21
DEMAND ESTIMATION
FOR PUBLIC TRANSPORT
NETWORK PLANNING

Wenzhe Sun and Jan-Dirk Schmöcker

Introduction

The classic public transport (PT) network planning process contains five main steps: (1) network route design, (2) frequency setting, (3) timetable development, (4) vehicle scheduling and (5) crew scheduling (Ceder & Wilson, 1986). Clearly, these steps are interdependent, and “loops” may be needed in the planning process. In particular, frequency setting and timetable development can often be integrated. Table 21.1 describes the framework of Ceder (2007) but in a simplified form. In Ceder (2007), multidirected interactions between the levels are illustrated and discussed. In this chapter, the aim is to emphasise that demand estimates are crucial for the first, and most influential, stages of the planning process. The latter two stages, vehicle and crew scheduling, mainly determine a balance between the demand-driven timetable and limited vehicle and human resources.

The different ways to obtain demand estimates are the main focus of this chapter. Classic tools such as surveys are still important, but the availability of massive passive PT data sources such as automatic passenger counter (APC) and automatic fare collection (AFC) data has triggered notable advances in deriving reliable demand estimates (Kurauchi & Schmöcker, 2016). As a consequence of this, however, the gap in the service quality provided by the “data-rich” and the “data-poor” practitioner is enlarged due to fundamental differences in the planning process. A “big data”-driven planning process can accommodate fluctuating user demand by efficiently making use of limited land, capital and human resources. It is necessary to be aware of this gap and to understand the potential in the datasets that are more common and cheaper to collect.

Two stages for PT demand estimation are defined. The first stage serves for the planning activities of initial network construction and initial timetable development. Since service performance data are not available at this stage, demand has to be inferred from population and traffic flows as well as comparative studies. If a service improvement for an existing service is required, actual demand data or demand estimates from the service in use can be obtained. In this chapter, both stages are discussed, but emphasis is placed on the second case, that is the demand estimation for an existing network, as the majority of the most recent research has focused on this stage.
The remainder of this chapter is as follows. The following section briefly reviews demand modelling based on four-step modelling, activity-based models and newer ideas that can contribute to demand estimation for new public transport services. This is followed by an introduction to the different levels of detail at which demand information can be obtained, as well as their role in the planning process.

The chapter then focuses on the demand estimation methods with data from an existing PT service. For this, methods using emerging massive passive PT data to estimate route-level demand in terms of stop boarding/alighting flows and “leg-Origin Destination (leg-OD)” are discussed. Leg-OD refers to a combination of boarding and alighting points for a line. The discussion includes straightforward methods based on observed passenger flows (on-board surveys, APC and AFC data) and a recently proposed idea of using bus automatic vehicle location (AVL) data. The feasibility of using AVL data to estimate demand offers great potential benefits to the “data-poor” practitioners. The penultimate section provides methods estimating network-level demand in terms of origin-interchange-destination flows, referred to as “Journey-OD”. Finally, the remaining issues and future research directions are discussed.

**Table 21.1 Public transport network planning process – adapted from Ceder (2007)**

<table>
<thead>
<tr>
<th>Independent input</th>
<th>Planning activity</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Land use characteristics;</td>
<td>Level A</td>
<td>Interchanges and terminals;</td>
</tr>
<tr>
<td>Authority constraints;</td>
<td>Network design</td>
<td>Fixed routes and stops</td>
</tr>
<tr>
<td>Demand by time of day and day of week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Service standards;</td>
<td>Level B</td>
<td>Alternative setting of frequencies and headways;</td>
</tr>
<tr>
<td>Time for first and last trips;</td>
<td>Timetable development</td>
<td>Selected public timetable</td>
</tr>
<tr>
<td>Running times;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deadhead times;</td>
<td>Level C</td>
<td>Minimum fleet size;</td>
</tr>
<tr>
<td>Recovery times;</td>
<td>Vehicle scheduling</td>
<td>Vehicle schedules</td>
</tr>
<tr>
<td>Schedule constraints;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost elements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crew list;</td>
<td>Level D</td>
<td>Crew schedules and duty rosters</td>
</tr>
<tr>
<td>Crew work rules, rotation rules, constraints</td>
<td>Crew scheduling</td>
<td></td>
</tr>
</tbody>
</table>

The remainder of this chapter is as follows. The following section briefly reviews demand modelling based on four-step modelling, activity-based models and newer ideas that can contribute to demand estimation for new public transport services. This is followed by an introduction to the different levels of detail at which demand information can be obtained, as well as their role in the planning process.

The chapter then focuses on the demand estimation methods with data from an existing PT service. For this, methods using emerging massive passive PT data to estimate route-level demand in terms of stop boarding/alighting flows and “leg-Origin Destination (leg-OD)” are discussed. Leg-OD refers to a combination of boarding and alighting points for a line. The discussion includes straightforward methods based on observed passenger flows (on-board surveys, APC and AFC data) and a recently proposed idea of using bus automatic vehicle location (AVL) data. The feasibility of using AVL data to estimate demand offers great potential benefits to the “data-poor” practitioners. The penultimate section provides methods estimating network-level demand in terms of origin-interchange-destination flows, referred to as “Journey-OD”. Finally, the remaining issues and future research directions are discussed.

**Demand estimation for new services**

If a new service is planned, no actual demand data in any form are available, but it is the latent and induced demand that need to be estimated. Latent demand describes those attracted to the service from other modes and induced demand the additional new demand generated by a better transport service.

Classic transport planning literature uses four-step models consisting of trip generation, trip distribution, mode choice and assignment to provide a basic estimation of demand (Ortúzar & Willumsen, 2011). For estimating latent demand, the existing traffic OD matrix of a city or region might be taken and then, assuming a specific public transport network design, discrete choice models can be applied to obtain modal splits to obtain public transport journey OD flows. Assignment models can then be subsequently applied to obtain leg-OD flows, expected boarding, alighting and on-board flows. A number of problems can arise, such as dealing with
Demand estimation for network planning

dynamics during a day and, over longer time periods, a consideration of user habits, as well as effects such as crowding and consideration of detailed timetable design that affect the attractiveness of the service which will require iteration between mode choice and assignment. There is a large body of literature; among others, Gentile and Noekel (2016) provide extensive coverage of the topic. Fonzone et al. (2016) furthermore contribute to the understanding of the circular effect of intelligent transport systems on demand patterns, discussing how information influences demand patterns during the different journey stages.

Despite all its shortcomings, it is fair to say that various modified versions of the four-step models are still commonly used by practitioners to obtain an initial, rough demand estimate. An increasingly popular improvement to the approach is the use of agent-based models. Compared to the “macroscopic” four-step model, here the population is “microscopically” modelled for a range of supply-side changes. Such models include “activity-based” approaches including the four steps from trip generation to assignment. Thus, not only latent demand but also induced demand can be better reflected. Software solutions such as MATSim are currently used to model large city-wide scenarios based on public transport data, as in Fourie et al. (2016), but also to model demand for totally new systems, as seen in Bischoff and Maciejewski (2016).

Finally, it can be noted that “big non-PT data” solutions are increasingly being used for demand estimation of new public transport services. For example, Bonnel et al. (2018) have been using mobile phone data to estimate OD matrices. Further, comparing the target city to a wide range of other cities and their socioeconomic, geographic, cultural and political situations can help to obtain initial overall demand estimates. Related to this is the potential of using map data to estimate latent and induced demand. Bast et al. (2015) have shown that very good population estimates can be achieved with map data. These and the distribution of “points of interest” in a city can then be used to estimate trip generation and trip distribution. Koca (2020) has been using OpenStreetMap data in combination with machine learning approaches to predict the OD matrix of taxi trip data for New York. Whether a similar approach to “learning” will also work to predict the demand for public transport needs further investigation. Compared to taxi trips that are less constrained by the infrastructure, public transport demand requires far more specifications in terms of network design and service characteristics.

The next section focuses on the cases where a service is already existing so that (limited) data are available. The demand for PT planning is classified at the three aforementioned levels: stop flows, leg-OD flows and journey-OD flows.

Levels of demand information and their roles in planning

Stop alighting/boarding flows and passenger loads

Stop alighting/boarding flows and passenger loads play a critical role in timetable development. This information can be collected in a variety of ways, including (automatic or manual) counters at the boarding and alighting door, cameras on-board or at the vehicle stop or smart card readers. Passenger loads, in terms of on-board passengers as a vehicle arrives at a stop, can then be reliably derived by processing the data streams of passenger boarding and alighting flows at a low cost relative to direct observation. These three datasets are termed the “stop-level demand profile”.

A simple time-tableing strategy is to schedule services with equally spaced headways (see also Chapter 32). In this case, headways can be determined based on the overall demand during the operational period of concern. An advanced alternative strategy is to equalise passenger loads per vehicle, considering the demand dynamics. The criteria for load balancing can be
the average load over the route, the load at a predetermined checkpoint or the maximum load. Different timetables are usually assigned to holiday periods due to their specific demand profile. Assuming that boarding demand follows a uniform distribution, the demand can be interpreted as the arrival rate per minute or per hour in the interval and the calculated expected passenger waiting time given a timetable plan.

Knowing the stop-level demand also supports the operator in inserting appropriate slack time into the timetable. The numbers of boarding and alighting passengers collectively illustrate the “busyness” of the PT stop. When large passenger flows frequently occur at a specific PT stop in an operating interval such as the morning peak, the delays due to the unplanned long dwell time are not random but predictable. They can therefore be removed by introducing slack time into the timetable in the planning stage. Ideally, the stop-level demand profile should not only include mean passenger numbers but also information on the expected variance.

Finally, stop-level demand can help to assess and modify timetables based on passenger-focused criteria beyond unweighted service regularity aspects. For example, Chen et al. (2009) measure the passenger-weighted reliability for a bus transit network, using a punctuality index, a headway deviation index and a headway evenness index. The former is based on the route and the latter two on a stop. Watanabe et al. (2017) show the trade-off between relaxed and tightened schedules considering the demand at each stop.

**At route level (leg-Origin Destination)**

Leg-OD captures the boarding and alighting stop of each PT ride and aggregates the rides by OD-pair for each line. Dominant OD-pairs can be observed. This information enables the operator to adjust capacity according to the expected route-section demand, by, for example, serving certain route sections during the day more frequently. It further allows for efficient short-turn strategies during disruptions and the potential of introducing express strategies for important OD pairs. Short turn is to schedule a portion of the fleet to serve short cycles on the route segments faced with high demand, and express services operate on the entire route but skip several normal stops (Cortés et al., 2011). Cortés et al. (2011) and Tirachini et al. (2011) distinguish passenger boarding demand at each stop based on whether the alighting stop is located within the short-turn cycle. This connection between boarding and alighting stops is only available in leg-OD or journey-OD data.

**At network level (journey-OD)**

Journey-OD connects a series of PT rides conducted by a specific user within the time and distance threshold. It is useful for transfer-oriented timetable development. Ceder et al. (2001) develop a mixed integer linear program to optimise the timetable with the objective of maximising the number of simultaneous bus arrivals at transfer nodes. Ibarra-Rojas and Rios-Solis (2012) maximise the transfer synchronisation and meanwhile reduce unnecessary bunching between the buses of different lines. They propose an integer programming method extending the work of Ceder et al. (2001) and enhance the computational efficiency by eliminating a series of decision variables. In light of the network-level demand information provided by the journey-OD matrices, practitioners can furthermore optimise the bus network design, route configuration and stop location, such as increasing or decreasing overlapping route segments, combining the bus stops to reduce walking time for transfer passengers. Automatic fare collection technologies have made it possible to filter and chain the trips of users throughout the multimodal PT network, including bus, metro, railway and ferry. With advances in the integration
Demand estimation for network planning

of fare collection and the emergence of smart mobility platforms, it should be possible to create journey itineraries for individuals in an extended transport network with new feeder modes such as bicycle and car sharing. Thus, knowledge on PT users’ (true) origin and destination can support evolving PT network design even considering competition and cooperation between traditional PT modes and emerging feeder modes.

Time-varying mean and distribution of passenger flows

For both stop and OD flows, variations over time are an important issue. Passive PT data make it easier for practitioners to collect the data from various operating intervals. Figure 21.1 illustrates the passenger demand profile in the morning and evening peak periods obtained from smartcard data from a bus line in Shizuoka, Japan. Significant differences between the demand patterns can be observed in these two periods. Bie et al. (2015) employ bus dwell time as the proxy for passenger alighting/boarding flows and develop a method to identify time-of-day operating intervals from AVL data for the bus operators.

Planning is usually based on mean passenger flows or rates in an interval, though the knowledge of the distribution during a time interval can offer a fuller picture to the practitioner with respect to the degree and probability that the demand can be outside the expectation. Figure 21.1(c) reveals that in the morning peak, the load can exceed 50 passengers between Stops 17 and 24, which indicates severe crowding on some buses. However, the mean loads in this segment remain between 30 and 37 passengers, showing a normal on-board condition. Taking Stop 20 (the mean maximum load point) as the checkpoint, it can be observed that the passenger load exceeds 50 for six times in 30 sampled runs during the morning peak hour. This high probability that passenger load can deviate by more than 50% upward from the mean deserves the operator’s attention.

For a bus route, passenger flows can be classified by run if the run ID is recorded when the passenger taps in. However, a passenger assignment model is required when it comes to a metro or railway route. The APC and AFC data for railway/metro route only record the time point when the passenger enters and leaves the PT station; thus, extra effort is needed to relate an individual passenger to a specific service. Zhu et al. (2017) develop a probabilistic approach for this. They connect AFC data and AVL data and attach the boarding probability to a series of candidate train vehicles for a single passenger (see also Chapters 33, 34, 35, 36). They presume that the time points of passengers entering/leaving the system and the arrival times of train

Figure 21.1a  Observed boarding flows
vehicles at origin/destination stations are given. With this information and the walking time from entrance to platform at the origin station and from platform to exit at the destination station, as well as the capacity constraints, they calculate the boarding probability. In this way, the flows and the leg-OD matrix for each vehicle are replicated.

**Leg-OD estimation**

Following on from the classification introduced in the previous section, this section details approaches to estimating leg-OD flows from different PT data before discussing the chaining techniques to estimate journey-OD flows from AFC data or by combing several data sources in the next section.

First, a review of the methods regarding estimating leg-OD matrices for a bus route from AFC or on-board survey data is shown, combining APC with survey data, as well as APC data by itself. Often, there is missing data, and so some discussion is included on the solutions used to address the missing origins or destinations in incomplete AFC data. Second, for practitioners who have difficulty in collecting data that directly observe passenger flows such as AFC and APC data, a recent method to infer expected OD flows using bus AVL data is provided.
Estimation with survey, APC and AFC data

Figure 21.2 illustrates the passenger flows for a general bus route. Consider that there are $K$ buses in total in a time-of-day operating period for a bus route having $N$ stops. For each bus run $k$ in this period, let $B_{i,k}$, $A_{i,k}$ and $O_{i,k}$ denote the boarding flow, alighting flow and passenger load at stop $i$, with $i$ and $j$ indicating the origin and destination stops, leading to $Q_{i,j,k}$, which represents specific OD flows for each bus run. The OD matrix of a general bus route is provided in Table 21.2, omitting the index of a bus run. There are $N(N−1)/2$ non-zero elements in the leg-OD matrix.

Boarding flows, alighting flow and passenger loads given observed OD flows are obtained using Equations (21.1)–(21.3). Note that $B_N$ and $A_1$ are zero, as no passenger would board at the last stop or alight at the first stop. The passenger load is defined as the number of on-board passengers at the arrival, not the departure, of the bus; $O_i$ is therefore also zero.

\begin{align*}
B_{i,k} &= \sum_{j=1}^{i-1} Q_{i,j,k} \quad i = 1,2,\ldots,N - 1 \quad \text{Equation 21.1} \\
A_{i,k} &= \sum_{j=1}^{i-1} Q_{i,k} \quad i = 2,3,\ldots,N \quad \text{Equation 21.2} \\
O_{i,k} &= O_{i-1,k} + B_{i-1,k} - A_{i-1,k} \quad i = 2,3,\ldots,N \quad \text{Equation 21.3}
\end{align*}

![Figure 21.2](image)

**Table 21.2** OD matrix for a transit route (omitting the bus run index $k$)

<table>
<thead>
<tr>
<th>Origin</th>
<th>Destination</th>
<th>Aggregates</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$Q_{1,2}$</td>
<td>$Q_{1,i}$</td>
</tr>
<tr>
<td>$i$</td>
<td>$Q_{2,i}$</td>
<td>$Q_{2,j}$</td>
</tr>
<tr>
<td>$N-1$</td>
<td></td>
<td>$Q_{i,N-1}$</td>
</tr>
<tr>
<td>$N$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Aggregates**

$A_K ≡ 0$  $A_2$  $A_i$  $A_{N-1}$  $A_N$
Hazelton (2001) analyzes the subtle distinction between replicating the OD trips and estimating mean OD trip rates (expectation). The discrepancy is demonstrated in the OD estimation for traffic flows. OD trips fluctuate in each homogeneous observational period, for example, different hours in the whole morning peak or the same hour in different days. Mean OD trips or OD trip rates can be employed to characterize these homogeneous periods in order to significantly reduce the estimation complexity. In the estimation for PT routes, the reconstruction problem, with the objective to accurately estimate the OD flows for each bus run in a time-of-day interval, should be distinguished from the estimation problem, with the objective to infer the expected OD flows in that period. For each time-of-day operating period, mean flows or arrival rates per minute (per hour) can characterize the demand pattern. The fluctuation in the flows of the bus runs in a same interval is usually the result of headway fluctuations. The OD estimation can therefore be parametrized if headway data can be collected. Generally speaking, there are two approaches. The first one is intuitive. It models the OD flows as OD-pair arrival rates $a_{ij}$ multiplied with the associated headway, as shown in Equation (21.4).

$$Q_{a,ij,k} = a_{ij} \Delta_{i,k}$$  \hspace{1cm} \text{Equation 21.4}$$

It is thus possible to directly estimate the rates if the leg-OD matrix is available for adequate bus runs, for example, collecting AFC data or large-scale on-boarding survey data. In an on-boarding survey, a leg-OD matrix sample can be constructed by asking the passengers of one bus trip to provide their boarding and alighting stop, but there might be “non-structural zeros” in the matrix sample (Ben-Akiva et al., 1985; McCord et al., 2010; Ji et al., 2015). These are the zero OD-flows that are not captured by the random survey, and they may still exist when aggregating surveys from different runs to create a less biased matrix unless the sample size is very large. Thus, the passenger OD rates drawn from these samples are likely to be biased. On the other hand, APC data provide passenger observations at a much finer scale, with the shortcoming that the passenger flows captured by the counter at boarding and alighting doors are not linked. Therefore, surveyed OD matrices can be taken as the base matrix and balancing techniques are employed to obtain the OD rates at a finer-scale.

To make use of the boarding and alighting counts, which are reliable control totals, a second approach to model the OD flows is introduced. It assumes stop-based boarding rates and alighting probabilities. Let $b_i$ denote the arrival rate for stop $i$; $c_i$ the alighting probability, defined as the proportion of the alighting passengers over the number of on-board passengers at stop $i$ and $\rho_{ij}$ the conditional alighting probability at stop $j$ given passenger boarding the bus at stop $i$. The flows can then be rewritten as in Equations (21.5) and (21.6). If $A, B, O$ and $\Delta$ are observed, $b$ and $c$ can then be calibrated by statistical analysis. This means that, given boarding and alighting counts, the estimation problem is transformed into the estimation of the conditional alighting probability matrix $\rho'$ as in (21.7) constrained by alighting proportion $c$. The right-hand side of (21.7) is then estimated with conditions (21.1) to (21.3).

$$B_{i,k} = b_i \Delta_{i,k}$$  \hspace{1cm} \text{Equation 21.5}$$

$$A_{i,k} = c_i O_{i,k}$$  \hspace{1cm} \text{Equation 21.6}$$

$$Q_{a,ij,k} = \rho'_{ij} b_i \Delta_{i,k}$$  \hspace{1cm} \text{Equation 21.7}$$
Ben-Akiva et al. (1985), McCord et al. (2010) and Mishalani et al. (2011) estimate probability matrix \( a' \) instead of arrival rate or volume matrix from unconnected boarding-alighting counts and base OD matrices derived from on-board survey data, using iterative proportional fitting (IPF). A problem remains, however, with the previously mentioned non-structural zeros. Ben-Akiva et al. (1985) point out that the non-structural zero entry to the surveyed matrix can be explained by a non-zero probability, which captures passenger travel patterns more correctly. McCord et al. (2010) introduce a null matrix assuming that a boarding passenger has equal probability to alight at all the downstream stops, and the starting point of the probability matrix is then a matrix weighted by a null and surveyed matrix whose weights are dependent on the survey size. This also addresses non-structural zeros. Mishalani et al. (2011) discuss how increasing the sample size of the on-board OD survey improves the estimation performance; in their case, the sample size refers not to the number of sampled runs but the number of passengers that report the leg-trip they made on this specific bus route. Under a data environment with large quantities of passenger counts provided by APC data and base OD matrix, Ji (2011) proposes a heuristic expectation maximisation (HEM) approach and comprehensively compares it with IPF, expectation maximisation (EM) and conditional maximisation (CM) algorithms. Ji et al. (2017) use stop-based farebox data as the proxy of boarding counts and the base matrix collected by Wi-Fi sensors.

The problem becomes increasingly difficult if the base matrix is totally unavailable. Consider a case when only APC data are available. As illustrated in Table 21.2, the number of unknown OD flows is \( N(N-1)/2 \), while the number of observations and linear equations determined for the system is only \( 2N \); the problem therefore becomes underspecified if \( N > 5 \). Several studies break down this problem by estimating the destination distribution for the boarding passengers at each stop. A naïve assumption is that the boarding passenger alights at each downstream stop with equal probability. Li and Cassidy (2007) assume a conditional alighting probability matrix and categorise all the bus stops into major and minor stops according to land use information and APC data. Li (2009) models the conditional alighting probability matrix with Markov chains. The probability of a passenger’s alighting at a specific stop is considered dependent only on whether s/he is on-board at the previous stop. Hazelton (2010) suggests that this is unrealistic and proposes a Markov chain Monte Carlo (MCMC) method using the transition probabilities derived by Li (2009)’s model as the initial “proposal distribution” in metropolis sampling. Ji et al. (2015) further find that the HEM method using only APC data can match the performance by IPF using APC data and a base matrix constructed by relatively few passengers. They confirm the performance match when more than 500 runs are collected on a bus route with 18 stops, 153 feasible OD pairs and 100 passengers being surveyed. Ji (2011) is recommended as additional reading material, as it comprehensively summarises the (up to then) existing leg-OD estimation methods considering their computational differences and similarities.

Inference of boarding and alighting stop with automatic fare collection data

All of the previous research becomes obsolete if complete AFC data that contain both tap-in and tap-out are recorded. If records are incomplete, with APC data also available, it is possible to obtain the percentage of smart card users among all users and scale up demand accordingly by inversely applying the calibrated percentage as the expansion factor. There may be a need to correct for biases, though, given that regular users are more likely on certain OD pairs and are also more likely to possess smart cards. Accordingly, the percentage of smart card versus
non-smart cards users may vary in different time-of-day periods, on different routes, and at different stops of a route due to the connectivity of the stop.

The data regarding boarding or alighting stops could be missing for a couple of reasons. First, the completeness of data is highly dependent on the fare collection policy and transfer policy, as only one record, either tap-on or tap-off, is needed if there is a flat-fare system. Among distance and zonal fare, structures that charge according to the actual distance travelled (or zones traversed) and those charging for ODs independent of the route need to be distinguished. For the latter fare structure, the AFC records usually do not entail records on transfer points. This means that, again, leg-OD matrices are not straightforwardly available from the data. In that case, one can specify the path by combining AVL data to analyze the time interval between passenger arriving time to the stop, arrival and departure time of transit vehicle and passenger leaving time (see also Chapters 33, 34, 35, 36).

For flat-fare PT networks, hence, the inference of boarding or alighting stops is required to complete the leg-OD matrix. Several studies propose methods by creating a database in which each passenger made at least two trips in one day (Barry et al., 2002; Trépanier et al., 2007; Munizaga & Palma, 2012). The missing information is inferred based on onward journeys as well as other day-to-day patterns, as also discussed in some more detail in the “journey-OD” section.

**Estimation based on automatic vehicle location data**

Sun (2020) considers the problem from the perspective of “data-poor” practitioners without access to AFC and APC data and provides a new approach using bus AVL data as the main data source to infer passenger flows. AVL data are collected by an in-vehicle chip that keeps recording the spatial-temporal coordinates locally or uploads them to the server at a predetermined frequency, for example, every 10 seconds. No privacy issue relates to this data, and it is usually managed by the bus operators themselves or through their data providers. If the data collection frequency is high enough, or if it includes “event data” such as door opening, door closure and stand-still at a bus stop, then bus AVL data contain a rich detail level regarding the time components in the bus trips, from which passenger activity times (boarding and alighting times) can be extracted. With the help of a small survey to obtain the information on individual boarding/alighting time, the connection between passenger flows and passenger activity times can be captured. However, it is in the form of a series of underspecified formulations; that is, the unique boarding and alighting number of passengers is not available.

Figure 21.3 shows the model framework to estimate OD flows. The demand for OD-pair specific arrival rates are the estimated parameters. Once these are known, stop-level flows are also derived. \( a_{ij} \) is assumed to follow a modified gravity model, as in Equation (21.7):

\[
a_{ij} = \frac{p_i^g p_j^a}{(j-i) + \frac{\alpha}{j-i}} \quad i = 1, 2, ..., N - 1; j = 2, 3, ..., N; i < j
\]

Equation 21.7

where \( p_i^g \) denotes the generation power of stop \( i \), and \( p_j^a \) denotes the attraction power of stop \( j \). The first term in the denominator is the travel distance, and the second term is to acknowledge that a passenger is not likely to make a short bus trip travelling only one or two stops. The two vectors \( p^g \) and \( p^a \), each of size \( N - 1 \), and parameter \( \alpha \) substitute the arrival rate \( a_{ij} \) as the new unknown parameters. The number of unknown parameters to estimate is hence reduced from \( N(N - 1)/2 \) to \( 2(N - 1) + 1 \). Sun (2020) turns to Bayesian inference, which has been shown
Demand estimation for network planning

to be powerful in inferring OD trips given limited observations or observations with uncertainty (Maher, 1983; Hazelton, 2001, 2008, 2010). A Hamiltonian MCMC algorithm is then employed to estimate the unknown parameters.

Results of this AVL-based approach are shown in Figure 21.4. Estimation results for a stop-level demand profile are illustrated in (a)–(c) in the form of the mean and 95% confidence interval, and the estimated versus observed OD flows are displayed in (d). The

Figure 21.3 A framework for OD estimation using bus AVL data. Items in grey boxes are the observations and input parameters; items in white boxes are the unknown and derived parameters.
Figure 21.4a  Boarding flows (observed and estimated from AVL data)

Figure 21.4b  Alighting flows (observed and estimated from AVL data)

Figure 21.4c  Passenger loads (observed and estimated from AVL data)
distribution of the observations presents the actual demand fluctuation of the bus runs dispatched during a time-of-day operating interval, 7 am–8 am in this case. The distribution of estimation results, however, are the possible values for mean passenger flows given the distribution of mean passenger rates drawn by the MCMC algorithm and the observed headways. Therefore, the variation in estimated passenger flows is less pronounced than those for the observed passenger flows. Nevertheless, the mean estimates appear reliable and useful for service planning.

**Journey-OD estimation**

Finally, the focus is on deriving complete journey itineraries across the multi-modal PT network for individual passengers in an observational period. If leg-ODs are available, estimating the journey-OD means to estimate transfer probabilities for all possible lines for alighting passengers. Similar to the discussion in the previous section, correlation between boarding and alighting flows at a stop and taking time differences between arriving and leaving services can be taken into account to “learn” transfer probabilities with Bayesian inference and other approaches.

A different problem formulation arises from incomplete AFC data with tap-in records only. In that case, “panel data” of users are available in the form of subsequent boarding points. The objective then is to connect these data points to obtain complete user trajectories. Consider a multi-modal network consisting of railway lines under a distance-based fare structure and bus routes under flat fare structure so that tap-out information is not available for bus trips. For rail-only trips, no inference is needed, but it is for the remaining four types of subtrips in a journey: single bus trips, trips with railway-to-bus or bus-to-rail transfers and trips with multiple bus legs. For a trip with a bus-to-railway transfer, the alighting point for the first leg is missing but can be easily inferred from the boarding point for the second leg. For the other three trip types, the origin of the next trip as information has to be leveraged so as to infer the alighting point of the previous trip. This might also include information derived from different days. A database containing one-day journey itineraries classified by passenger smart card ID can be feasibly constructed from AFC data. These basic principles of estimation are first proposed by Barry et al. (2002): 1) users will return after their activity to the alighting point of the previous trip; 2) in the same day, the destination of the last trip is identical to the origin of the first trip; 3) the unknown alighting location of a bus trip is the stop closest to the origin of the next trip. Trépanier et al.
Wenzhe Sun and Jan-Dirk Schmöcker

(2007) extend the database by constructing multiple-day itineraries, identifying mirrored journeys in a longer timespan. Zhao et al. (2007) apply maximum walking distances of 400 m or 5 min to detect the alighting stop given the next boarding stop in addition to the basic assumptions. Munizaga and Palma (2012) propose an algorithm minimising the total travel time of a journey, in particular the transfer time, filtering out the options not likely taken by transfer users. Alighting stop detection based on trip symmetry can be time consuming and data demanding.

Alternatively, Gordon et al. (2018) relax the data requirement and attempt to build journey itineraries with the data in an observational period shorter than one day, such as a time-of-day period. As the boarding stop of the next leg is not available for the destination inference, they reduce the number of OD pairs in a PT network. They attach two action labels to each station (entry or exit at the station) and generalise for bus nodes by attaching two direction labels to a whole route (northbound or southbound). The journey OD flows are therefore captured and classified by all possible combinations of itineraries. A checklist-based method identifies each unique itinerary, and the flows are aggregated by itinerary. For the example network shown in Figure 21.5 there are 38 feasible itineraries, such as “Bus Route 3 northbound, Station C entry, Station B exit”. The base counts for these itineraries are directly observed from AFC data or inferred using the basic principles introduced previously. Control totals are collected at the railway station gates and bus fareboxes, which are used to obtain expansion factors. The journey expansion algorithm is based on IPF.

**Conclusion**

The aim of this chapter has been to provide a systematic introduction on the demand estimation approaches serving PT network planning. The new service construction problem has been distinguished from the problem of estimating demand for existing services. For new services, demand estimation is mainly based on population and land use data as well as flows observed for existing transport modes. Information about existing “potential” demand can be estimated from an increasingly broad range of data. The role of map data has been discussed, but other data sources such as mobile phone records, Wi-Fi records, sales data or information that can be inferred from social media have not been discussed in detail.

By estimating demand for existing services, the gaps between demand and supply can be identified for further improvement purposes. In this stage, actual passenger demand patterns are available in the form of massive passive PT data, such as AFC and APC data. Three basic levels
have been defined, namely the stop-level demand profile for a transit route, leg-OD matrix for a transit route and journey-OD matrix for a PT network. The chapter then discussed approaches for each of these levels. There is a significant body of literature using AFC and APC data. It is pointed out that AVL data are also a potential information source if the higher-quality direct passenger flow data are not available. This has been the recent research topic of the authors, and an example is shown. As a practical note, low-cost actions can significantly enhance the quality of AVL data and further improve the quality of demand estimation. One example is automatically recording the stop and time information for each door opening and closing activity. This can help to extract precise passenger activity time and make it possible to estimate passenger flow, even with low-frequency bus AVL data.

More generally, it is hoped that the chapter can “broaden the mind” of operators and researchers, showing how different (incomplete) data can be used and combined to obtain better information for service planning. The abundance of passive data in combination with machine-learning methods offers large potential that is starting to be “standard” via integration in mainstream planning software.

A range of challenges still remain with respect to demand forecasting for network changes or if disruptions occur where a gap appears to remain between the approaches discussed in this chapter. The vast majority of studies using passive big PT data concern the reconstruction of demand. More recently, four-step and agent-based modelling has started to benefit from the findings of big data studies, but comprehensive frameworks as to how AVL, APC and AFC data can be used to derive sensitivities to key transport policy variables are still in their early stages and require further research.

This chapter was largely written before the impact of the COVID-19 pandemic was appreciated. The methodologies discussed in this chapter are clearly not directly affected by the current crisis, but the implications of the pandemic around the globe, its lockdown and the emergence from lockdown are beginning to be realised. First, the avoidance of contact in general and in public transport in particular will encourage customers and operators to utilise cashless payment systems even more. For demand estimation, this is advantageous, as often only the electronically collected data are used for analysis. Furthermore, the increase in (anonymous) tracking data as well as the increased utilisation of pre-booking a seat in public transport services might be valuable data for activity estimation as well as demand prediction. Second, changes in passenger behaviour following the pandemic will lead to a need to re-calibrate demand estimation such as sensitivity in choice with respect to other alternatives. In this chapter, this has particular relevance to possible changes in the relationship between dwell time and demand. More social distancing might mean longer boarding time per passenger. For the demand estimation approach with dwell time data presented in this chapter, this would spell good news, as the effect of each passenger on the observed dwell time would be pronounced. Furthermore, the crowding-averse tendency simplifies the dwell time function in the estimation, as the boarding passengers are more willing to wait until the alighting process ends, and friction due to crowding might less often occur. As a result, post-COVID-19 models might be easier to calibrate.

References


