Paper No. 03-2179

Analytical Model for Paratransit Capacity and Quality-of-Service Analysis

Liping Fu

An analytical model could potentially be used by paratransit service planners to predict fleet requirements, system capacity, and quality-of-service measures for specific operating conditions. The model has a sound theoretical origin and was calibrated using data from a large number of simulated cases representing a wide range of operating conditions and quality of service. This model was shown to have a strong explanatory power, capable of capturing the complex relationship between fleet size, travel demand, quality-of-service measures, and other operating condition variables. With this model, analytical procedures similar to those provided in the *Highway Capacity Manual* and the *Transit Capacity and Quality of Service Manual* for other transportation facilities and services could be developed for paratransit systems.

The first edition of the *Transit Capacity and Quality of Service Manual* (TCQSM) has provided the transit industry with methodologies and techniques for capacity and quality-of-service analysis of the two main transit modes, namely, fixed route bus service and rail service (*1*). The current edition does not, however, provide an in-depth coverage on the demand-responsive transit mode, Paratransit. Technical procedures similar to those for bus and rail transit are not available for paratransit capacity and quality-of-service analysis. The manual provides the following reason and suggestion:

"To date, no national studies have been performed on demand-responsive person capacity, particularly for dial-a-ride types of service, so this chapter does not provide calculation procedures for estimating demand-responsive capacity. However, the following general statement about capacity can be made: a demand-responsive vehicle's person capacity is inversely related to the size of its service area and also is inversely related to the number of potential origins and destinations it must serve. The best method for estimating demand-responsive person capacity is to identify a well-used demand-responsive system serving an area similar to one for which service is contemplated, and to identify the number of passengers per hour or per day that system is capable of serving" (*1*).

The goal of this research is to address the methodological gap associated with paratransit capacity and quality-of-service analysis. The specific focus of the research is on the following planning and design questions:

- What is the minimum number of vehicles or fleet size required for specific operating conditions and quality of service?
- What is the maximum number of trips that can be serviced by a given fleet of vehicles?
- What quality of service could be maintained with a given fleet of vehicles?

Department of Civil Engineering, University of Waterloo, 200 University Avenue, W., Waterloo, ON, Canada N2L 3G1.

To address these questions in a technically sound way, models that relate parameters such as fleet size, travel demand, level of service, and environmental variables must be developed. Paratransit service systems are, however, notoriously complex. They are dynamic, involving a set of vehicles traveling from place to place to provide transport services in a stochastic environment; their operations require solving routing and scheduling problems that are not amenable to optimal solutions; and they are full of peculiarity in operating conditions such as service coverage, network topology, traffic congestion, fleet mix, and spatial and temporal variations in travel demand. Developing an analytical formula that is capable of capturing all these complexities has been considered extremely difficult, if not impossible. As a result, simulation method has become the most favorable choice among various modeling methodologies for paratransit analysis (*2–4*). Simulation models have been found effective in representing complex relationships such as those observed in paratransit systems. However, they also require large amounts of data, significant preparation efforts, and long computational time, and they are therefore expensive and time consuming to use and unsuitable for parametric analysis. Furthermore, estimates obtained from a simulation model may differ considerably between different runs, depending on the choices made on scheduling algorithm and parameter values. These limitations could be serious from practitioners' points of view, for what they really want is a tool that can provide quick solutions to their service planning and design problems.

Limitations of the simulation method had also motivated early attempts to develop analytical models to predict performance measures and fleet requirements for dial-a-ride paratransit systems. The two models of particular relevance are Wilson et al.'s empirical model (*5*) and Daganzo's theoretical model (*6*), as reviewed in detail in the next section. Both models assumed idealized operating conditions and were developed for dynamic dial-a-ride systems in which all trips are demand trips that must be scheduled in real time. As a result, they are no longer applicable for today's mostly reservation-based paratransit service systems. Another notable study was conducted by Stein (*7*), who developed several asymptotic models for estimating route length and average service time for a dial-a-ride system. Again, the models were for extremely idealized conditions.

This paper describes the development of an analytical fleet size and performance model that could potentially be used by paratransit agencies to conduct paratransit resource requirements and quality-of-service analysis. The paper first presents an overview of two existing analytical models, and the proposal of a new analytical model. That proposed model is subsequently calibrated for a set of idealized conditions. The impacts of prevailing conditions on minimum fleet size are then investigated, with the objective of quantifying the magnitude of these impacts for postadjustments.

Finally, a practical example is used to illustrate the application of the proposed model.

MODEL STRUCTURE

The structure of the proposed model was motivated by two existing models found in literature. Both models were developed for demand-responsive paratransit systems in which all trips are assumed to be demand trips that need to be scheduled in real time as they enter the system. The first model was proposed by Wilson et al. (*5*), which was calibrated statistically based on operating data generated by a simulation model. The model is of the following form:

$$FS = \frac{0.68A + 0.072\lambda}{E^{0.5}} \tag{1}$$

where

- FS = minimum fleet size or number of vehicles required for a given dial-a-ride system,
- A = size of the area covered by the service (mi²),
- λ = average trip demand or trip arrival rate (trips/h), and
- E = level-of-service indicator defined as the average ratio of excess service time to direct driving time of all the users; the excess service time of a given user is the difference between the total time in system (waiting time plus ride time) and the direct driving time.

The relationships depicted by Equation 1 are intuitive because the minimum number of required vehicles increases as the service area (A), service demand (λ), and quality of service increase. This model, however, has two limitations:

1. The operating conditions assumed by the model are no longer the true representation of current practice. For example, the equation was developed for dynamic service systems in which all trips are assumed to be same-day demand trips. Most paratransit services currently in operation are, conversely, reservation based, in which trips are mostly known in advance and scheduled before service starts. This type of operation is expected to be more efficient than a dynamic service system, for most trips are known in advance and thus better routes and schedules could be generated.
2. The model was calibrated using data with limited variations in operating conditions as acknowledged by the developer. Furthermore, variations in travel speed, dwell time, algorithm settings, characteristics of service area, network topology, and demand clustering were not considered.

It should be noted that, from Equation 1, the average excess travel time ratio (E) could be expressed as a function of fleet size, service area, and demand density. This average excess travel time ratio is often considered as one of the quality-of-service measures in paratransit industry. Thus, the resulting equation can be used to predict the corresponding quality-of-service measure that would result with a given fleet size.

The second model was proposed by Daganzo (*6*) for the same type of systems as represented by Wilson's model. Daganzo's model was, however, derived analytically on the basis of geometric probability and queuing theory. By assuming a random demand distribution in time and space, and simple routing logic (e.g., nearest insertion algorithm), Daganzo was able to develop several approximate models for average service time in system (waiting time plus ride time) as related to fleet size, demand rate, and other operating parameters. Although not explicitly stated, one of the intermediate results led to the relationship between fleet size and other parameters as follows:

$$FS = \lambda \cdot \left[\tau + \frac{1}{V}\left(\frac{A}{n}\right)^{0.45}\right] \tag{2}$$

where

- FS = fleet size,
- V = average travel speed based on Euclidean distance (mi/h),
- τ = boarding plus alighting time (h), and
- n = average number of requests waiting to be picked up.

Note that Equation 2 should be viewed as the theoretical minimum fleet size and is valid only in an asymptotic probabilistic sense under idealized conditions such as random demand, no service time constraints, and simple routing methods.

Proposed Model

The two existing models discussed in the previous section, although not applicable to current practice, revealed the general intuitive relationship between minimum fleet size and system characteristics. For example, minimum fleet size should be directly related to trip demand (λ) and service area (A), and it should be inversely related to acceptable ride time ratio (E) and average travel speed (V). Both models, however, have advantages and disadvantages pertaining to the kind of system factors represented. For example, average ride time ratio, a service quality indicator, was included in Wilson et al.'s model (*5*) but not in Daganzo's model (*6*), whereas the opposite is true for the parameters average travel speed and dwell time. As a result, we proposed the following model structure as an attempt to combine the advantages of the two models:

$$FS = \beta_1 \frac{\lambda_T}{E^{\beta_4} T^{\beta_5}} \cdot \left[\tau + \frac{\beta_2}{V}\left(\frac{A}{\lambda_T \cdot T}\right)^{\beta_3}\right] \tag{3}$$

where

- A = size of the service area (km²);
- T = a quality-of-service constraint defined as the trip service (pickup/delivery) time window (h);
- E = quality-of-service measure defined as the maximum allowable ratio of excess ride time to direct driving time; the excess ride time of a given user is defined as the difference between the total ride time and the direct driving time;
- λ_T = peak trip rate defined as the equivalent hourly rate at which trips need to be serviced by the system within the evaluation time interval T (trips/h) (the term $\lambda_T\, T$ represents the number of trips during the peak period of duration T);
- V = average travel speed based on Manhattan distance (km/h);
- τ = boarding plus alighting time (h); and
- $\beta_1, \beta_2, \ldots, \beta_5$ = model parameters to be calibrated.

Note that the proposed model deviates from the two existing models in three important aspects. First, it includes a new parameter called trip service time window (T) to account for its possible impact on fleet size requirement. It is intuitive that the smaller the service time window, the larger the required fleet size. Second, the parameter E is now defined as maximum allowable excess ride ratio instead of average ride ratio as in Equation 1. The former is a policy variable that is commonly used as an input to a scheduling process and can be easily obtained without actually scheduling the trips. In contrast, average excess ride time ratio is an outcome of a scheduling process and therefore is not known until trips are scheduled. Finally, the concept of peak trip rate (λ_T) is introduced to account for the fact that maximum fleet size primarily depends on the peak demand during a day. The peak demand is defined as the maximum trip rate during the prespecified time interval T (usually less than 1 h) in a typical operating day. The reason for using trip service time window (T) to define peak demand is that in scheduling trips, the actual pickup and delivery time of each trip could be moved around within the time window T, which means that variation of demand within the subinterval T could be smoothed out, and its effect on fleet size could be ignored.

It should also be noted that the proposed model does not include many other important factors, such as the topology of the service area and street network, the spatial distributions of trip demand, and vehicle size and fleet mix. It was assumed that the impacts of these factors would be captured by the model parameters ($\beta_1, \beta_2, \ldots, \beta_5$) obtained from a calibration process. In recognizing the difficulty in capturing all these factors and their variations, a two-stage calibration method was proposed. The first stage is to fit the proposed model to a set of idealized conditions, and the second stage is to develop a set of modifiers that can be used to adjust the equation for conditions that deviate from the idealized conditions. This approach is consistent with the methodology that has been adopted by the *Highway Capacity Manual* (*8*) and TCQSM (*1*) for capacity and level-of-service analysis of other types of transportation facilities and systems.

CALIBRATION OF A MODEL FOR IDEAL CONDITIONS

This section discusses the calibration of the proposed model for paratransit services operating under a set of idealized conditions, which have the following characteristics:

- Geometrically square service areas:
  – Each service area is covered by a uniform grid road network with all neighboring nodes (intersections) connected by two links, one in each direction. Each link has a length of 500 m and a uniform constant speed.
  – Rectangular distance and average link speed can be used to calculate the exact travel time between any two points (or network nodes).
- Random demand:
  – Trip origins and destinations are uniformly distributed over the service area.
  – Each trip specifies either a desired pickup time (inbound trips) or a desired drop-off time (outbound trips) that is uniformly distributed during a 3-h service period.
- Uniform service vehicles:
  – All vehicles are assumed to be identical, with an extremely large seating capacity.
  – All vehicles start and end their routes at the same depot, which is located at the center of the service area.
  – There is no limitation in service length or shift length, which means the fleet size required during the peak period determines the minimum fleet size required for the whole day.
- Service constraints:
  – Users must be picked up or delivered within their desired time windows. For outbound trips, the delivery time window is $[D - T, D]$, where D is the desired delivery time and T is the time window. For inbound trips, the pickup time window is $[D, D + T]$.
  – The excess ride time of each trip must not exceed the maximum allowable ride time ratio (E).

The number of vehicles required also depends on how vehicles are scheduled and how the underlying dial-a-ride scheduling problem is solved. If schedules are created manually by schedulers, the outcome will depend on the level of skill and experience of the schedulers. For the computer-aided scheduling method, different representations of service policy and objectives in the scheduling algorithm (e.g., minimize fleet size, minimize total service time, or minimize user ride time) would lead to different solutions for the same operating conditions. That is because optimal algorithms are not feasible to solve practical dial-a-ride scheduling problems, and heuristic algorithms are often the only options available. As a result, for the development of a representative and consistent analytical model, a standard scheduling process must be specified. The scheduling process adopted for calibrating the proposed model has the following specifications:

- A computer scheduling software was used called FirstWin (L. Fu, *User's Guide: FirstWin—A Tool for Routing and Scheduling Dial-A-Ride Paratransit Vehicles,* Department of Civil Engineering, Waterloo, Ontario, Canada, unpublished data, 2002) to generate data for simulated operating environments. Operating cases, which are characterized by service area and network, travel demands, and service vehicles, were generated using a utility program. Trips are then scheduled using a procedure discussed later to determine the minimum number of vehicles required for a given operating condition and level of service.
- Two objectives are considered in scheduling: minimize fleet size and minimize the total travel time. The objective of minimizing fleet size was achieved by using a neighborhood-based sequential insertion algorithm (NSI), which schedules vehicles one at a time and uses trip-clustering knowledge in the insertion process (L. Fu, unpublished data, 2002). On the basis of numerous experiments with both simulated and real cases, it was found that this algorithm performed better than other algorithms, such as the parallel insertion algorithm, in regard to minimizing the number of vehicles required to service a given set of trips (*9*). The minimization of total travel time was explicitly considered in selecting trips and identifying optimal insertion positions in the scheduling algorithm.
- The maximum allowable ride time ratio and service time window are considered as hard constraints that must be satisfied for all trips. These constraints define the minimum level of service that must be guaranteed for all trips. Note that, in all tests, an unlimited fleet with large seating capacity is used to eliminate the influence of capacity constraint and the possibility of any trip rejection.
- The scheduling procedure was as follows:
  – Step 1. Import network/vehicle/trip data.

– Step 2. Set the maximum allowable ride time ratio (E) and service time window (T).

– Step 3. Schedule all trips using the NSI algorithm.

– Step 4. Remove all trips from those vehicles that have fewer than three trips assigned and try to reassign removed trips to other scheduled vehicles by using the Swap and Reinsertion algorithm included in FirstWin (L. Fu, unpublished data, 2002). This is an attempt to reduce the number of vehicles required for delivering the service.

– Step 5. Apply the Reinsertion and Exchange improvement algorithm to improve the generated schedules and go back to Step 3. This step continues for one iteration only.

– Step 6. Record the scheduling statistics including the number of vehicles that have been scheduled with trips (FS).

To generate data with sufficient variations in system conditions, we considered five factors, each of which was varied between two to three levels. The following combinations of settings were used:

- Service area (A, km^2): 10 × 10, 15 × 15, and 20 × 20;
- Time window (T, min): 20, 30, and 40;
- Maximum allowable excess ride time (E): 0.5, 1.0, and 1.5;
- Demand (trip) density (trips/km^2/h): 0.5, 1.0, and 1.5; and
- Vehicle average velocity (V, km/h): 20, 30, and 40.

A total of 243 (3^5) cases were simulated. Table 1 provides some sample data obtained from this process.

Because Equation 3 is nonlinear and cannot be transformed into a linear equation, we used the nonlinear regression procedure available in the statistics analysis software SPSS to fit the equation to the simulated observations (*10*). Similar to that of the linear regression method, the goal of a nonlinear regression process is to identify the best values for the model coefficients in regard to estimation error. Initial experiments indicated that the coefficient β_1 was approximately equal to 1.0 and β_5 was near zero. We therefore decided to drop the coefficient β_1 and the term β_5. The final regression equation is the following:

$$FS = \frac{\lambda_T}{E^{0.20}} \cdot \left[\tau + \frac{4.62}{V}\left(\frac{A}{\lambda_T \cdot T}\right)^{0.31}\right] \quad (4)$$

The R^2 value of the final regression equation was 97.8% and all coefficients passed the t-test at a level of significance of 1%, which suggests that the proposed model has a significant explanatory power. Figure 1 shows the scatter graph comparison between model estimates and observations (simulated). The standard error of the model estimates is 3.3 vehicles for an average fleet size of 36 vehicles, indicating a relative average estimation error of 9%. Both the scatter graph and the regression statistics show an exceptionally high quality of fit between the model and the simulated data.

We can also obtain the following quality-of-service equation from Equation 4 that can be used to predict the maximum allowable ride time ratio E with a given fleet:

$$E = \left\{\frac{\lambda_T}{FS} \cdot \left[\tau + \frac{4.62}{V}\left(\frac{A}{\lambda_T \cdot T}\right)^{0.31}\right]\right\}^5 \quad (5)$$

In addition to excess ride time ratio, there are other measures that often need to be considered in evaluating the level of services of a paratransit system (*1*). Examples include response time or advance notification time for reservation, percentage of denials and percentage of subscription trips. Most of these quality-of-service measures can, however, be easily obtained from transit agencies based on their service policies and practices; that is, analytical models are not required to determine the values of these measures.

FACTORS IN MINIMUM FLEET SIZE UNDER IDEALIZED CONDITIONS

Equation 4 was developed based on a set of specified standard conditions and may not hold for applications with deviating conditions. The objective of this section is to analyze the impacts of deviations in conditions on the required minimum fleet size. Five major factors are identified and discussed.

Shape of Service Area

The shape of the geographic area covered by a parantrasit service may not be square as assumed in the ideal conditions. This section examines the possible impact of the shape of the service area on fleet

TABLE 1 A Sample of Simulated Scheduling Data

Case	Service Area (A) Km^2	Trip Density trips/km^2/h	Trip Rate (λ) trips/h	Velocity (V) km/h	Dwell Time (τ) min	Time Window (T) min	Max. Excess Ride Ratio (E)	Fleet Size (FS)
1	400	1.0	400	30	2.0	20	2	115
2	400	1.0	400	30	2.0	40	1.5	109
3	400	1.0	400	30	2.0	40	2.5	89
4	400	1.0	400	30	2.0	20	1.5	125
5	400	1.0	400	30	2.0	20	2.5	106
6	400	1.0	400	30	2.0	20	2	99
7	400	0.8	300	30	2.0	20	2	92
8	400	0.8	300	30	2.0	20	1.5	101
9	400	0.8	300	30	2.0	20	2.5	86
10	400	0.8	300	30	2.0	40	2	72
11	400	0.8	300	30	2.0	40	1.5	84
12	400	0.8	300	30	2.0	40	2.5	69
13	400	0.8	300	20	2.0	40	2	98

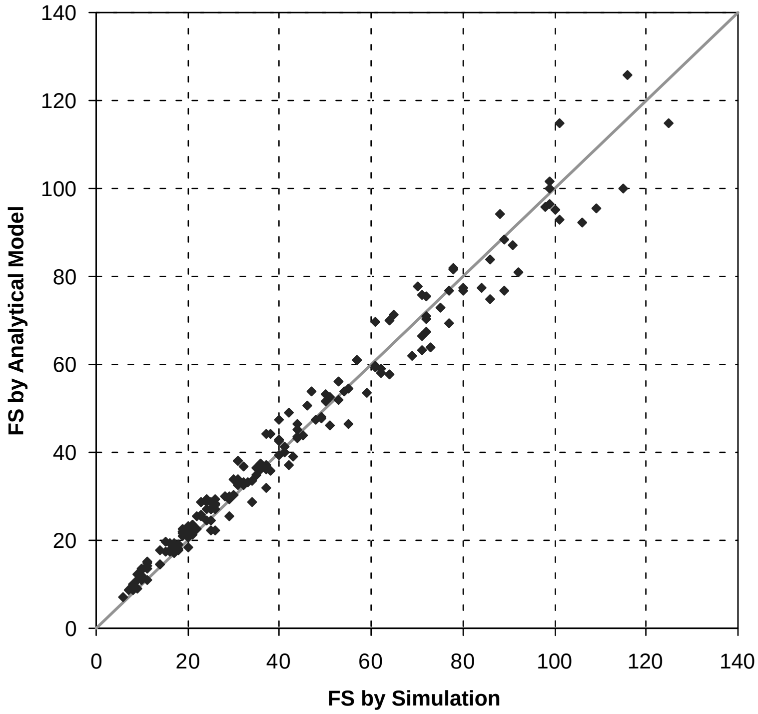

FIGURE 1 Model calibration: analytical model versus simulation.

size. Three rectangular service areas of the same size (100 km^2) but different length-to-width ratios (1.0, 2.0, and 4.0) were investigated. Figure 2 shows the relationship between the increase in fleet size as related to shape ratio under different levels of demand. Two observations can be made from this result. First, there appears to be no direct dependency between fleet size and shape of the service area, which suggests that the variable area size sufficiently explained variation in the minimum required fleet size caused by service area. Second, the variation in fleet size under an elongated service area falls within 10% of the model estimates, which is close to the standard error of the model estimates.

Vehicle Seating Capacity

The ideal operating conditions assume unlimited vehicle capacity; that is, the routing and scheduling process is not restricted by vehicle capacity, but by other constraints such as user ride time and service time window. This assumption should not be very restrictive if the fleet mix is optimally configured for maximum productivity. Nevertheless, because of the complication of multiple seating requirements that often arises in practice and the inability of most scheduling algorithms to optimize the use of available vehicle seats of a mixed fleet, seating capacity could become a factor in influencing the

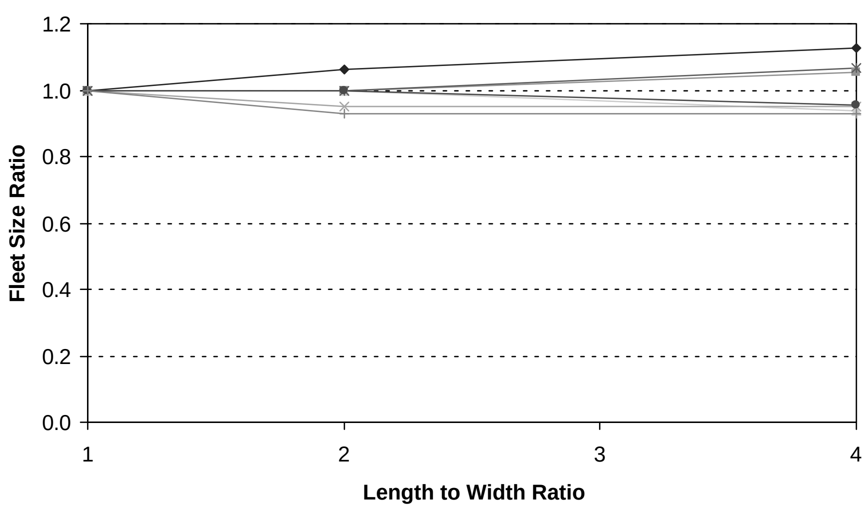

FIGURE 2 Impact of the shape of service area on minimum fleet size.

minimum number of vehicles required in some service conditions. To examine this impact, we constructed a set of cases with a uniform fleet but varying vehicle capacities. Figure 3 presents the simulation results in which curves represent the relationship between the increase in fleet size and vehicle capacity under three levels of travel demand. Other factors are fixed. *V* is 30 km/h, *T* is 30 min, and *E* is 1.0. The results are generally expected. There is an upper bound in vehicle capacity (10 seats in this example), beyond which vehicle capacity is no longer a constraint in the scheduling process. The smaller the service vehicles are, the more vehicles are required to meet the same demand. In the extreme case that only one seat is available for each vehicle, ride sharing and efficient routing become impossible and therefore a much larger fleet is required compared with the cases with vehicles of larger size. In practice, most agencies use a mixed fleet of large and small vehicles, with which the impact of seating capacity on fleet size is likely to be small.

Trip Clustering

In practical operating conditions, trips are often clustered in both space and time. Spatial clustering is reflected in the fact that trip ends are usually concentrated at certain activity centers such as hospitals, shopping malls, and downtown. As a result, a lot of trips are many-to-one or many-to-few trips for which more compact and efficient routes could be developed compared with the cases of purely random, many-to-many trip distribution that is considered in the ideal situation. The number of vehicles required to service clustered trips should therefore be lower than what would be needed for the same number of trips that are randomly distributed. The same is true for clustering of desired pickup and delivery times owing to the schedule constraints of our daily activities.

To quantify the impact of trip clustering on fleet requirements, we simulated a set of demand patterns with varying percentage of clustered trips and numbers of spatial clustering centers (SC) and temporal clustering points (TC). We consider a square service area of 10×10 km with a set of combinations in SC and TC, and the percentage of trips that has at least one trip end located at a SC point. The location of each SC and the time of each TC are generated randomly. To limit the variation, we assumed clustered trips are distributed between the SC and TC at an equal probability. The other system parameters are fixed. The time window (*T*) is 30 min, maximum excess ride (*E*) is 1.0, and demand density is 1.0 trips/km^2/h. Figure 4 shows the fleet size ratio as a function of the percentage of trips clustered for combinations of SC and TC. Two observations can be made. First, as expected, trip clustering did reduce fleet requirements and the reduction increased as the degree of clustering increased. Second, unless the percentage of trips clustered is high (e.g., more than 50%) and at the same time most of them have the same desired time (e.g., 1 TC), the reduction in fleet size owing to trip clustering should be small.

It should be pointed out that the effect of trip clustering on minimum fleet size also depends on the capacity of each available vehicle. For example, spatial trip clustering would have less of an effect on reducing fleet size if a small fleet rather than a large fleet of vehicles is used. However, under the assumption that fleet mix is always optimized before fleet requirement is analyzed (i.e., there is always a sufficient number of large vehicles in a fleet), the effect of such interactions is expected to be small.

Scheduling Method and Algorithm

As discussed earlier, how trips are scheduled and what scheduling algorithm is used may also have some impacts on the number of vehicles needed for a given service condition. This section shows the possible variation in minimum fleet size caused by differences in the scheduling algorithm. The variation was simulated using different scheduling algorithms and settings available in FirstWin, with the aim of mimicking scheduling methods and algorithms of varying degrees of quality that could be seen in practice.

Figure 5 shows the calculated fleet sizes under various configurations of the scheduling algorithms in FirstWin for a problem involving 300 trips randomly distributed over a 10×10 km area. There were striking differences in fleet size requirement among different algorithms and parameter settings, ranging from 5% to more than 35% of an increase compared with those of the base case, Scheduling Method 1, which was also used for the proposed model.

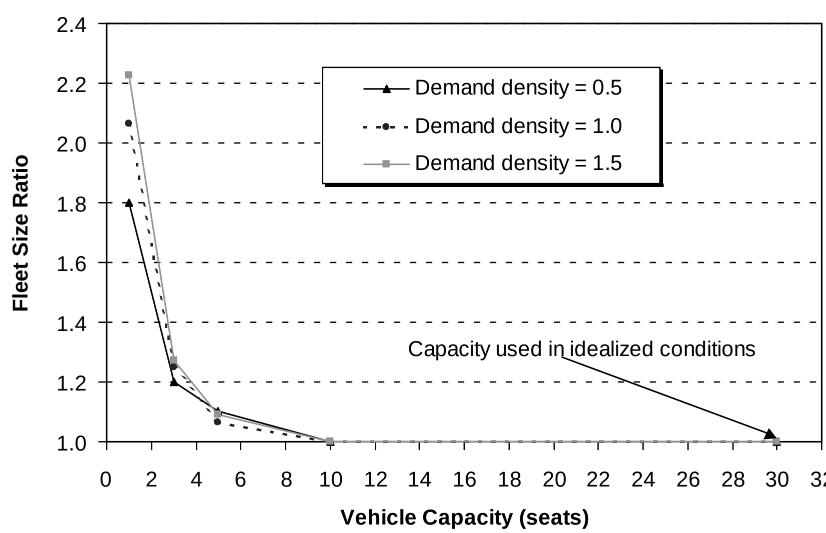

FIGURE 3 Impact of vehicle seating capacity on minimum fleet size.

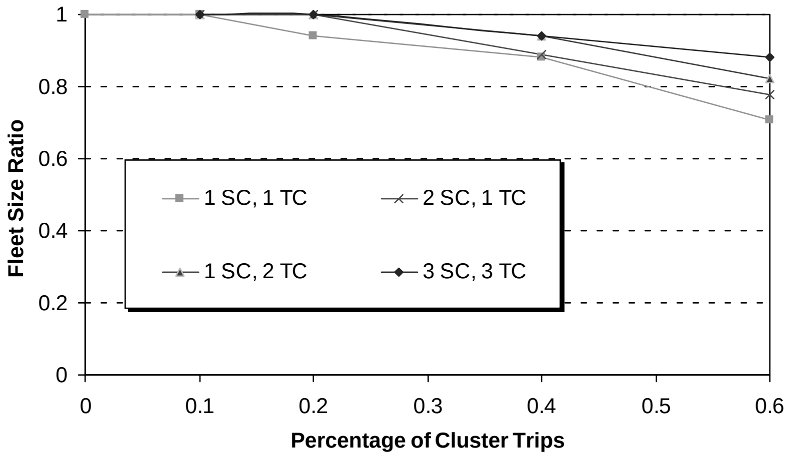

FIGURE 4 Impact of trip clustering on fleet sizing.

These differences are much larger than the estimation errors of the analytical model. This result further supports our argument that using a simulation model may lead to inconsistency.

EXAMPLE OF AN APPLICATION

The objective of this section is to demonstrate the validity of the proposed model using a real-life example and how it can be applied to predict fleet requirements in practical situations. The example consists of a weekday service covered by the Disabled Adult Transportation System (DATS) in the city of Edmonton, Alberta, Canada. A total of 620 trips consisting of both ambulatory and wheelchair trips during the period of 8:00 to 11:00 a.m. were extracted for analysis. A fleet of vehicles with capacities ranging from [4, 0] to [13, 8], in which the first number represents ambulatory seats and the second number represents wheelchair seats, is available to provide the service. The road network includes all arterial streets and freeways within the service area. The following steps were used to generate model estimates on minimum fleet size:

1. Specify user service constraints (T and E). A service time window (T) of 0.5 h and a maximum excess ride ratio (E) of 1.0—that is, 100% excess ride time—were used.

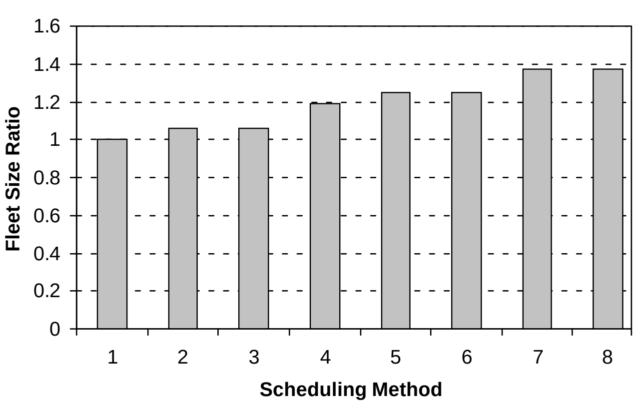

FIGURE 5 Impact of scheduling method on fleet size.

2. Determine peak period demand rate. First plot trip distribution during the 3-h period was done with an interval size equal to half of the service time window (T), that is, $T/2 = 15$ min, as shown in Figure 6. The two consecutive intervals with the highest combined trip rate were identified as the peak demand period. In this example, the peak period is 9:00 to 9:30 a.m., with a peak trip rate of 370 trips/h.
3. Determine average total boarding and alighting time (τ). This parameter was obtained directly from the trip database. The average total boarding and alighting time is 3.57 min/60 = 0.060 h.
4. Determine service area (A). The service area is determined based on the coordinates of the pickup and delivery points of all 620 trips. Figure 7 shows a map of the trip distribution over a two-dimensional space. The service area (A) is determined by identifying the smallest rectangular frame that encompasses all the trip points. In this example, it is approximately equal to 561 km^2, or 416 km^2 if the two outliers are excluded (frame with a thin line). This process may be automated on the basis of trip coordinates and the orientation of the service area.
5. Determine average travel speed (V). An average speed of 30 km/h was used. In practice, this parameter could be calibrated from daily operating records.
6. Calculate fleet size using Equation 4 (FS):

$$FS = \frac{\lambda_T}{E^{0.20}} \cdot \left[\tau + \frac{4.62}{V}\left(\frac{A}{\lambda_T \cdot T}\right)^{0.31}\right]$$

$$= \frac{370}{1^{0.20}} \times \left[0.06 + \frac{4.62}{30} \times \left(\frac{561}{370 \times 0.5}\right)^{0.31}\right] = 103 \text{ vehicles}$$

If the smaller area ($A = 415$ km^2) were used, then the minimum fleet size would be 96 vehicles. To validate this result, FirstWin was also used to schedule the same trips, which resulted in 97 vehicles assigned with trips. As can be seen, the model estimates are very close to the simulation result even without any adjustments. There could be two possible explanations to this close match: either the effects of the prevailing conditions were small, or the effects of different factors had canceled out one another. For example, limitation in vehicle capacity would increase the fleet size while trip clustering would induce more efficient routes and thus demand a smaller

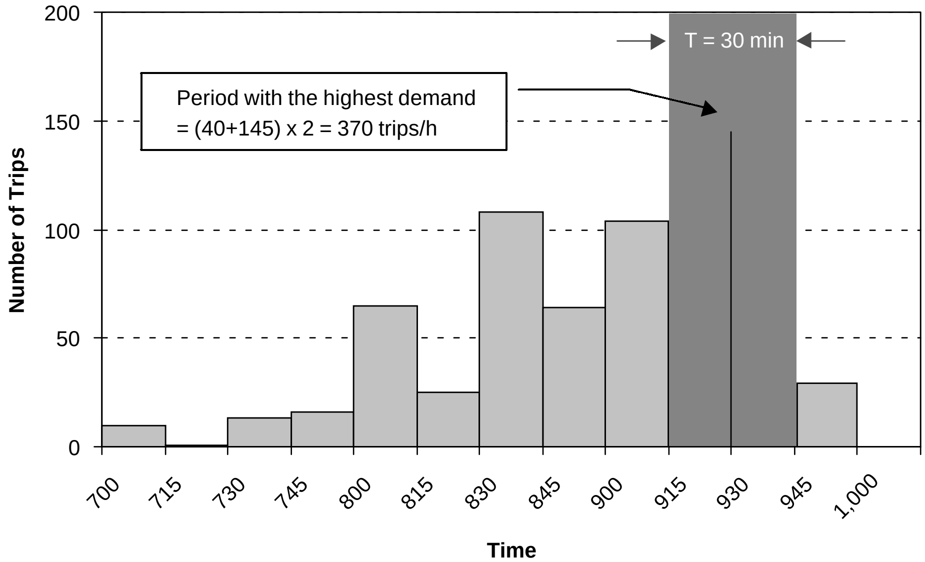

FIGURE 6 Trip: temporal distribution and peak demand.

number of vehicles. The city of Edmonton DATS is currently using a fleet of 106 vehicles, which is larger than the estimates from the analytical model and simulation. This discrepancy between that number of vehicles and the estimate might be attributed to the constraint by the length in vehicle shift, as enforced in the city's operations.

CONCLUDING REMARKS

The planning and design of paratransit service systems call for methodologies and tools that can be used by paratransit planners and service providers to analyze resource requirements, system capacity, and level of service for specific operating conditions. This paper

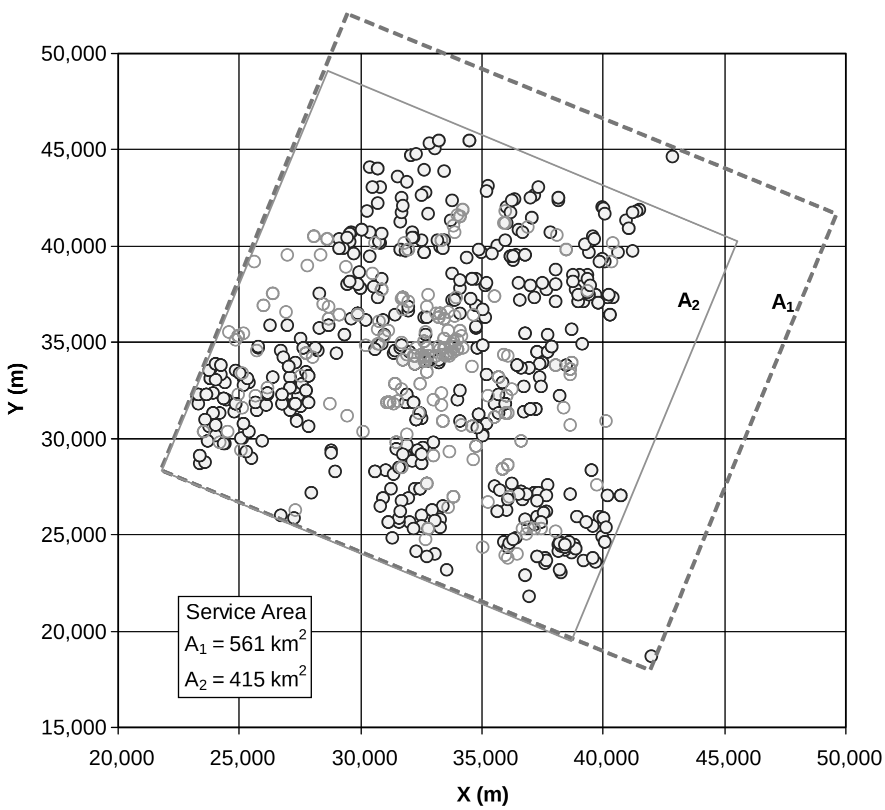

FIGURE 7 Trip: spatial distribution and service area.

presents an approximate model that allows analytical investigation of changes in fleet size and quality of service when parameters associated with a service system are varied, without the need to resort to simulation. Because of its proper model structure, the proposed model was found to fit the simulated data particularly well. Simulation data also indicated that although the model was calibrated for ideal operating conditions, it could be conveniently adjusted to more accurately reflect the prevailing conditions.

The significance of the presented work is perhaps in providing a start point toward developing an analysis procedure that could be integrated into the future edition of the TCQSM for capacity and level-of-service analysis of paratransit services. Such an approach would be consistent with the convention adopted by the TCQSM and the HCM, in which all analysis procedures are based on analytical models instead of simulation. However, further research is needed to quantify the exact impacts of those conditions that do not match the idealized conditions, and to develop corresponding adjustment factors based on both simulated and field data.

ACKNOWLEDGMENTS

This research was supported by the Natural Sciences and Engineering Research Council of Canada. Thanks go to the Disabled Adult Transportation System of Edmonton, Alberta, for providing the data on which a case study was performed.

REFERENCES

1. *Transit Capacity and Quality of Service Manual* (1st ed.). TCRP Web Document 6. January 1999. www4.nas.edu/trb/crp.nsf/All+Projects/TCRP+A-15. Accessed June 2003.
2. Wilson, N. H. M., J. M. Sussman, B. T. Higonnet, and L. A. Goodman. Simulation of a Computer-Aided Routing System (CARS). In *Highway Research Record 318,* HRB, National Research Council, Washington, D.C., 1970, pp. 66–76.
3. Fu, L. A Simulation Model for Evaluating Intelligent Paratransit Systems. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1760,* TRB, National Research Council, Washington, D.C, 2002, pp. 93–99.
4. *TCRP Project B-23: Resource Requirements for Demand-Responsive Transportation Services,* in progress.
5. Wilson, N. H. M., J. M. Sussman, H. K. Wong, and T. Higonnet. *Scheduling Algorithms for a Dial-a-Ride System.* Massachusetts Institute of Technology, Cambridge, 1971.
6. Daganzo, C. F. An Approximate Analytical Model for Many-to-Many Demand-Responsive Transportation Systems, *Transportation Research,* Vol. 12, 1978, pp. 325–333.
7. Stein, D. M. Scheduling Dial-A-Ride Transportation Systems. *Transportation Science,* Vol. 12, No. 3, 1978, pp. 232–249.
8. *Highway Capacity Manual.* TRB, National Research Council, Washington, D.C., 2000.
9. Fu, L. Scheduling Dial-a-Ride Paratransit Under Time-Varying Stochastic Congestion. *Transportation Research B,* Vol. 36, 6, July 2002, pp. 485–506.
10. SPSS, Inc. www.spss.com. Accessed Jan. 2002.

Publication of this paper sponsored by Committee on Paratransit.