-internal (and vice-versa), and through trips is useful for ‘meso’ scale and metropolitan level as well as small area studies. This classification scheme is useful for developing forecasting procedures/models as well as policies and strategies for accommodating travel because strategies for each of these classes of travel would be different. For example, through trips may require a by-pass facility. External-internal trips may have to be intercepted before they reach a heavily congested area such as the central business district (CBD).

The origins and destinations of trips along with some other characteristics such as trip purpose and mode of travel can be determined in different ways:

a. Home interviews (for internal travel)

b. Roadside interviews at cordon stations (for external-internal and
c. On-board survey on transit vehicles.

All of these three techniques involve “sampling” and require careful planning before their execution. Several years ago Federal Highway Administration and the former Urban Mass Transportation Administration developed detailed guidelines for these survey procedures documented in reports titled *Urban Origin-Destination Surveys*, and *Urban Mass Transportation Travel Surveys* respectively, which are included in the list of references. The reliability of the results of an O-D survey depends on its sampling scheme and sample size, and this issue was examined by Makowski, Chatterjee and Sinha (1974).

Full-scale origin-destination surveys were widely used during 1960’s and 1970’s to develop a variety of information including “desire lines” of travel. Their use has decreased because of the cost and also due to the use of synthetic or borrowed disaggregate travel models, which require less survey data.

### 4.1.3 Traffic Volume and Passenger Counts

For determining the use of various roadway facilities and assessing their level of service, vehicle counts are taken at selected locations along roadways. Short-count techniques are useful provided appropriate expansion factors are developed based on previous or ongoing research on fluctuations of traffic by hour, by weekday, and by month. All state Departments of Transportation (DOTs) have extensive programs for gathering traffic volume data on an annual basis. These vehicle counts usually are taken with machines.

For urban transportation studies ‘screen lines’ and ‘cut-lines’ are established in
the study area to select traffic count locations and take counts in an organized manner so that the major travel movements can be measured and analyzed. These counts are also used for checking the results of travel forecasting models. Similarly traffic counts are taken at special traffic generators such as an airport and a large college/university to capture their unique travel generating characteristics.

For analyzing the use of a transit service, passenger counts are taken on-board transit vehicles and/or at selected stops or stations. These passenger counts usually are taken by observers who are assigned to specific transit vehicles and/or transit stops/stations according to a survey plan.

5 FORECASTING FUTURE DEMAND

The need for travel demand forecasts arises in various contexts of planning -- short-range as well as long-range. Travel forecasting is one of the most important and difficult tasks of transportation planning. There are different types of travel prediction techniques, and the one to be used in a particular case must be compatible with the nature of problem and scope of planning. Constraints of available time and resource also influence the selection of a technique.

5.1 Predicting Response to Service Changes Using Elasticity Coefficients

For short-range planning or a Transportation Systems Management (TSM) study, it is often necessary to predict the effect of a proposed change in transportation service that can be implemented in the near future. For example, a planner may be asked to evaluate the impact on transit ridership of improving transit service in a travel corridor by providing increased frequency and/or limited stop service. The impact of changing the fare structure on transit
ridership also may be of interest. In these cases demand ‘elasticity coefficients,’ if available from past studies, would be useful. Typically an ‘elasticity coefficient’ is developed with respect to a specific ‘factor’ such as travel time or fare based on actual observation. The coefficient should exclude the effect of other factors that also may be influencing demand at the same time. A report authored by Mayworm, Lago and McEnroe (1980) titled * Patronage Impacts of Changes in Transit Fares and Services* contains information on demand elasticity models. More information on the elasticity of transit use with respect fare may be found in a paper by Parody and Brand (1979), and also in another paper by Hamberger and Chatterjee (1987).

5.2 Stated Preference Surveys and Concept Tests for Forecasting

For transit planning it is sometimes necessary to ask people about their preference, and their likes and dislikes for various service characteristics. These surveys are used for determining how to improve an existing service and/or designing a new service, and also for forecasting ridership on a new service. These attitudinal and ‘stated preference’ surveys need sound statistical design for selecting the sample and analyzing the results. A discussion on stated preference survey may be found in the *Travel Survey Manual*. In the field of marketing, ‘concept tests’ are performed for estimating the potential demand for a new consumer product, and this approach can be extended to ridership forecasts for new/innovative transit services.

Hartgen and Keck (1976) describes a survey based method of forecasting ridership on a new dial-a-bus service. The interpretation of results of opinion-based surveys must be done carefully in order to account for any bias reflected in apparent results. A paper
by Chatterjee, McAdams and Wegmann (1983) presents a case study involving non-commitment bias in public opinion on the anticipated usage of a new transit service.

5.3 Forecasting Future Travel on Road Segments and/or Transit Lines

A variety of forecasting procedures are available ranging from the extrapolation of past trends to complex mathematical models involving several steps. A transportation planner must recognize the advantages and disadvantages of each procedure. Two procedures are examined for illustration.

5.3.1 Direct Estimation of Traffic Volume by Trend Analysis

If traffic volume data are available for a road segment or a transit line of interest for several years in the past, the historical trend can be identified and extrapolated to estimate future volumes. This approach, of course, is appropriate if the trend is expected to continue, which commonly is true for short-range forecasts. Trend based forecasts are appropriate also for aggregate values such as total VMT or transit rides in an urban area. However, major changes in the land development pattern and/or transportation network can cause substantial changes in the travel pattern, and if such changes are likely then trend extrapolation would not be appropriate. Therefore, for long-range forecasts of traffic volumes on individual segments of a road network or the number of passenger trips on individual transit routes, trend analysis is not used.

5.3.2 Stepwise/Sequential Procedure

A widely used travel estimation procedure for long-range forecasts of traffic volumes on a highway network uses several steps in a sequence as shown in the flow chart of Figure 3.
Each step requires a particular type of model or procedure, and there are different choices of models at each step. One of the major advantages of this procedure is its ability to reflect several types of changes, which may occur in the future:

a. Changes in trip making rates;

b. Changes in development pattern resulting in altered travel pattern;

c. Changes in transportation mode usage; and

d. Changes in transportation network.

Another advantage of the stepwise, or sequential, procedure is that it generates several types of useful information at the end of various steps. The disadvantage of the procedure is that it needs a large amount of data for model development. It also requires a sound knowledge of one of the available computer software that is specially designed for developing and applying these models. A great deal of research has been performed and is still being continued to improve this stepwise modeling procedure. It should be acknowledged that the staff of the former Bureau of Public Works and later the staff of Federal Highway Administration and Urban Mass transportation Administration made tremendous contribution to the development of various procedures and also computer software. An historical overview of the development of planning and modeling procedures in the United States is presented in a report by Weiner (1992).

The stepwise procedure is popularly known as the four-step modeling process as this procedure includes four major steps: trip generation, trip distribution, mode choice, and traffic assignment. Additionally network analysis must be done to develop a few types of
information that are needed for the other steps. These steps and procedures involved with each are discussed in detail in the following sections.

6 TRIP GENERATION

Trip generation is the first step of the four-step modeling procedure. It is a very important step since it sets up not only the framework for the following tasks but also some of the controlling values such as the total number of trips generated in the study area by location and trip purpose. The commonly used units for trip generation analysis usually include a household, a ‘dwelling unit’ (DU), and a business establishment. However, the results of a trip generation analysis for a study area are aggregated based on larger areas known as ‘traffic zones’.

A typical classification scheme of trips used for trip generation analysis is presented in Figure 2. A detailed discussion of this classification scheme is presented in a paper by Chatterjee, Martinson, and Sinha (1977). A thorough analysis of all these types of trips shown in the figure requires a large amount of data. These data are collected by using a variety of survey techniques, commonly referred to as origin-destination (O-D) surveys, which are discussed briefly in an earlier section. The discussion of this chapter will focus primarily on trip generation models for internal passenger trips made by households.

6.1 Models for Internal Passenger Trips: Aggregate and Disaggregate Models

The goal of trip generation models for internal passenger trips is to estimate total number of ‘trip ends’ for each purpose generated in each traffic zone based on socioeconomic and/or land use data for the respective zones. This task can be accomplished with either aggregate or disaggregate models. For aggregate models the total number of trips (trip ends) generated in a zone is used as the dependent variable, whereas for disaggregate models trips made by a
household (or a business establishment) is used as the dependent variable. When using disaggregate models the trip ends generated by households, and/or any other trip generating units such as business establishments, in a zone are combined to produce the zonal (total) value. Both disaggregate and aggregate trip generation models are used in planning studies.

6.2 Trip Generation by Households

Household generated trips comprise a major portion of all trips in an urban area. Actually more than 80% of trips in an urban area are generated by the residents of households in the area. Trips by non-residents and a variety of other vehicles including commercial vehicles such as taxis and trucks, and public utility and public service vehicles comprise the remaining portion of total travel. For the purpose of modeling, the trips generated by households are classified as home-based and non-home-based. Home-based trips have one end, either origin or destination, located at the home zone of the trip maker. If both ends of a trip are located in zones where the trip maker does not live, it is considered as a non-home-based trip.

6.2.1 Definitions of Productions and Attractions and Trip Purpose

Because of the predominance of home-based trips in an urban area, the model development is simplified if it is assumed that the home end of a trip is a ‘production’ (P) in the zone of the trip maker’s residence irrespective of whether it represents the origin or destination of the trip. According to this approach, the non-home end of a home-based trip is considered to be ‘attraction’ (A). For a non-home based trip, which, have neither its origin nor its destination at the trip maker’s residence, ‘production’ and ‘attraction’ are synonymous with ‘origin’ and
‘destination’ respectively. This definition of productions (P’s) and attractions (A’s) is depicted in Figure 4. It should be noted that for home-based trips the activity at the non-home end determines the trip purpose, and that non-home-based trips usually are not further stratified by purpose.

6.3 Cross-Classification or Category Models for Household Trip Generation

Household trip rates have been found to vary significantly according to certain socio-economic characteristics and the size of a household. Household Characteristics that have been found to be significant for trip generation and are commonly used in trip generation models are:

1. Household size

2. Auto ownership

3. Income

A hypothetical example of a trip generation model for households is presented in Table 1, which includes trip production rates for different types or categories of households defined in terms of different combinations of household size and auto ownership. This type of models is referred to as ‘cross-classification’ or ‘category’ models, and these are used widely for estimating trip productions by households.
Table 1.
A Hypothetical Category Model for Total Person Trips per Household

<table>
<thead>
<tr>
<th>Household size (persons per household)</th>
<th>Car Ownership</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No car</td>
</tr>
<tr>
<td>1 to 2</td>
<td>3.2</td>
</tr>
<tr>
<td>3 to 4</td>
<td>5.2</td>
</tr>
<tr>
<td>5 or more</td>
<td>7.2</td>
</tr>
</tbody>
</table>

Alternative techniques for a statistical analysis of these models are discussed in a paper by Chatterjee and Khasnabis (1973). When the households in a traffic zone are cross-classified by size and auto ownership, total trips made by households in a zone for a specific purpose \( P \) = Summation of (No. of households of a category) x (Trip rate for households of that category and for that specific purpose). In terms of mathematical notations this relationship is shown below:

\[
P = \sum_{kl=1}^{n} (HH_{kl})(TR_{kl}) \quad \ldots \quad (I)
\]

where, HH and TR stand for households and trip rates respectively, and \( kl \) represents a particular combination of household size \( k \) and auto ownership \( l \), and \( n \) is the number of combinations or categories.

The choice of household characteristics to be used for developing the various categories for
trip production rates may vary for one study to another. One advantage of disaggregate models is that for developing these models a full-scale O-D survey is not needed. A carefully selected small sample of households may be used for developing trip production rates as long as the number of cases for each category, or cell of the matrix, is statistically adequate.

### 6.4 Models for Trip Attractions

It is a common practice to use aggregate models in the form of regression equations for trip attractions. The dependent variable for these aggregate models is the total number of trip attractions for a specific trip purpose in a traffic zone. The independent variables usually are employment related and they represent zonal total values. Hypothetical examples of trip attraction models are presented below:

\[
\begin{align*}
(HBW A)_j &= 1.5 \text{ (Total Employment)}_j \\
(HBNW A)_j &= 8.5 \text{ (Retail Employment)}_j + 1.0 \text{ (Non-Retail Employment)}_j + 0.9 \text{ (Dwelling Units)}_j \\
(NHB A)_j &= 3.0 \text{ (Retail Employment)}_j + 2.0 \text{ (Non-Retail Employment)}_j + 0.8 \text{ (Dwelling Units)}_j
\end{align*}
\]

where,

\[
\begin{align*}
(HBW A)_j &= \text{ Home-Based Work Attractions in zone } j, \\
(HBNW A)_j &= \text{ Home-Based Non-work Attractions in zone } j, \\
(NHB A)_j &= \text{ Non-Home-Based Attractions in zone } j.
\end{align*}
\]

The development of aggregate models usually requires a full-scale O-D survey. The coefficients of the regression equations would vary from area to area. The choice of
independent variables and trip purpose categories also may vary from one study to another.

### 6.5 Balancing of Productions and Attractions

Due to the definition of productions and attractions, home-based productions in a zone may not be equal to the corresponding attractions in the same zone. Non-home-based productions in a zone should equal to corresponding attractions in the same zone. However, areawide (total) productions of any trip purpose -- home-based or non-home-based -- should be equal to the corresponding areawide (total) attractions. Thus,

\[
\sum (\text{HBW } P_i) = \sum (\text{HBW } A_j) \quad \text{(2)}
\]

\[
\sum (\text{HBNW } P_i) = \sum (\text{HBNW } A_j) \quad \text{(3)}
\]

\[
\sum (\text{NHB } P_i) = \sum (\text{NHB } P_j) \quad \text{(4)}
\]

When synthetic or borrowed models are used, the estimated areawide (total) productions would not be equal to the estimated areawide (total) attractions. Therefore, to achieve a balance zonal attractions are adjusted proportionately such that the adjusted areawide attractions equal areawide productions. Adjustment or scaling factors for Attractions are calculated as follows:

\[
\text{Adjustment Factor for HBW } A_j's = \frac{\sum (\text{HBW } P_i)}{\sum (\text{HBW } A_j)} \quad \text{(5)}
\]

\[
\text{Adjustment Factor for NBNW } A_j's = \frac{\sum (\text{HBNW } P_i)}{\sum (\text{HBNW } A_j)} \quad \text{(6)}
\]
6.6 Commercial Vehicle Traffic in an Urban Area

It should be pointed out that although internal trips made by residents on passenger vehicles account for a large proportion of total trips in an urban area, the other categories of trips must not be overlooked. The classification scheme presented in Figure 2 shows the other categories. The proportion of each category of trips varies according to the size and other characteristics of an urban area. For example, the proportion of through trips usually is larger in smaller size areas. In some cases, trips of one or more of these other categories may be the cause of major problems and thus would require special attention. For example, through traffic may be the major issue in the case of a small or medium sized urban area, and the planners may have to analyze these trips thoroughly. Similarly, the movement of large trucks may be of major interest in some urban areas. A comprehensive study should pay attention to travel demand of all categories although the level of detail may vary.

The analysis of commercial vehicle travel has been neglected in most urban transportation studies. These vehicles are garaged in non-residential locations and include trucks of all sizes, taxicabs, rental cars, service vehicles of plumbers and electricians, etc. There are a few useful references on how to estimate truck traffic in urban areas, and these include an article by Chatterjee, et al. (1979), a book by Ogden (1992), and a NCHRP Synthesis report by Fischer (2001).
When developing trip generation models the availability of data for the independent variables of the models is an important issue that can influence the selection of a variable. Usually the availability of data for the base year is less problematic than that for future years. Of course, if data for an independent variable is not available for the base year, it cannot be used in model development. However, what the model developer must recognize before building and adopting a model is whether the independent variables used in the model can be forecast by the responsible planning agency, and if such forecasts would be very difficult then it may be desirable to avoid using those variables in the model. Sometimes transportation planners have to develop a procedure or model to be used for making such forecasts.

Usually aggregate values of socioeconomic parameters used in trip generation models are not very difficult to forecast with the existing state of the art. The difficulty usually involves the task of disaggregating socioeconomic data at the zonal level. For example, it may not be very difficult to predict the total number of households in each zone along with their average size and auto ownership. However, it would be difficult to cross-classify the predicted number of households in a zone according to specific categories based on household size and auto ownership. Similarly predicting the average income of households in individual traffic zones may not be very difficult, but to develop a breakdown of the households in every zone by income groups would be difficult. The disaggregate trip generation models thus present a challenge to planners for making detailed forecasts of socioeconomic characteristics for future years. In order to provide assistance for making forecasts in a disaggregate form, a few procedures have been developed, and examples of such household stratification models can be found in a paper by Chatterjee, Khasnabis and
Slade (1977), and also in a report titled *Trip Generation Analysis* prepared by the staff of Federal Highway Administration (1975).

7 TRIP DISTRIBUTION

The purpose of the trip distribution step of the stepwise travel modeling procedure is to estimate ‘zone to zone’ movements, i.e., trip interchanges. This step usually follows trip generation analysis. In some cases trip distribution may come after trip generation and modal split analysis but that is not commonly found. The inputs to a trip distribution model are the zonal productions \( P_i \) and attractions \( A_j \). The model strives to link the productions and attractions based on certain hypotheses/concepts.

When trip distribution phase precedes modal split analysis, productions and attractions include trips by all modes and the distribution model should be multimodal in nature. In actual practice, however, multimodal trip distribution models are seldom, and in most cases highway oriented models have been used to distribute trips of all modes. It should be noted that in the rare case where trip distribution phase follows modal split analysis, mode–specific distribution models are needed. It is generally believed that ideally trip distribution should be combined with modal split analysis because decisions related to destination and travel mode usually are made simultaneously.

In this section widely used trip distribution technique – gravity model – will be discussed in detail followed by a brief overview of other types of models. A good review of commonly used trip distribution models can be found in an article by Easa (1993).
7.1 Formulation of a Gravity Model

The basic hypothesis underlying a gravity model is that the probability that a trip of a particular purpose \( k \) produced at zone \( i \) will be attracted to zone \( j \), is proportional to the attractiveness or ‘pull’ of zone \( j \), which depends on two factors. One factor is the magnitude of activities related to the trip purpose \( k \) in zone \( j \), and the other is the spatial separation of the zones \( i \) and \( j \). The magnitude of activities related to trip purpose \( k \) in a zone \( j \) can be expressed by the number of zonal trip attractions of the same purpose, and the effect of spatial separation between zones \( i \) and \( j \) can be expressed by a friction factor, \( F_{ij}^k \), which is inversely proportional to an appropriate measure of impedance, usually travel time. The attractiveness or pull of zone \( j \) with respect to zone \( i \) is proportional to \( A_j^k F_{ij}^k \). The magnitude of trips of purpose \( k \) produced in zone \( i \) and attracted to zone \( j \), \( T_{ij}^k \), of course, also depends on the number of trips being produced at zone \( i \), \( P_i^k \). This can be expressed mathematically as follows:

\[
T_{ij}^k = f(P_i^k, A_j^k, F_{ij}^k).
\]

The above formulation is not sufficient for estimating the \( T_{ij}^k \) values because it yet does not reflect any considerations for other zones that are competing as alternative destinations for the trips \( P_i^k \). Actually, the effective attractiveness of a zone is relative to others and it can be expressed as the ratio of its own attractiveness with respect to the total. Thus the relative attractiveness of a zone for trips of purpose \( k \) being produced in zone \( i \) is expressed by the ratio \( A_j^k F_{ij}^k / \sum_{j=i}^n A_j^k F_{ij}^k \). Dropping the subscript \( k \), the trip distribution model can be written as follows:
\[ T_{ij} = \frac{P_i A_j F_{ij}}{\sum_{j=1}^{n} A_j F_{ij}} \quad \text{…… (9)} \]

7.2 Application of Gravity Model Concept

The application of the gravity model concept for the trip distribution step of the stepwise travel forecasting procedure was introduced by Alan M. Voorhees (1955). The classic example of an application of a gravity model as presented by Voorhees is shown in Figure 5. This application is based on the assumption that the effect of spatial separation with respect to trip making is proportional to the inverse of the square of travel time between the respective pairs of zones. The calculations presented below shows how the accessibility and attractiveness of Zone 4 changed due to a new expressway resulting in an increased number of trip attractions. The number of shopping trips attracted to Zone 4 from Zone 1 was 28 without an expressway, and it increased to 80 as a result of a new expressway. The increase in trips attracted to Zone 4 resulted in a decrease of trips attracted to the other zones.

1. Situation Without Expressway

<table>
<thead>
<tr>
<th>Existing ‘Pulls’ From Zone 1</th>
<th>% Pull</th>
<th>No. of Trips From Zone 1 to Zone 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zone 1 = ( \frac{0}{1^2} = 0 )</td>
<td>0/70 = 0 %</td>
<td>200 x 0 = 0</td>
</tr>
</tbody>
</table>
Zone 2 = \( \frac{1000}{5^2} = 40 \)  
\( \frac{40}{70} = 57 \% \)  
\( 200 \times .57 = 114 \)

Zone 3 = \( \frac{2000}{10^2} = 20 \)  
\( \frac{20}{70} = 29 \% \)  
\( 200 \times .29 = 58 \)

Zone 4 = \( \frac{4000}{20^2} = 10 \)  
\( \frac{10}{70} = 14 \% \)  
\( 200 \times .14 = 28 \)

Total ‘Pull’ = 100  
100  
200

II. Situation With Expressway

Existing ‘Pulls’ From Zone 1

<table>
<thead>
<tr>
<th>No. of Trips From Zone 1</th>
<th>% Pull</th>
<th>No. of Trips From Zone 1 to</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zone 1 = ( \frac{0}{1^2} = 0 )</td>
<td>0/100 = 0 %</td>
<td>200 x 0 = 0</td>
</tr>
<tr>
<td>Zone 2 = ( \frac{1000}{5^2} = 40 )</td>
<td>40/100 = 40 %</td>
<td>200 x .40 = 80</td>
</tr>
<tr>
<td>Zone 3 = ( \frac{2000}{10^2} = 20 )</td>
<td>20/100 = 20 %</td>
<td>200 x .20 = 40</td>
</tr>
<tr>
<td>Zone 4 = ( \frac{4000}{20^2} = 40 )</td>
<td>40/100 = 40 %</td>
<td>200 x .40 = 80</td>
</tr>
</tbody>
</table>

Total ‘Pull’ = 100  
100  
200
7.2.1 Calibration of Gravity Model Parameters

The three basic parameters of a gravity model are zonal trip productions and attractions, $P_i$’s and $A_j$’s and friction factors $F_{ij}$’s. Whereas the $P_i$’s and $A_j$’s are estimated by trip generation models, friction factors must be derived as a part of the trip distribution phase. The basic concept of gravity model implies the following form for friction factors,

$$F_{ij} = \frac{1}{(\text{travel time between } i \text{ and } j)^x} \quad \ldots \ldots (10)$$

In the early application of the gravity concept, the exponent of the travel time was assumed to be a constant. However, further empirical analysis suggested that the exponent varies over the range of values for travel time. The actual values of friction factors are derived by a trial and error procedure and these vary according to trip purpose.

For each trip purpose an arbitrary set of friction factors for a range of travel time values at an increment of one minute is assumed at the beginning of the calibration process and the results of this initial application of the model is evaluated with respect to the actual trip distribution obtained from the origin-destination (O-D) surveys. The evaluation is made by comparing the trip length frequencies generated by the model with those derived from the O-D survey, and if the results are not similar, the friction factors are adjusted and the gravity model is applied again with the new factors. This trial and error procedure is continued until the trip length frequencies of a model appear similar to those of the O-D survey. It may be noted that the absolute values of these factors have no special implications and that it is the relative weight with respect to each other that is important. The respective set of these frictions factors can be scaled up or down by a constant factor.
7.2.2 Balancing a Gravity Model

It must be pointed out that due to the basic nature of the gravity model formulation, the zonal productions obtained from the model application must equal the values of $P_i$’s originally used as inputs to the model. However, the same is not true for the zonal attractions and the model results must be compared with the original $A_j$’s. In the cases when model generated $A_j$’s do not match closely with original $A_j$’s, a balancing procedure is used by adjusting the input values of $A_j$’s until the model results are satisfactory.

8 OTHER TYPES OF TRIP DISTRIBUTION MODELS

The gravity model is by far the most widely used trip distribution technique. There are also other techniques that are used in urban transportation planning. One type of these alternative techniques utilizes growth factors for each traffic zone and uses an iterative balancing procedure to project a base year trip matrix to the future year. The most popular of the growth factor techniques is that introduced by Fratar (1955) and it is known as the Fratar technique. The limitation of a growth factor procedure is that they basically are extrapolation techniques and cannot be used to synthesize movements between zone pairs if the base year trips are zero. However, the Fratar technique is utilized regularly for projecting through trips in an urban area and sometimes even for external - internal trips.

A somewhat complex trip distribution technique that was used by the Chicago Area Transportation Study, the transportation planning agency for Chicago, is the Intervening Opportunities Model. The trip distribution theory underlying this model states that the probability that a trip originating in a zone $i$ will find a destination in another zone $j$, is
proportional to possible destinations in zone \( j \) and also the possible destinations in zones closer to the origin of the trip. This model is rarely used by any agency today.

9 MODAL SPLIT

One of the widely researched step/phase of the sequential travel modeling procedure for urban transportation planning is the modal split analysis which involves the allocation of total person trips (by all modes) to the respective modes of travel, primarily automobile and public transit. It should be noted, however, that many studies for small and medium sized urban areas omit this step by developing and using models for automobile trips only. This omission is justified in areas where transit trips constitute a very small fraction of total trips and are made primarily by captive riders.

Modal split models basically relate the probability of transit usage to explanatory variables or factors in a mathematical form. The empirical data necessary to develop these models usually are obtained from comprehensive O-D surveys in specific urban areas. In applying these models to predict the future transit usage, one must make the implicit assumption that the variables which explain the present level of transit usage will do so in much the same manner in the future.

9.1 Factors Affecting Mode Choice

Factors that may explain a trip maker’s choosing a specific mode of transportation for a trip are grouped commonly as follows:
9.1.1 Trip Makers Characteristics:

a. Income

b. Car-Ownership

c. Car Availability

d. Age

9.1.2 Trip Characteristics:

a. Trip Purpose - work, shop, recreation, etc.

b. Destination Orientation - CBD vs. non-CBD

c. Trip Length

9.1.3 Transportation Systems Characteristics

a. Waiting time

b. Speed

c. Cost

d. Comfort and Convenience

e. Access to terminal or transfer location
9.2 Categories of Modal Split Models

The possible sequence of different types of modal split models with respect to the other steps of travel modeling procedure is shown in Figure 3.

9.2.1 Pre-Distribution (or Trip End) Models:

This type of a modal split model is used to separate the trip productions in each zone into the different modes to be distributed by mode-specific trip distribution models. The primary disadvantage of these models is that they cannot include variables related to transportation system characteristics. Pre-distribution models are not used commonly.

9.2.2 Post-Distribution (or Trip Interchange) Models:

This type of modal split models is very popular as it can include variables of all types. However, conceptually it requires the use of a multi-modal trip distribution model and currently such distribution models are not used commonly. Figure 6 illustrates the sequence of application of a post-distribution model.

9.2.3 Simultaneous Trip Distribution and Modal Split Models:

This type of a model strives to estimate the number of trips between two zones by specific modes in one step directly following the trip generation phase. Conceptually and theoretically this type of a model has a sound basis, but it is not commonly used at this time.

9.3 Developing a Modal Split Model

Modal split models are developed from observed data on trip making available from home-interview surveys. The analysis involves the processing of a variety of data for both demand
9.3.1 Aggregate Model

Modal split models of 1960’s and early 1970’s in most cases were based on an ‘aggregate’ approach, which examined the mode choice of trip makers and their trips in groups based on similar socioeconomic and/or trip characteristics. These mode choice models usually involved two modes only - auto and transit. A detailed stratification scheme was used, and the share of each mode was determined for each stratified group of trips, which then was correlated with selected independent variables. The dependent variable was ‘percent transit’ applicable to a group of trips of similar characteristics made by similar trip makers. Commonly used independent variables include: the ratio of travel time by transit to that by automobile; the ratio of travel cost by transit to that by automobile; and the ratio of accessibility by transit to that by automobile. The relationship of the dependent variable, percent transit, with the independent variable, say ratio of travel times, commonly was expressed by a set of curves. These curves sometimes were referred to as modal diversion curves.

The development of ‘aggregate’ modal split models requires a large amount of data. Discussion of procedures used for developing different types of ‘aggregate’ modal split models along with examples of these models can be found in papers by Weiner (1969), and Chatterjee and Sinha (1975).

9.3.2 Disaggregate Behavioral Logit Models

During late 1970’s a new approach known as disaggregate behavioral method was developed and refined by a number of researchers. This approach recognized each individual’s choice of
mode for each trip instead of combining the trips in homogeneous groups. The underlying premise of this modeling approach is that an individual trip maker’s choice of a mode of travel is based on the principle called ‘utility maximization’. Another premise is that the utility of using one mode of travel for a trip can be estimated using a mathematical function referred to as the ‘utility function’, which generates a numerical utility value/score based on several attributes of the mode (for the trip) as well as the characteristics of the trip maker. Examples of a mode’s attributes for a trip include travel time and costs. The utilities of alternative modes also can be calculated in a similar manner. A trip maker chooses the mode from all alternatives that has the highest utility value for him/her.

A mathematical function that was used to represent the correlation of the probability of a trip maker’s choosing a specific mode for a specific trip with a set of utility values is known as the ‘logit’ function. Therefore, these models are also referred to as ‘logit’ models. Binomial logit models deal with two modes, whereas multinomial logit models can deal with more than two modes. An example of the mathematical formulation of a multinomial logit model is given below:

\[ p(k) = \frac{e^{U_k}}{\sum_{x=1}^{n} e^{U_x}} \quad \ldots \ldots \ (11) \]

\[ p(k) \quad = \quad \text{probability of using mode } k \]
\[ U_k \quad = \quad \text{utility of using mode } k \]
\[ U_x \quad = \quad \text{utility of using any particular mode } x \]
\[ n \quad = \quad \text{number of modes to choose from} \]
A special statistical procedure known as the maximum likelihood technique is used to derive an equation that combined different variable/factors in a meaningful way to calculate a utility (or disutility) value. The coefficients of each variable included in the utility (or disutility) function reflect certain behavioral aspects of a trip maker. Usually transportation related variables used for a utility function include such items as access (or, egress) time to (or, from) transit stops/stations, wait time, line-haul time, and out-of pocket costs, and the coefficients of these variables are negative. Thus the combined utility value comes out to be negative, which indicates ‘disutility’ of using a mode. A trip maker’s characteristics such as ‘income’ are also built into the utility function.

One of the advantages of disaggregate mode choice models is that it does not need a full-scale O-D survey with household samples from every traffic zone. A carefully selected sample of 1,500 to 2,000 households would be adequate for developing these models. The mathematical theory related to multinomial logit models for mode choice analysis is fairly complex and beyond the scope of this chapter. Numerous articles and reports have been published on the subject of behavioral logit models, and these include an article by Reichman and Stopher, and another by McFadden. A report by Horowitz, Koppelman and Lerman also has detailed information about disaggregate mode choice modeling.

10 TRAFFIC ASSIGNMENT

The task of the traffic assignment process is to develop the loadings, or user volumes, on each segment of a transportation network as well as the turning movements at intersections of the network. The user volumes may be the number of vehicles, the number of total persons, the number of transit riders, or any other units of travel demand that can be described by an origin
and destination. For highway networks, user volumes are in terms of the number of vehicles whereas for transit assignment, the numbers of riders/passengers represent volumes. The relationship of the traffic assignment phase with respect to the other phases of the sequential travel simulation procedure is shown in Figure 3.

10.1 Inputs to Traffic Assignment Process

The two basic inputs to the assignment process are the transportation network and the zone to zone trip interchanges. The transportation network of automobiles, trucks and taxis are analyzed separately from that of public transit systems, and usually traffic assignments are made separately for highway and transit systems. The typical inputs of a highway traffic assignment are shown in Figure 7. Transit network assignments are limited to internal person trips only.

10.2 Highway Network Analysis

For the purpose of computer analysis a highway network is represented by links and nodes and the traffic zones are represented by centroids, which are connected to the network. The characteristics of each link, such as the distance, speed, capacity, turn prohibitions and functional classification are coded. One of the primary tasks of network analysis is to determine the minimum time routes between each pair of centroids and this task is performed utilizing the Moore’s algorithm. This algorithm does not require all possible routes between an origin and destination to be individually investigated to find the shortest route. Rather, a minimum “tree” is developed by fanning out from the origin to all other accessible nodes in an increasing order of their impedance summation from the origin.
A ‘tree’ is defined as the set of shortest routes from an origin to all other points in a network. Example of a path tree is shown in Figure 8. The travel time between a pair of zones is obtained by adding up the times on the individual links comprising the minimum time route, and this is repeated for every pair of zones. A ‘skim tree’ usually refers to the interzonal travel time matrix.

It should be pointed out that the coding of a network for analysis with a computer-based algorithm requires a great deal of care and experience. There are many detailed issues and questions that come up with reference to such items as centroid connectors, representation of interchange ramps, whether to include certain roads or not, etc. Coding errors also can cause problems, and there are certain checks that can be done to minimize errors. A paper by Easa (1991) and a report by Ismart (1990) discuss some of these issues and techniques for coding a network.

### 10.3 Alternative Techniques for Highway Traffic Assignment

A traffic assignment technique basically allocates the trips between each zone pair to the links comprising the most likely travel routes. The trips on each link are accumulated and the total trips on each link are reported at the end of the assignment process. Alternative assignment techniques vary in terms of the criteria for route selection.

#### 10.3.1 All or Nothing Assignment (AON):

This procedure assigns all trips between a zone pair to its minimum time route. This is the most commonly used technique, although the realism of its basic hypothesis is questionable.
It should be noted that other more advanced techniques make use of this technique as a part of their more involved procedure.

10.3.2 Diversion Techniques:

A diversion technique allocates the trips between a zone pair to more than one route. The most commonly used diversion technique considers two routes. One of these routes uses freeways and the other is the quickest alternative non-freeway arterial route. The procedure assumes that a proportion of trips as determined by a diversion curve will be diverted from an arterial route to a freeway route based on the ratio of ‘time via freeway’ with respect to ‘time via quickest alternate arterial route’. This procedure is documented in *Traffic Assignment Manual (1964)*.

Diversion techniques were widely used in the early 1960’s. Their advantage is in getting a spread of traffic between competing routes, and these techniques appear to be more realistic than the all-or-nothing assignment. With the introduction of the capacity restraint assignment procedure, diversion techniques are rarely used today for network assignments, although its usefulness should not be overlooked for corridor type applications.

10.3.3 Capacity Restraint Assignment:

The capacity restraint procedure explicitly recognizes that as traffic flow increases the speed of traffic decreases. In this procedure several assignments are made based on the “all or nothing” concept. At the end of each assignment, however, the assigned volume on each link is compared with the respective capacity and the travel time is adjusted according to a given formula. A new set of minimum time routes is computed for the next assignment.
The original capacity restraint procedure developed by the then Bureau of Public Roads, which is documented in *Traffic Assignment Manual (1964)*, assumed that the relationship between travel time and the volume peculiar to each link in a highway network can be expressed by the following equation:

\[
T = T_0 \left[ 1 + 0.15 \left( \frac{V}{C_p} \right)^4 \right] \quad \text{.................. (12)}
\]

where:  
- \( T \) = Travel time at which assigned volume can travel on the subject link.  
- \( T_0 \) = Base travel time at zero volume = travel time at practical capacity x 0.87.  
- \( V \) = Assigned Volume  
- \( C_p \) = Practical capacity.

This process may be continued for as much iteration as desired. Usually four iterations are adequate. The analyst has the choice to accept the results of any single iteration. Sometimes the link volumes obtained from all iterations are averaged to produce the final result. This procedure strives to bring the assigned volume, the capacity of a facility, and the related speed into a proper balance.

### 10.3.4 Equilibrium Assignment:

Traffic assignment has been a subject of intense research for many years, and the research has resulted in several alternatives to the all-or-nothing and capacity restraint techniques, which were widely used during the 1960’s and 1970’s. One example of an assignment technique that was developed after the capacity restraint technique is the probabilistic multipath assignment
technique, which was developed by Robert Dial (1971). The most widely used procedure today is the ‘user equilibrium’ assignment. Equilibrium assignment technique is based on the notion that traffic flows on network links are adjusted to an equilibrium state by the route switching mechanism. That is, at equilibrium, the flows will be such that there is no incentive for route switching. As mentioned before, the travel time on each link changes with the flow and therefore, the travel time on several network paths changes as the link flows change. A stable condition is reached only when a traveler’s travel times cannot be improved by unilaterally changing routes. This network state characterizes the user equilibrium (UE) condition.

The UE condition strives to optimize the utility of individual drivers. If the analysis is focused on optimizing a system wide travel measure such as minimum aggregate travel time, then the problem is called a system optimal equilibrium (SOE) problem. Both UE and SOE problems rely on mathematical programming methods for developing the formulation and deriving a solution.

**UE Problem Statement:** Given a generalized function $S_a$ that relates arc/link costs to traffic volumes, find the equilibrium traffic volumes on each arc/link of a directed graph, $G(N,A)$, with $N$ nodes, $A$ arcs and a total number of origin-destination zones, $Z$. This user equilibrium traffic assignment problem may be formulated in the following non-linear optimization form.

$$\text{Minimize} \quad f(x) = \sum_a \int_0^{V_a} S_a(x) dx \quad \text{...... (13)}$$

$$\text{subject to:} \quad V_a = \sum_i \sum_j \sum_k \delta_{ij}^{ak} x_{ij}^k \quad \text{......(14)}$$
\[ \sum_{k} x_{ij}^k = T_{ij} \] \quad \text{......(15)}

\[ x_{ij}^k \geq 0 \] \quad \text{......(16)}

where,

- \( i \) subscript indicates an origin zone/node \( i \in Z \);
- \( j \) subscript indicates a destination zone/node \( j \in Z \);
- \( k \) indicates a path between the origin zone (root) \( i \) and the destination zone \( j \);
- \( a \) subscript for link/arc, \( a \in A \);
- \( u \) subscript for volume category, \( u \in U \);
- \( T_{ij} \) number of trips (all modes) originated at \( i \) and destined to \( j \);
- \( C_a \) capacity of arc \( a \); and
- \( V_a \) total volume in category \( u \) on arc \( a \) in current solution (total volume across all classes when no class is specified);
- \( W_a \) all-or-nothing volume of \( u \) trips on arc \( a \) in current solution;
- \( S_a(V_a) \) generalized travel time (cost) function (also known as the link performance function) on link \( a \) which is determined by total flow \( V_a \) on each link (independent of vehicle operating mode classes);
- \( x_{ij}^k \) number of total trips from \( i \) to \( j \) assigned to path \( k \);
- \( \delta_{ij}^a \) 1 if link \( a \) belongs to path \( k \) from \( i \) to \( j \), 0 otherwise;
Most common form of the link performance function used for equilibrium traffic assignment problems is shown in equation 17.

\[ T = T_0 \left[ 1 + a \left( \frac{V}{C_p} \right)^b \right] \] .......................... (17)

where \( a \) and \( b \) are constants (note that when \( a = 0.15 \) and \( b = 4 \), equation 17 reduces to the form of BPR function shown in equation 12.

Even for very small networks, it is very difficult to obtain a mathematical solution to this non-linear optimization problem. Frank-Wolfe (1956) developed a heuristic solution that decomposes the problem into a number of steps. The solution is popularly known as Frank-Wolfe decomposition of the user equilibrium problem. Presented in Figure 9 is a schematic of computer implementation of Frank-Wolfe decomposition.

After several iterations, the heuristic reaches a situation where UE condition is satisfied. As can be seen in Figure 9, the user equilibrium assignment technique utilizes a convex combination method called the Golden Section Search for direction finding. The mathematical concepts and optimization techniques underlying the user equilibrium assignment is fairly complex and beyond the scope of this chapter. More details about UE and other optimization solutions for network assignment problems can be found in Sheffi (1985).

11 MODEL ADJUSTMENTS AND USE OF SYNTHETIC/BORROWED MODELS

When the models at different steps of the sequential modeling process are developed based on detailed data collected in the study area using large samples as prescribed for O-D and other
survey procedures, the results at each step can be verified against survey generated data that represent true values of the dependent variables of the models, such as trip ends (productions and attractions), trip length frequencies, mode choice proportions, and traffic volumes on network links. The availability of detailed data allows model developers make adjustments at respective steps as needed. For example, if during trip generation analysis it is found that the trip ends at a certain zone cannot be estimated closely by trip production or attraction models, a special investigation may be done, and an off-model procedure may be used for that zone. A special generator analysis for a college campus or an airport is an example of this type of a case. For gravity models also a special adjustment factor has been used in some cases to reflect the impact of certain physical features such as a river crossing on the attractiveness (or ‘pull’) of a group of traffic zones. Traffic counts taken at selected screen-lines are useful for comparing model generated travel pattern with actual volumes of traffic crossing screen-lines and making adjustments, if needed.

At the traffic assignment step, network related parameters such as travel speed and capacity of certain links may need adjustments in order to produce results that match closely existing traffic counts. For this purpose traffic counts taken at cut-lines are utilized for comparison with volumes generated by traffic assignment model. For example, if it is found that along a freeway corridor the assigned volumes on freeway links are too high whereas those along a parallel arterial highway are too low in comparison to ground counts, the speeds and capacities of selected links along the corridor should be reexamined and adjusted. A good source of useful ideas for model adjustments including network coding is a report written by Dane Ismart (1990).

Making adjustments to models to replicate the existing situation more closely is
an important task of model development, and this requires some experience and sound understanding of how the models work. Adjustments to models assume a greater role in the case of synthetic or borrowed models. In this case a set of pre-existing models for one area is transferred and adopted for another study area. Then the original values of key parameters such as trip generation rates and friction factors are adjusted, if necessary, to produce desired results. This process is referred to as model calibration. In the case of transferred/borrowed models data from O-D survey are not available for checking the results at each step of the sequential modeling process, and the only data to check with is traffic volume counts on various road segments, which are to be replicated by the results of traffic assignment.

Synthetic models have been used widely especially for small and medium sized urban areas in order to avoid or reduce the cost and time required for full-scale O-D surveys. Transportation planners of North Carolina DOT’s Planning and Research Branch have developed and used synthetic models for urban areas of a variety of sizes in North Carolina for many years. Papers written by Chatterjee and Cribbins (1972), Bates (1974), Modlin (1974), Khasnabis and Poole (1975), and Chatterjee and Raja (1989) describe some these procedures. Two NCHRP reports authored by Sosslau, et al. (1978), and Martin and McGuckin (1998) respectively contain considerable information on the use of borrowed models and transferred parameters.

12 TRAVEL DEMAND MODELING SOFTWARE PACKAGES

The principles and the steps involved in the travel demand modeling process were first implemented in the form of computer software programs by the former Bureau of Public Roads and later refined by Federal Highway Administration. This software package originally
was called PLANPAC/BACKPAC, and it was primarily highway oriented. A report titled FHWA Computer Programs for Urban Transportation Planning (1974) contains the details of these programs. In early 1970’s the Urban Transportation Planning System (UTPS) by developed by the former Urban Mass Transportation Administration to add the capability for transit planning. Later the scope of UTPS was expanded to include both highway and transit networks. UTPS is an IBM mainframe computer based software system, which has individual modules that are capable of performing a specific task. For example, the UMATRIX module performed matrix computations and the ASSIGN module performed AON or capacity restrained assignment. As with the case of many computer programs in the 1960s and 1970s, the capabilities of early version of UTPS were very limited. Performing a complete run of the four-step process took several days of work related to input preparation, debugging and output analysis. Until early 1990s several MPOs were still using the UTPS based planning software.

With the advent and penetration of microcomputers in early 1980s, different commercial versions of travel demand modeling (TDM) software were developed and marketed. Among the first of these kind is a software package called MINUTP, which was developed and marketed by COMSIS Corporation. MINUTP was a MS-DOS based command driven modeling package and similar to FHWA’s PLANPAC software. Included among the command driven TDM packages that were popular till late 1990s and even in early 2000s are TRANPLAN and QRS-II.

Since the advent of Microsoft Windows operating system, the travel demand modeling software landscape has changed even more dramatically. Current high-end TDM packages are not only capable of performing travel demand modeling, but are also compatible with Geographic Information Systems (GIS). For example, TransCAD™ (Caliper
Corporation) is a travel demand modeling package as well as a GIS software package. Other TDM packages include CUBE, TP+ (Citilabs), EMME/2, T-Model (Strong Concepts) and Saturn (UK).

13 APPLICATIONS OF TRAVEL MODELS

Depicted in Figure 10 is the traditional long-range planning process for a region (MPO) or sub-region. This process involves the identification of transportation-related problems followed by the determination of the future travel demand for a given situation. This in turn is followed by an attempt to find future transportation improvement that would meet the need of the future travel demand. Traditionally in the planning process, the main criterion that is used objectively to evaluate alternative projects is congestion relief by capacity improvement, which typically involves building new highways, widening of existing highways, and improving transit services. Land use related alternatives also are examined. Travel forecasting models help assess the effectiveness of each alternative in reducing traffic congestion. Since traffic congestion of serious nature usually occurs on major highways – primarily arterials – the travel forecasting procedure usually pays more attention to these highways, and this was reflected in network coding. Typically local roads and some minor collectors are not included in the network used for traffic assignment.

For the design of a new highway and/or the widening of an existing highway, the estimated traffic volume for the facility is main item of interest that highway design engineers expect from travel forecasting models. For other related information such as the proportion of design hourly volume with respect to average daily traffic, directional split, and truck percentage, highway design engineers use other sources of information. However, this
situation changed when more and more urban areas had to assess the air quality impact of highway networks as more detailed and accurate information related to travel was needed for air quality analysis.

TRANSPORTATION AND AIR QUALITY PLANNING

In recent years, the planning process has given considerable emphasis on the assessment of the effect of transportation alternatives on the environmental consequences especially air quality impacts. The Environmental Protection Agency (EPA) developed several versions of a model over time for estimating emission factors of air pollutants from mobile sources in terms of grams per mile. One of the recent versions of this model – MOBILE5 – was used widely during 1990’s. Currently the latest version, MOBILE6, is being used. The state of California, a different emission factor model called EMFAC is used. These emission factor models need a variety of travel related measures for the estimation of emissions from vehicular travel, and this need uncovered several deficiencies of the traditional travel forecasting models and led to various refinements and advancement of the modeling procedure. The integration of travel models with emission factors models is illustrated in Figure 11.

The travel related inputs required for mobile source emissions estimation are discussed in a paper by Miller, et al. (1992), and the deficiencies of the four-step models are examined in a paper by Stopher (1993). Another source of information on this subject is the NCHRP Report 394 titled Improving Transportation Data for Mobile Source Emission Estimates authored by Chatterjee, et al. (1997). A few examples of these weaknesses and refinements are discussed below:
14.1 Travel Speed.

The amount of emissions released by vehicles when traveling varies considerably with speed, and, therefore, an accurate estimation of travel speed on each link of a highway network is important. However, the traffic assignment procedures usually focus on the accuracy of the assigned traffic volumes and travel speeds are adjusted to produce better results for traffic volumes. In many cases the travel speeds generated by travel forecasting models are not accurate and little effort was made in the past to improve the speed values because, as was mentioned earlier, for the tasks of capacity-deficiency analysis and design of highway improvements, predicted traffic volume is the item of interest and there was no urgent need for improving the accuracy of speed estimates. During 1990’s, however, in response to the need of emission factor models considerable research was performed to improve the model generated speed values. For this purpose feedback loops as shown in Figure 12 and post-processing procedures for speed calculations based on volume/capacity ratios were introduced, and these are widely used at this time. There are alternative methods for feedback and iteration that can be used, and also there are different convergence criteria to choose from. A paper by Boyce, Zhang and Lupa (1994) contains a discussion on alternative methods of introducing ‘feedback’ into the four-step procedure, and a paper by Dowling and Skabardonis (1992) provides good information on speed post-processing.

14.2 VMT.

For the calculation of total emissions from mobile sources in an urban area, a reliable estimate of total VMT for each class of roads is needed. The estimation of VMT generated on local and minor collector roads is important for this purpose. However, as discussed earlier, most
of the local and minor roads usually are not included on the highway network used for travel demand modeling. Thus network-based models are not capable of generating reliable estimates of VMT on local and minor collector roads. To fill this gap a few off-model procedures have been developed in recent years, and some of these are discussed in NCHRP Report 394 authored by Chatterjee, et al. (1997).

14.3 Vehicle Class Mix and Vehicle Age Distribution.

For traditional transportation planning, which determines future needs of new and improved highways and transit services, detailed information regarding different types of vehicles operating on various highways is needed. Usually it is sufficient to use a rough estimate of the proportion of large trucks for capacity analysis. For air quality analysis, however, MOBILE5 provided separate emission rates for eight different types of vehicles. The MOBILE6 model requires input parameters for up to 28 different types of vehicles. The proportions of these classes of vehicles actually vary according to the functional class of roads and also by time of day. The data need becomes more demanding and complicated, when it is recognized that emission rates depend not only on the size and weight of a vehicle but also on its age and mileage accumulation, and that this information too has to be developed to take advantage of the MOBILE model’s capabilities.

It is clear that commonly used travel forecasting models are not designed to develop detailed information on vehicle class and age, and that other off-model procedures must be used for this purpose. However, the responsibility for developing this information lies with transportation planners who are working in the area of travel demand analysis.
14.4 Start Versus Running Emissions

Earlier versions of MOBILE (version 5 and earlier) combined start emissions and running emissions into one composite emission factor. Owing to the realization that start emissions are much more pronounced than the running emission, the latest version of MOBILE (version 6), approached emissions modeling in such a way that start emissions are separated from running emissions. Inputs on starts are provided in the form of start distributions by vehicle class.

14.5 Cold and Hot Soak Periods and Operating Modes

MOBILE 6 accounts for engine temperatures in the form of soak distributions (cold and hot soak distributions). The time lag between two successive trips of a vehicle determines the emissions at the beginning (start emissions) and end (evaporative emissions) of a trip. Inputs on this time lag are provided to the emission models in the form of soak period distributions. Venigalla and Pickrell (2002) described procedures to obtain soak period inputs from large travel survey databases.

The operating modes of vehicles are related to engine temperature. Earlier versions of the MOBILE model classified operating modes into two broad categories – transient and hot stabilized modes. The transient mode is further categorized into two separate sub-categories – cold start and hot start modes. EPA uses a few criteria based on engine soak period to determine the operating mode of a vehicle at a particular time. A few researchers used these time-based criteria in conjunction with travel surveys and innovative traffic assignment procedures to demonstrate how operating mode fractions can be estimated analytically. Papers by Venigalla, et al. (1995a, 1995b, 1999), and Chatterjee, et al. (1996), explain these procedures.
More and more the regulatory burdens of air quality modeling related to transportation projects is shouldered by transportation planners. Air quality models are constantly being updated with new knowledge gained on transportation related emissions. The current state of the art in emissions modeling requires more from the transportation planning community than ever before. For example, the concept of trip ends and trip chaining are easily extended to deriving travel related inputs to the emission factor models (Chalumuri, 2003). While model improvement efforts undoubtedly improve the state of the practice, additional burdens are placed on transportation modeling community to develop innovative methods to derive travel related inputs to emissions model. In order to accommodate the needs of the transportation related air quality modeling in the foreseeable future, the transportation planners are expected to develop new methods or adapt existing methods.

**List of References**


Chalumuri, S. *Emission Impacts of Personal Travel Variables*. A thesis submitted in partial fulfillment of Master of Science degree in Civil and Infrastructure Engineering Department, George Mason University, Fairfax, VA 2003.


Figure 1
The Complexity of Trip Making at a Typical Suburban Household
Alternative format

Figure 2
Classification of Travel within a Metropolitan Area
Figure 3
Sequential Travel Estimation Procedure
Figure 4
Definitions of Productions and Attractions

Zone | Production | Attraction (Purpose) |
-----|------------|---------------------|
1    | 4 (2 work, 1 personal business, 1 shop) | 0 |
2    | 1 (Nonhome-Based) | 1 (Personal business) |
3    | 0 | 2 (1 Nonhome-Based, 1 shop) |
4    | 0 | 2 (work) |
Total | 5 | 5 |
Figure 5
Example of Gravity Model Concept as Introduced by Alan M. Voorhees
Travel Simulation Procedures with Post-Distribution Modal Split Models
Figure 7
Inputs and Outputs of Traffic Assignment
Figure 8
An Example of a Minimum Path Tree from Home Zone 4
Start

Insert Initial Solution, or Set \( V_a = 0 \) for all \( a \in A \)

Update Travel Times Based on Link Flows, \( V_a \)

Set Origin Zone, \( i = 1 \)
Set \( W_a = 0 \) for all \( a \in A \)

Build Min Path Trees for \( i \)

Set Destination Zone, \( j = 1 \)

Assign AON Volumes
\[
W_a = W_a + T_{ij}
\]

Set \( j = j + 1 \)

Is \( j > Z \)

Set \( i = i + 1 \)

Is \( i > Z \)

Determine the Equilibrium Combination Factor, \( \lambda \)

Obtain New Current Solution
\[
V_a = (1-\lambda) V_a + \lambda W_a
\]

Solution Converged?

Stop

Network Database
Trip Exchange Matrix \((T_{ij} \text{ where } i, j \in Z)\)

Golden Section Search

Figure 9
An Example of a Minimum Path Tree from Home Zone 4
Figure 10
The Long Range Planning Process
Figure 11
A Schematic Representation of TDM Integrated with Air Quality Modeling
The Planning Process with Feedback Loops

Figure 12

The Planning Process with Feedback Loops