Estimating Passenger Miles, Origin–Destination Patterns, and Loads with Location-Stamped Farebox Data

David S. Navick and Peter G. Furth

Integrating an electronic farebox with a location system can provide location-stamped records of passenger boardings, a valuable source of information on passenger travel patterns. However, this information is of small value unless the pattern of passenger alightings can also be determined, since most relevant measures of interest—passenger loads, passenger miles, and origin–destination (O-D) patterns—require a knowledge or at least estimate of passenger alightings by stop. The assumption of symmetry—that the pattern of passenger alightings in one direction mirrors the daily boardings pattern in the opposite direction—is explored. Estimation methods using this assumption are tested at the trip, route, and system levels using a full-day's set of on–off counts on five Los Angeles area routes. Tests at the route level indicate that although perfect symmetry does not exist, patterns are substantially similar on many routes. Based on the Los Angeles data, it can be found that systemwide estimates of passenger miles made using this method satisfy U.S. National Transit Database precision requirements; however, this finding should be confirmed using data from other cities. Proposed and tested, with a small amount of success, is a method for estimating trip-level O-D patterns using location-stamped farebox data based on the symmetry principle and a gravity model. Location-stamped farebox data can also be used to estimate passenger loads in real time to support control measures such as conditional priority at traffic signals without requiring automatic passenger counters.

For both effective planning and operation control, transit agencies need not just information on how many passengers they are carrying, but on where these passengers boarded and alighted. This information is used to estimate systemwide passenger miles, a robust measure of system use and a reporting requirement of the U.S. National Transit Database (*1*), to determine passenger loads along a route, useful for both schedule analysis and for real-time control, such as giving transit priority at a traffic signal based on vehicle occupancy (*2*), and to determine route-level load profiles and origin–destination (O-D) matrices for analyzing routing and scheduling options such as short-turn service, a small amount of stop service, and express service (*3*).

Historically, gathering location information on passenger boardings and alightings has required labor-intensive ride checks (manual on and off counts by stop) or automatic passenger counters (APCs). However, APCs are expensive and their widespread adoption is still many years off. We do not know of a recent, complete count of APC users; however, a 1998 survey (*4*), targeted in part at transit agencies that were known to be APC users, obtained information on eight North American transit agencies that were regular APC users and another five who were in a demonstration phase. Because of their cost, APCs are usually installed on only a sample of the fleet, typically 6% to 10%, which is rotated through the system to provide a sample of passenger on–off data, an arrangement that cannot support real-time load monitoring.

Electronic fareboxes with location-stamped transactional data offer another alternative to getting passenger location data. "Transactional data" means that a record is kept of each farebox transaction—essentially, each boarding—rather than keeping records of accumulated counts and revenue as in the traditional electronic farebox. The latest generation of fareboxes incorporates a transactional data system (*5*). However, it is also possible for a traditional farebox to supply the signals necessary for an external processor to create transactional records. Transactional records are typically time stamped and include information about fare type and route and trip identifiers. "Location stamped" means the records also report where the boarding took place, that is, the most recent stop at which the door opened.

Providing a location stamp requires an automatic vehicle location (AVL) system and its integration with the farebox. Vehicle location systems, using technologies including geographic positioning systems, dead reckoning, and wayside beacons, are more widely used than APCs. A 1997 report (*6*) lists 19 operational North American systems and describes a surging demand: 70 North American transit agencies had implemented or received funds to implement AVL, with 20 receiving funding in the year before the report. When implemented, AVL, like electronic fareboxes, is typically installed fleetwide and therefore is not subject to the typical sampling limitations of APC.

Integration of AVL with transactional fareboxes is technologically feasible and anticipated in plans for the "smart bus" (*7, 8*). Three possible modes of integration have been described (*5*):

- A transactional farebox accepts inputs from the location system and uses them to stamp its records;
- The location system accepts inputs from the farebox and creates its own transactional records; and
- The location system, farebox, and other devices such as the destination sign and an operator's control head are interconnected by means of a local area network known as a vehicle area network (VAN) featuring a common data bus and communication standards. In this last arrangement, the computer storing the time-stamped records could be the AVL computer, the farebox computer, or another onboard computer connected to VAN.

D. S. Navick, College of Engineering 230 SN; and P. G. Furth, Department of Civil and Environmental Engineering, 400 SN; Northeastern University, 360 Huntington Avenue, Boston, MA 02115.

The first and second alternatives listed above have two advantages—they do not require a transactional farebox system, and they better protect the integrity of the farebox data, which is important for revenue control. As both transit agencies and suppliers contemplate farebox–location system integration, it is important to explore its possible benefits in order to guide product development and procurement specifications.

Allowing that the technological challenges of integrating farebox and location data can be overcome, another challenge remains. Although ride checks and APC count both where passengers board and alight, fareboxes can at best record only where passengers board. The thesis of this discussion is that the symmetry of travel patterns exhibited by most transit lines, in which passenger boarding and alighting patterns in one direction mirror those of the other direction on a daily basis, makes it possible to estimate the alightings pattern from location-stamped boardings data with a reasonable degree of accuracy.

First discussed is whether the full distribution of alightings mirrors the boardings pattern in the opposite direction; data from full-day on–off counts on five Los Angeles bus routes are used. Then the accuracy of passenger-miles estimates are explored and a summary of the boardings and alightings profile is provided, which is made from location-stamped farebox data using the symmetry assumption at three levels:

1. Route level, for which passenger-miles estimates serve as a measure of monitoring route productivity;
2. Systemwide, for National Transit Database reporting; and
3. Trip or period level, which is appropriate for analyses of routing and scheduling options and estimates for which involve trip distribution modeling.

Finally, discussed are possible uses of location-stamped farebox data for real-time control.

BOARDINGS–ALIGHTINGS SYMMETRY

Although an electronic farebox integrated with a location system can register boardings by stop, it cannot register alightings. Therefore the following symmetry assumption is proposed: the boarding pattern for a route in one direction is equivalent to the alighting pattern in the opposite direction over the course of an entire day. This assumption is based on the general behavior of return travel. Roughly speaking, passengers get on in one direction where they got off in the other direction.

To test this assumption, Los Angeles Metro provided us with ride checks data for five routes, 10/11, 60, 105, 236/240, and 418. The data set for each route consists of on–off counts at every stop on every trip on a single day. [Detailed information on these routes and on other aspects of this discussion can be found in the doctoral dissertation on which it is based (*9*).]

The analysis was complicated by the fact that all of these routes, as is common in Los Angeles, combine a trunk (a set of stops served on almost every trip) with a number of route deviations called branches. An example of a branch is a small deviation from the trunk that passes by an elementary school serving four off-trunk stops on only the few trips matching the school's opening and closing times. These branches complicate the process of comparing the full-day boarding and alighting patterns. Another complication is the lack of a one-to-one correspondence between stops in opposite directions even on the trunk. A route may have a greater number of stops in one direction or may not stop at the same intersection in both directions.

To accommodate these variations, an "ideal route" was developed for each study route in each direction following the trunk route. Where stops in the subject direction do not mirror stops in the opposite direction, estimating alightings in the subject direction from boardings in the opposite direction involved allocating the opposite direction boardings to the stop locations in the subject direction in proportion to segment length. For example, to estimate the alightings pattern inbound from the boardings pattern outbound, if an outbound stop lay 40% of the way between two inbound stops, 60% of its boardings were allocated to the nearer inbound stop and 40% to the farther inbound stop. Boardings and alightings at branch stops were likewise allocated to the appropriate stops on the ideal route.

The test for similarity in boarding patterns with alighting patterns in the opposite direction was conducted using the Kolmogorov–Smirnov test, a test that compares whether two distributions are equivalent based on the maximum distance between their cumulative distribution functions, the test statistic. The cumulative distributions of boardings and alightings are meaningful measures, because their difference at any given point is passenger volume passing that point.

In a test comparing two empirical distributions compiled from data (boardings counts in one direction compared with alightings counts in the other), the test statistic can be compared with the critical value at the .05 significance level, which is given as (*10*, pp. 221–222)

$$KS_{.05} = 1.36\sqrt{\frac{(n_1 + n_2)}{n_1 n_2}} \tag{1}$$

where n_1 and n_2 are the number of passengers counted in the two directions. There were 10 tests, one for each route in each direction. As an example, the test for Route 10/11 eastbound is shown in Figure 1, comparing the cumulative distribution of eastbound boardings with westbound alightings, with the westbound alighting allocated to the eastbound stops.

Listed in Table 1 for each route direction analyzed is the test statistic KS_{max}, the greatest absolute deviation between the cumulative distribution of boardings in the subject direction and alightings in the opposite direction. Comparing against the critical value, the test fails for 9 of the 10 cases tested, indicating an absence of perfect symmetry. The sample sizes are so large—on the order of 10,000 observations per route direction—that a near-perfect fit is required to pass the test. This result is reasonable; symmetry is a general pattern, but not a law. Not all passengers take return trips that mirror their first trip. Also, the modeling approximation necessary to fit stops in the two directions to each other and to fit branches to a trunk is bound to lead to some errors.

Although symmetry is not perfect, it may nevertheless be a good approximation and therefore a good means of estimation. Therefore, also tested was whether opposing boarding and alighting are substantially different using the metric that a maximum variation in cumulative distribution of less than 5% shows no meaningful difference, a maximum of variation 5% to 10% a small difference, a maximum variation of 10% to 15% a mild difference, and a maximum variation greater than 15% a substantial difference. It was found that 5 of the 10 cases showed no meaningful difference, 2 showed a small difference, 1 a mild difference, and only 2—the two directions of Route 418—showed a substantial difference. An examination of Route 418 reveals that its lack of symmetry can be explained by the fact that it shares a long section with other routes

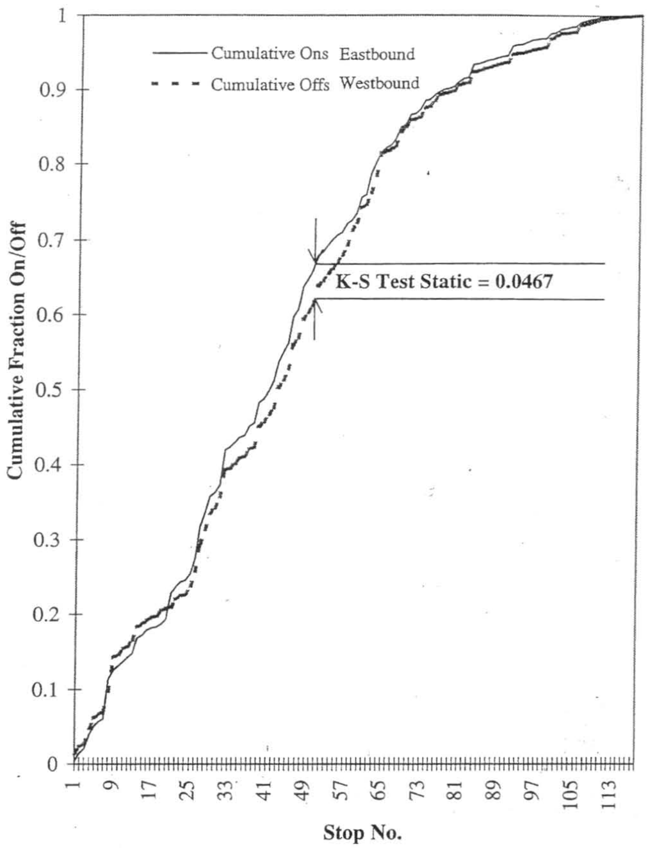

FIGURE 1 Kolmogorov–Smirnov test of symmetry for Los Angeles Route 10/11 eastbound.

from downtown Los Angeles to the San Fernando Valley via the Golden State Freeway. Passengers whose trips begin and end on the common section, or who transfer to the route on its common section, can use one route in one direction and another on their return trip, depending on which comes first or is more likely to afford a seat. Much of Route 418's ridership consists of passengers transferring from other routes, for which a small change in arrival time can determine whether a busload of transferring passengers will take one route or another along the Golden State Freeway.

In a related study, Furth (*11*) tested for symmetry on two light rail lines in Pittsburgh. In that study, the boardings and alightings data came from a small set of sampled trips. Rather than testing the entire distribution, that study tested for similarity in means—whether the centroid of the alightings distribution in one direction equals the centroid of the boardings distribution in the opposite direction. The boarding centroid, C_b, is a weighted (by boardings at each stop, b_i) average of the distance to each stop, D_i, from the beginning of the route:

$$C_b = \frac{\sum_i b_i D_i}{b.} \qquad (2)$$

where $b.$ represents total boardings. The alighting centroid is found analogously. In three of the four cases (each route direction affords a test), the centroids were statistically equivalent. However, there was a large discrepancy in one route direction because of the large number of riders who transfer from buses to this route for the last 3 km of their trip to downtown Pittsburgh in the morning but use a different route in the evening.

In conclusion, symmetry can be a valuable tool for estimating the alightings pattern on many routes. However, some routes depart markedly from a pattern of symmetry. Some checking against on–off data is appropriate before the symmetry assumption can be taken with confidence to estimate the alightings pattern on a given route.

ESTIMATING DAILY ROUTE-LEVEL PASSENGER MILES

As Furth (*11*) pointed out, average passenger trip length on a route is the distance between its boarding centroid and its alighting centroid, and passenger miles are simply the average passenger trip length expanded by total boardings. This relationship is illustrated in Figure 2. If the symmetry assumption is valid, the boardings centroid in one direction equals the alightings centroid in the opposite direction. A corollary, then, of the symmetry assumption is that average passenger trip length is the same in both directions, and is the difference between the boardings centroid in the one direction and the boardings centroid in the opposite direction. Multiplying this difference by total boardings will yield total passenger miles.

Therefore, using the symmetry assumption, average trip length and total passenger miles for any route day can be estimated from location-stamped boardings data—data that will not be subject to sampling error because, assuming every bus has a farebox integrated with a location system, data will be collected on every operated trip.

TABLE 1 Symmetry Hypothesis Test Results

Route	Direction	Total Boarders, n_1	Total Alighters, n_2	KS_{max}	$KS_{0.05}$	Accept Symmetry Hypothesis?
LA 10	EB	9,129	9,642	0.047	0.020	NO
LA 10	WB	9,729	9,141	0.068	0.020	NO
LA 60	NB	12,393	13,134	0.045	0.017	NO
LA 60	SB	13,134	12,393	0.106	0.017	NO
LA 105	NB	8,657	8,089	0.024	0.021	NO
LA 105	SB	8,386	8,198	0.033	0.021	NO
LA 240	NB	1,828	1,799	0.066	0.045	NO
LA 240	**SB**	**1,799**	**1,828**	**0.026**	**0.045**	**YES**
LA 418	EB	492	402	0.199	0.091	NO
LA 418	WB	448	378	0.135	0.095	NO

NOTE: Row in bold indicates success of test, indicating absence of perfect symmetry.

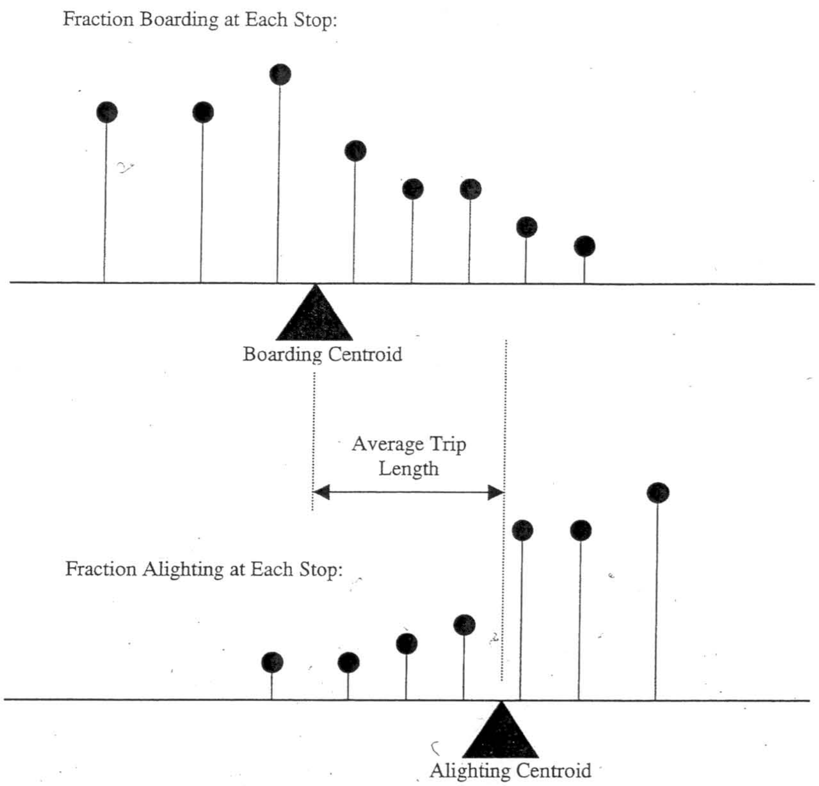

FIGURE 2 Average trip length calculated via centroids.

Another summary of the passenger boarding and alighting patterns is the volume or load profile, which is simply the difference between the cumulative boardings and cumulative alightings. The volume profile is a valuable tool for visualizing service utilization along the route and for suggesting possible scheduling changes such as short-turn service to improve operation efficiency (*2*). Passenger miles can also be determined from the volume profile by multiplying the volume on each segment by segment length and summing over all segments, as described elsewhere (*1*). Using the cumulative boardings distribution in the opposite direction to estimate cumulative alightings (scaling the opposite direction's cumulative boardings distribution to match total boardings in the subject direction) and taking the difference, one can estimate a route's volume profile in each direction from location-stamped boardings counts. Another corollary of the symmetry assumption is that the daily volume profiles in the two directions are the same, except for a scalar to reflect a possible difference in total passengers carried by direction.

Data from the five Los Angeles bus routes were used to compare route–day level passenger miles as estimated using the symmetry estimator with location-stamped boardings data against passenger miles calculated directly from on–off counts. Estimating passenger miles using the symmetry method required not only fitting the data to the ideal routes previously described, but also assigning to each stop on the ideal route a distance from the start of the route, which is complicated because the distance of a stop from the start of the route measured along the actual path taken can vary depending on which branch a trip is following. The ideal route uses distances for each stop that reflect the average distance traveled:

$$d_{k*} = \frac{\sum_b d_{kb} q_{kb}}{\sum_b q_b} \qquad (3)$$

where

d_{k*} = distance from the start of the route to stop k on the ideal route,
d_{kb} = distance from the start of the route to stop k on branch b (here the trunk is treated as a branch), and
q_b = number of bus trips per day operated on branch b.

Any change in distance between the branch endpoints is then allocated proportionally to the trunk segments between those endpoints.

The passenger-mile comparisons are shown in Table 2. Both overprediction and underprediction are found with all absolute errors for the routes less than 9% of the time. Overall there was an overprediction error of 3.2%. The relative standard deviation of between-routes estimation error was 9.9%. To two significant digits, these summary statistics apply equally to the 10 route-direction totals and to the 5 route totals. By coincidence, the average prediction error of the symmetry estimator was also 3.2% on the two Pittsburgh light rail lines analyzed elsewhere (*11*).

A *t*-test supports the hypothesis that there is no underlying overestimation or underestimation bias. Given the small number of routes in the sample and the large between-routes variation, that overall level of overprediction found is statistically insignificant—it can be explained by the expected level of sampling variation. This finding supports the notion of using an estimate aggregated over many routes, the subject of the next section.

TABLE 2 Daily Passenger-Mile Estimate Errors for Los Angeles

Route	Direction	Actual Passenger-Miles	Predicted Passenger-Miles	Difference	Relative Error
LA 10/11	EB	24,605	26,446	1,841	7.5%
	WB	27,586	28,272	686	2.5%
	Both	**52,191**	**54,718**	**2,527**	**4.8%**
LA 60	NB	47,757	51,541	3,784	7.9%
	SB	49,479	53,862	4,383	8.9%
	Both	**97,236**	**105,403**	**8,167**	**8.4%**
LA 105	NB	22,594	20,659	-1,935	-8.6%
	SB	20,220	19,719	-501	-2.5%
	Both	**42,814**	**40,378**	**-2,436**	**-5.7%**
LA 236/240	NB	4,384	4,084	-300	-6.8%
	SB	4,078	3,866	-212	-5.2%
	Both	**8,462**	**7,950**	**-512**	**-6.1%**
LA 418	EB	6,180	5,667	-513	-8.3%
	WB	5,977	5,612	-365	-6.1%
	Both	**12,157**	**11,279**	**-878**	**-7.2%**
Mean		42,572	43,946	1,374	3.2%
Standard Deviation				4,202	9.9%
Standard Error				1,879	4.4%

ESTIMATING SYSTEM PASSENGER MILES

In the United States, most bus systems are required to report annual estimates of systemwide passenger miles to the National Transit Database. A standard sampling plan may be used, requiring ride checks on 549 or more randomly sampled trips (*1*). Alternatively, any other method may be used that meets the specified accuracy level of ±10% precision at the 95% confidence level. This section examines whether the symmetry estimator, applied to all routes and aggregated systemwide, can be expected to provide that level of accuracy.

When aggregating over all routes, sampling variation disappears because all days, trips, and routes are sampled. Likewise, variation among routes, large as it is, disappears because all routes are sampled. Variation in the final estimate, then, is based only on the estimator's overall bias (overprediction or underprediction) for the subject city.

As explained in the previous section, the Los Angeles data from five routes are consistent with the hypothesis of no overall bias. However, with data from only five bus routes in one city, more work is clearly needed to confirm this hypothesis or, better still, to estimate the range of bias expected from one city to another. Nevertheless, even if the estimator's bias for a given city is as great as the observed Los Angeles bias of 3.2%, the resulting precision at the 95% confidence level is ±6.4%, which is still well within National Transit Database specifications. (It is worth clarifying that although intentional bias is certainly unacceptable, statistical bias—particularly when the analyst does not know whether the bias is positive or negative—is a routine part of many often-used estimators. If the magnitude and direction of the bias become known through further analysis, it is then a simple matter to adjust the estimator to compensate.)

The findings of this data analysis suggest that location-stamped farebox data may be a suitable, stand-alone source for estimating passenger miles to the level of accuracy required by the National Transit Database. This finding alone should be enough to spur farebox-location system integration. However, analysis on more bus routes and more cities should be done to confirm that the level of modeling and estimation error expected in using this method lies within the National Transit Database standard.

ESTIMATING TRIP-LEVEL AND PERIOD-LEVEL O-D MATRICES

As mentioned earlier, trip-level and period-level route O-D matrices are useful for analyzing routing and scheduling options such as short-turn service, expressing, and route splitting. A partial aggregation of the trip O-D matrix, the load profile (passenger load on each segment), is useful for analyzing schedules and assessing passenger crowding. The question is how to make use of the symmetry principle to estimate a trip-level and period-level route O-D matrix using location-stamped farebox data. This problem differs from the daily O-D matrix estimation problem because return travel, the basis for symmetry, generally occurs in the time frame of a day, not a scheduling period such as morning peak, and certainly not as a trip. Nevertheless, symmetry can still be exploited. If a full-day's worth of trip O-D matrices is estimated, the symmetry assumption can provide a constraint on total daily alightings by stop, shaping the estimated alightings profile of every trip.

A trip-level O-D matrix, t_{iju}, indicates the number of passengers boardings at each stop i and alighting at each stop j for a given trip u. It can be visualized in two dimensions as a stack of trip-level O-D matrices, with each trip's square O-D matrix stacked above the next. In this representation, each row total—boardings at a given stop for a given trip—will be known from location-stamped farebox data. Using the symmetry assumption, each column total—alightings at a given stop summed over all the trips of a day—will likewise be known, using the assumption of symmetry applied at the daily level. Given an initial or seed matrix, the stack of O-D matrices was estimated using the biproportional method, also called "iterative proportional fit," by iteratively scaling the rows and columns to match their given totals (*12*). The resulting estimate has the form

$$t_{iju} = s_{ij} R_{iu} C_j \tag{4}$$

where

$s_{ij}(d_{ij})$ = the seed, the estimated propensity of making a trip of a distance d_{ij} = distance from stop i to j;

R_{iu} = internally computed balancing factor that represents the popularity of origin stop i on trip u; and
C_j = internally computed balancing factor that represents the popularity of destination stop j.

This model is equivalent to a doubly constrained gravity model for trip distribution.

Following Navick and Furth (*13*), a distance-based propensity function is used for the seed matrix. Although that study used a gamma distribution as the propensity function, for this research a new form for the propensity function was developed tailored to the behavioral characteristics of urban bus travel. For shorter trips, there is competition from walking, and so travel propensity increases with distance. An exponential decay model, decreasing with distance, can be used to model the propensity to walk; its complement is used to model the propensity to take the bus when trips are short. However, as trips become longer, propensity to travel by bus is expected to fall with travel distance, modeled with another exponential decay function. The propensity function is the product of these two effects:

$$s_{ij} = [1 - \exp(-\alpha \cdot d_{ij})] \cdot \exp(-\beta \cdot d_{ij}) \quad (5)$$

in which α represents resistance to walking longer distances, and β is the resistance to riding the bus for longer distances. This propensity function is illustrated in Figure 3. Moreover, Furth and Navick (*14*) have shown that, when travel is unidirectional and origin and destination totals are constrained, β cannot be identified; that is, every value of β will yield the same estimated matrix. Therefore its value can be set to the convenient value of zero.

The parameter α was estimated from O-D data from Boston area Bus Routes 1, 66, and 77 collected using a "no questions asked" survey (*15*). A maximum-likelihood methodology was used to determine the α that led, after applying the biproportional method, to O-D cell estimates that best fit the Boston data at the trip level. The estimated α was slightly greater than zero, giving the propensity function an initial positive, nearly linear slope.

To assess the accuracy of the estimated trip-level O-D matrices, passenger miles were compared as calculated from the estimated matrices with passenger miles as calculated from the original on–off data. Because the total boardings value is given, this is equivalent to comparing average passenger trip length. In Figure 4 the cumulative distribution is shown of the relative absolute errors for all the trips in 1 day for the five Los Angeles routes analyzed. Median absolute error using the stacked-matrices method was 13%, indicating that the estimation method is rather weak. Errors in period-level estimates were not analyzed, which are more valuable than trip-level estimates for route planning; because they represent an aggregation of trips, estimation errors are likely to be smaller.

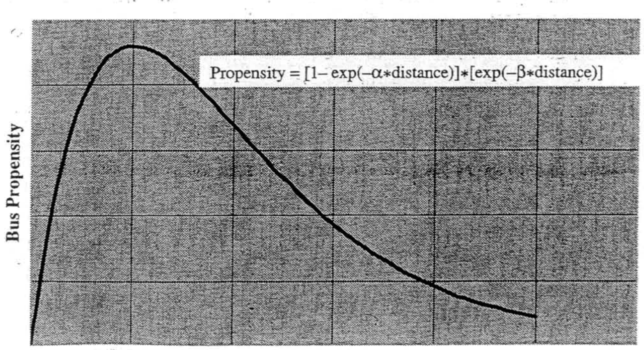

FIGURE 3 Bus travel propensity function.

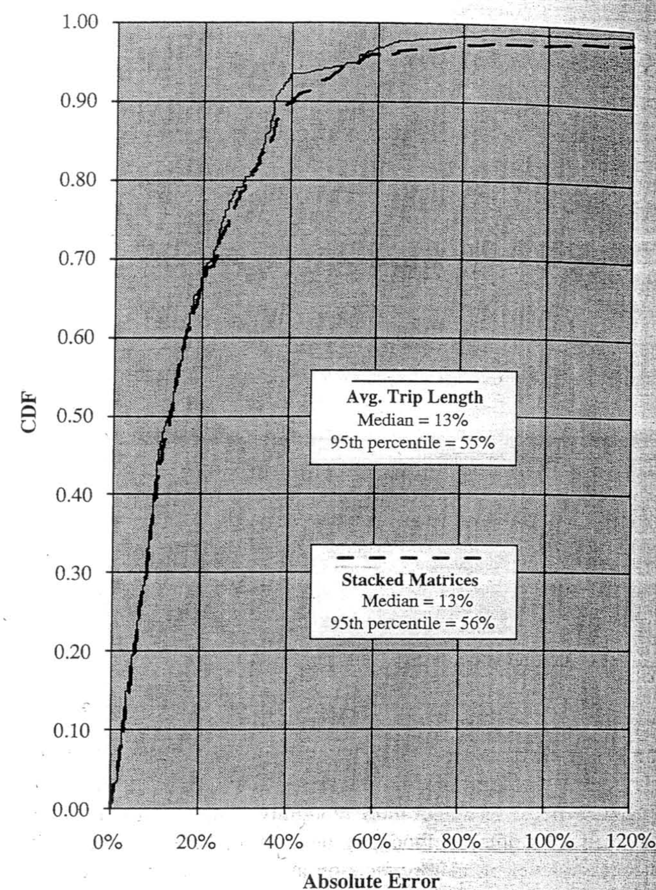

FIGURE 4 Trip-level passenger-mile estimates, showing cumulative distribution function (CDF) relative to absolute errors for 1 day for five Los Angeles routes.

Passenger-miles estimation errors of the stacked matrices method were compared with errors resulting from assuming the same average trip length on every trip of the day. In the latter method, average trip length was determined using the symmetry method applied to a full day's boardings counts by stop. The results were virtually identical to the stacked matrices methodology, suggesting that little information about the trip-level O-D patterns can be deduced from a symmetry-based alightings estimate based on a day's, not a trip's, data.

This finding lends weight to using simplified methods to estimate trip-level passenger travel patterns in real-time monitoring and control applications, for which the full day's data are not yet available. To estimate passenger load on a bus in real time using location-stamped farebox data—say, to determine whether to give the bus priority at a traffic signal, or whether to dispatch a relief bus—one could rely on historical data to estimate for each boarding stop a survival function r_{ij} equal to the fraction of the passengers who boarded at i who are still on board at j. The symmetry method described in this discussion could provide the estimates of the historical O-D data, making manual or APC counts unnecessary. An interesting test for further work would be to compare the accuracy of estimates made using boarding-stop-specific survival functions with a simpler, non-route-specific survival function r_j equal to the fraction of onboard passengers who remain on board as the bus passes j. Estimating this kind of one-dimensional survival function requires historical on–off data (less detail than O-D data). Again, location-stamped farebox data, with the symmetry method, could be used to estimate the historical data needed. Although passenger-load estimates made using survival functions are bound to be rough, they are probably adequate for signal priority control decisions, and can serve as an effective screening device for other methods of operational control.

CONCLUSIONS

Integrating an electronic farebox with a location system in order to provide location-stamped boardings records offers a valuable source of information on passenger travel patterns. On many routes, the assumption of symmetry can be used to estimate the daily passenger alightings pattern from the daily boardings pattern in the opposite direction. With boardings recorded by stop and alightings estimated by stop, valuable travel-pattern measures including passenger miles, load profiles, and O-D matrices can be estimated. Data from five Los Angeles bus routes suggest that systemwide passenger-mile estimates made using this method by aggregating over routes will satisfy U.S. National Transit Database precision requirements; however, this finding should be confirmed using data from other cities. A method is proposed for using location-stamped farebox data to estimate trip-level O-D patterns. These techniques also make it possible to use location-stamped farebox data to monitor passenger loads in real time without requiring APCs.

ACKNOWLEDGMENTS

Research for this paper was supported by the Transit-IDEA program, a cooperative program administered by TRB.

REFERENCES

1. *Sampling Procedures for Obtaining Fixed-Route Bus Operation Data Required Under the Section 15 Reporting System.* FTA Circular 2710.1A. FTA, U.S. Department of Transportation, 1990.
2. Casey, R. F., L. N. Labell, and R. Holstrom. *Advanced Public Transportation Systems: The State-of-the-Art Update '96.* Report DOT-VNTSC-FTA-95-13. FTA, U.S. Department of Transportation, 1996.
3. Furth, P. G., and F. B. Day. Transit Routing and Scheduling Strategies for Heavy Demand Corridors. In *Transportation Research Record 1011,* TRB, National Research Council, Washington, D.C., 1985, pp. 23–26.
4. Boyle, D. K. *TCRP Synthesis of Transit Practice 29: Passenger Counting Technologies and Procedures.* TRB, National Research Council, Washington, D.C., 1998.
5. Furth, P. G. Integration of Fareboxes with Other Electronic Devices on Transit Vehicles. In *Transportation Research Record 1557,* TRB, National Research Council, Washington, D.C., 1996, pp. 21–27.
6. Okunieff, P. A. *TCRP Synthesis of Transit Practice 24: AVL Systems for Bus Transit.* TRB, National Research Council, Washington, D.C., 1997.
7. Schiavone, J. J. *TCRP Report 43: Understanding and Applying Advanced On-Board Bus Electronics.* TRB, National Research Council, Washington, D.C., 1999.
8. Furth, P. G. *Transit–IDEA Program Project Final Report: Management Information Benefits of Integrating Electronic Fareboxes with Other On-Board Equipment.* TRB, National Research Council, Washington, D.C., Sept. 1995.
9. Navick, D. S. *A Model for Estimating Load Profiles and Passenger-Miles from Location-Stamped Transactional Farebox Data.* Ph.D. dissertation. Northeastern University, Boston, Mass., 1997.
10. Hoover, S. V., and R. F. Perry. *Simulation: A Problem-Solving Approach.* Addison-Wesley Publishing Company, Inc., Reading, Mass., 1990.
11. Furth, P. G. Innovative Sampling Plans for Estimating Transit Passenger-Kilometers. In Transportation Research Record 1618, TRB, National Research Council, Washington, D.C., 1998, pp. 87–95.
12. Ben-Akiva, M. E., P. P. Macke, and P. S. Hsu. Alternative Methods to Estimate Route-Level Trip Tables and Expand On-Board Surveys. In *Transportation Research Record 1037,* TRB, National Research Council, Washington, D.C., 1985, pp. 1–11.
13. Navick, D. S., and P. G. Furth. Distance-Based Model for Estimating a Bus Route Origin-Destination Matrix. In *Transportation Research Record 1433,* TRB, National Research Council, Washington, D.C., 1994, pp. 16–23.
14. Furth, P. G., and D. S. Navick. Bus Route O-D Matrix Generation: Relationship Between Biproportional and Recursive Methods. In *Transportation Research Record 1338,* TRB, National Research Council, Washington, D.C., 1992, pp. 14–21.
15. Stopher, P. R., L. Shilltio, D. T. Grober, and H. M. A. Stopher. On-Board Bus Surveys: No Questions Asked. In *Transportation Research Record 1085,* TRB, National Research Council, Washington, D.C., 1986, pp. 50–57.

Publication of this paper sponsored by Committee on Public Transportation and Fare Policy.