Generating Route-Level Mutually Exclusive Service Areas

Comparative Study of Alternative Methods

Sang Gu Lee, Daoqin Tong, and Mark Hickman

Willing-to-walk distance was investigated as a means of measuring spatial accessibility of bus stops and of examining the effectiveness of alternative methods of generating mutually exclusive transit service areas at the route level. First, the walking distance to and from a transit stop was investigated with onboard survey data. Two methods in geographical information systems—the combination of Thiessen polygon and buffer and the network distance-based service area—were compared as strategies for generating mutually exclusive service areas. For the examination of the effectiveness of these two methods, all mutually exclusive service areas were validated with a spider diagram generated from an onboard survey. Measures of urban form were also statistically tested for comparison of the two methods. A case study of a single route, serving the Minneapolis–Saint Paul, Minnesota, metropolitan area, was performed with data from various sources, such as Google's General Transit Feed Specification, an onboard survey, parcel-level land uses, and the U.S. Census street network. Validation with onboard survey data demonstrates the strengths of each method. Results also show that the network-based service area, a popular geographical information system method for service area analysis, does not yield a more meaningful strategy for generating mutually exclusive transit catchment areas, especially when spacing between stops is very small.

Forecasting transit demand is one of the major and necessary tasks in urban transportation planning. Careful transit demand modeling requires the input of very detailed information for determining transit ridership and other explanatory variables within each transit-route service area. Therefore, delineating the service area of each transit stop is critical in demand modeling, for it defines the geographic area where passengers are expected to access to and from the stop. More technically, this service area also provides a spatial unit over which each variable should be measured and integrated. Since it is often assumed that land use patterns are highly correlated with ridership generation, a transit demand model requires accurate estimation of mutually exclusive service areas, and determination of areas within walking distance of each transit stop. This study investigates the willing-to-walk distance in measuring spatial accessibility of bus stops, and it examines the effectiveness of alternative methods for generating mutually exclusive transit service areas at the route level.

Investigated first is the walking distance to and from a transit stop using onboard survey data, and the empirical walking distance distribution is derived. Then, a feature-based proximity analysis based on the estimated walking distance threshold is conducted to delineate the service area of a stop. The proximity analysis in geographical information systems (GIS) can be implemented by either buffering based on a straight-line distance or generating service areas based on a network distance. Two methods, the combination of Thiessen polygon and buffer (CTPB) and the network distance-based service area (NDSA), are used to examine the effectiveness of alternative strategies for generating mutually exclusive service areas (MESAs). These methods are applied to a case study using four data sources: Google's General Transit Feed Specification (GTFS), onboard survey data from Metro Transit (operating in the Minneapolis–Saint Paul, Minnesota, metropolitan area), parcel-level land use data from MetroGIS, and street network data from the U.S. Census. Next, all MESAs are validated with a spider diagram generated from an onboard survey. Measures of urban form are also statistically tested to compare the two methods.

The remainder of this paper is organized as follows. The next section begins with the background for this study, including a literature review, followed by the data description and preparation. The methodology section explains the development of the model. Then is a presentation and discussion of results from the analyses. Finally, there are concluding remarks and proposed future work.

LITERATURE REVIEW

Spatial data are more complex compared with traditional business data, and the traditional relational database constructs are not adequate for handling them (*1*). According to Peng and Dueker, spatial representation in transportation network elements generally consists of the following (*2*):

- Point. Basic unit of location: transit stop, intersection, transit center, or park-and-ride;
- Link. Line between two points: entire route or segment; and
- Polygon. Enclosed area bounded by three or more links: transit service area, census tract, or block.

The relationship between these spatial features can be established by point-on-line, point-in-polygon, line-in-polygon, and polygon overlay (*2*). The focus of this paper is mainly on point-in-polygon,

---

S. G. Lee and M. Hickman, Department of Civil Engineering and Engineering Mechanics, University of Arizona, 1209 East 2nd Street, P.O. Box 210072, Tucson, AZ 85721-0072. D. Tong, School of Geography and Development, University of Arizona, 1103 East 2nd Street, P.O. Box 210076, Tucson, AZ 85721-0076. Corresponding author: M. Hickman, mhickman@email.arizona.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2350, Transportation Research Board of the National Academies, Washington, D.C., 2013, pp. 37–46.
DOI: 10.3141/2350-05

related to a topological overlay procedure. This procedure also allows determining of the spatial coincidence of points (e.g., transit stops) and polygons (e.g., catchment area or service coverage).

Spatial Characteristics in Demand Modeling

Typically, transit demand models can be developed at the stop, route segment, and route levels (*2*). These three levels of models have their own advantages and disadvantages. It is difficult for the route-level approach to take into account the stop-level land use characteristics and stop-specific amenities, such as shelters and sidewalks. The route-level approach assumes that the social and demographic characteristics along the route are homogeneous, and that is invalid in most cases. In the stop-level model, point data (transit stops) often have a directional component, and that can be treated as symmetric and nonsymmetric (*2*), depending on whether stops are considered the same in both directions. Differing from rail stations, bus stops are usually represented as points that have different directions. Although the stop-level model reduces aggregation errors in many cases by considering the directional effect, it can introduce redundancy as a result of double counting if spacing between transit stops is very small. A stop aggregation model (SAM) (*3*) provides the level of aggregation between the stop- and route segment-level approaches, by relaxing the directional feature of the individual stops.

Despite the prevalence of bus transit, previous direct demand models have dealt primarily with rail transit (*4–6*). These studies have derived demand for transit at specific stations using various, multiple catchment bands: for example, within a quarter- or half-mile of a station (*5, 6*). In a station-level direct demand model, there might be an issue of what catchment areas are most appropriate, especially if mutually exclusive catchment areas are desired. Ultimately, to delineate the catchment area of a stop, there are two critical components: the willing-to-walk distance and a method of generating MESAs.

Walking Distance

To determine a transit service area, one might consider an accessibility model. The accessibility model usually requires either a maximum fixed distance (e.g., willing-to-walk distance) or a distance decay function [e.g., a negative exponential function by Zhao et al. (*7*), or a negative logistic function by Kimpel et al. (*8*)]. For this study, transit accessibility is defined as the ability of people to reach transit facilities, including bus stops (*7*). This study adopts a clear-cut boundary of spatial accessibility to define service areas, since the study focuses mainly on MESAs.

Walking is usually regarded as the primary access mode to and from a transit stop or station. In other words, most riders walk from their origins to a transit stop and from a transit stop to their final destination. So, different riders have various preferences for the amount of distance they are willing to walk. The majority of previous studies attempted to find the suitable walking distance using survey data or a theoretical approach. While models employing a decay function are likely to improve the models' predictive power, most studies apply a fixed distance threshold, such as a quarter mile or 400 m, or multiples of a quarter mile, such as a half mile or 800 m, as the access distance for service planning. Table 1 summarizes the willing-to-walk distance and distance thresholds in those applications found in the literature.

Generating Mutually Exclusive Service Areas

To delineate the catchment area of a stop, a feature-based proximity analysis in GIS is usually considered, and it can be implemented by buffering—creating polygons around input features using a straight line (Euclidean) distance or network distance. Although the ArcGIS Network Analyst (*20*) extension can calculate service areas (network buffering) without any overlapping issues, the buffer (straight-line buffering) in ArcGIS produces individual buffering polygons for all features regardless of overlap, or a single polygon after dissolving all buffering polygons. Either way, one cannot avoid the overlapping of polygons in service areas. For this reason, Thiessen polygons (also known as Voronoi diagrams) can be used to define the individual influence regions around each point feature. The Thiessen polygon approach is not confined by a specific distance but can provide mutually exclusive coverage areas for an input feature. This approach is also available within common GIS packages, and intersecting Thiessen polygons with buffered polygons around bus stops provide a way to generate mutually exclusive service areas. Upchurch et al. introduce this procedure as "nonoverlapping circles divided by Thiessen polygon boundary," and point out that this approach ignores the geometry of the actual street network (*21*). Although transportation networks are normally represented with a vector model involving points, lines, and polygons, those authors proposed the linked on–off network method with a raster model. Through the raster-based techniques, various spatial phenomena can be modeled continuously over space. However, since all attributes must be stored as a separate layer in a raster model, a vast number of layers may bring intense and complex computation. Also, the choice of raster cell size should be also considered, for finer resolution can increase computation significantly. More important, compared with raster-based approaches, vector-based approaches are much more efficient in managing and processing network topology. For these reasons, a vector-based approach is more suitable for modeling a transportation network than a raster-based one.

Straight-Line and Network Distance-Based Buffer Approaches

Proximity analysis is generally implemented by buffering based on a Euclidean distance. However, a Euclidean distance seems unrealistic in some areas because it may ignore any non-straight-line paths or obstructions, or it may define an area inaccessible to pedestrians. In these cases, a network distance can represent a more realistic understanding of travelers' behavior. O'Neill et al. suggest that the network distance is preferred, as opposed to the Euclidean distance, for identifying streets with access to a transit system (*22*). This has been confirmed by some additional studies (*11, 23*). In addition, Landex and Hansen point out that using the network distance-based method can improve the accuracy in catchment area delineation if a detailed street network is available (*24*). The network distance method provides systematically better estimates of transit ridership than the Euclidean distance method (*25*).

While a network distance-based service area delineation has been a popular GIS method for service area analysis (*8, 26*), it also has limitations in regard to both applying distance parameters and defining the service area. The network may be either underspecified, inaccuracy and lack of connectivity (e.g., isolated network such as a park, or pedestrian-level network such as shortcuts or parking lots), or overspecified, inaccessibility to pedestrians (e.g., auto-only-network such as a freeway or on-ramp). Some recent studies point out these issues

TABLE 1 Summary of Willing-to-Walk Distance

Study	Walking Distance: Distance Threshold	Measurement and Application	Location	Remark
Lam and Morrall (*9*)	292 m	Median walking distance	Calgary, Canada	Average: 327 m; 75th percentile: 450 m
O'Sullivan and Morrall (*10*)	326 m: CBD; 649 m: suburban	Average walking distance	Calgary, Canada	Distinguish between walking to LRT stations in the suburbs and in the CBD
Hsiao et al. (*11*)	400 m	Buffer	Orange County, Calif.	According to the 1990 on-board survey, more than 80% of bus riders would walk up to 400 m
Polzin et al. (*12*)	800 m	Buffer	Tampa, Fla.	800-m buffers for zonal coverage have been drawn around each route
Zhao et al. (*7*)	800 m	Buffer	Southeast Florida	By applying a decay function, a long walking distance (800 ~ 1,600 m) may be unnecessary
Kittelson and Associates (*13*)	400 m	Average walking speed of 5 km/h	North American cities	Most passengers (75% to 80% on average) walk 400 m or less to a bus stop
Alshalalfah and Shalaby (*14*)	231 m: downtown; 454 m: suburban	Median subway access distance	Toronto, Canada	
Utsunomiya et al. (*15*)	Distribution of minimum daily access distance	Estimated access distance	Chicago, Ill.	In the case of Chicago Card customers, walking access distance vary significantly between rail and bus
Kimpel et al. (*8*)	536 m (⅓ mi)	Buffer	Portland, Ore.	Initial distance of ⅓ mi and then a distance decay function is applied
Alshalalfah and Shalaby (*16*)	60% of users live within 300 m from their stop	Buffer (an interval of 100 m)	Toronto, Canada	Overall, 80% live within a distance of 500 m
Hoback et al. (*17*)	580 m	True walking distance	Detroit, Mich.	On average 1,300 m per round trip (e.g., home–transfer–work–transfer–home)
Foda and Osman (*18*)	400 m	Buffer	Alexandria, Egypt	Ideal and actual stop-accessibility indices and stop coverage ratio index are introduced
Daniels and Mulley (*19*)	454 m: inner Sydney	Mean walking distance	Sydney, Australia	From home to bus stop
	502 m: outer Sydney			From home to bus stop
	759 m: inner Sydney			From home to rail station
	873 m: outer Sydney			From home to rail station

NOTE: CBD = central business district; LRT = light rail transit.

(*21, 26, 27*). Because the NDSA delineation assumes that service areas are generated along available paths only, the delineation becomes quite complicated when spacing between facilities is very narrow. In addition, the majority of direct demand models rely on a straight-line distance to simplify the prediction process and significantly reduce labor and time (*6*).

Although a considerable body of research on transit service area delineation has been carried out, there is still a need for an improved model. The authors explored what they believe is the first attempt to validate MESAs against transit onboard survey data. Additionally, the types of land use in MESAs are evaluated through urban form measures, to investigate which MESA generation method is the best representation of transit catchment areas.

METHODOLOGY

The research methodology includes the GIS applications presented previously and analytical procedures as presented in Figure 1. First investigated is the willing-to-walk distance using onboard survey data. Spatial analyses (CTPB and NDSA) are briefly described and used to generate MESAs at the route level, based on the distance threshold determined from the willing-to-walk analyses. Then, results are validated with onboard survey data. Finally, land uses in delineated MESAs are evaluated through two measures of urban form.

Analysis of Onboard Survey Data

Onboard survey data provide important information about passengers. For this study, these data indicate how far they are willing to walk. The distribution of passengers' walking distances from and to bus stops is explored. All analyses are implemented in *R*, a language and environment for statistical computing and graphics (*28*).

Point Data Analysis with Catchment-Based SAM

Lee et al. propose a SAM to better understand the spatial and temporal interaction between transit demand and land use patterns (*3*). A catchment-based SAM (CBSAM) is developed that allows a stop to be treated either as a spatially unique stop, or as part of a group with the nearest stops in the opposite direction. As a case study, Route 6 of Metro Transit, in the Minneapolis–Saint Paul region, is selected to

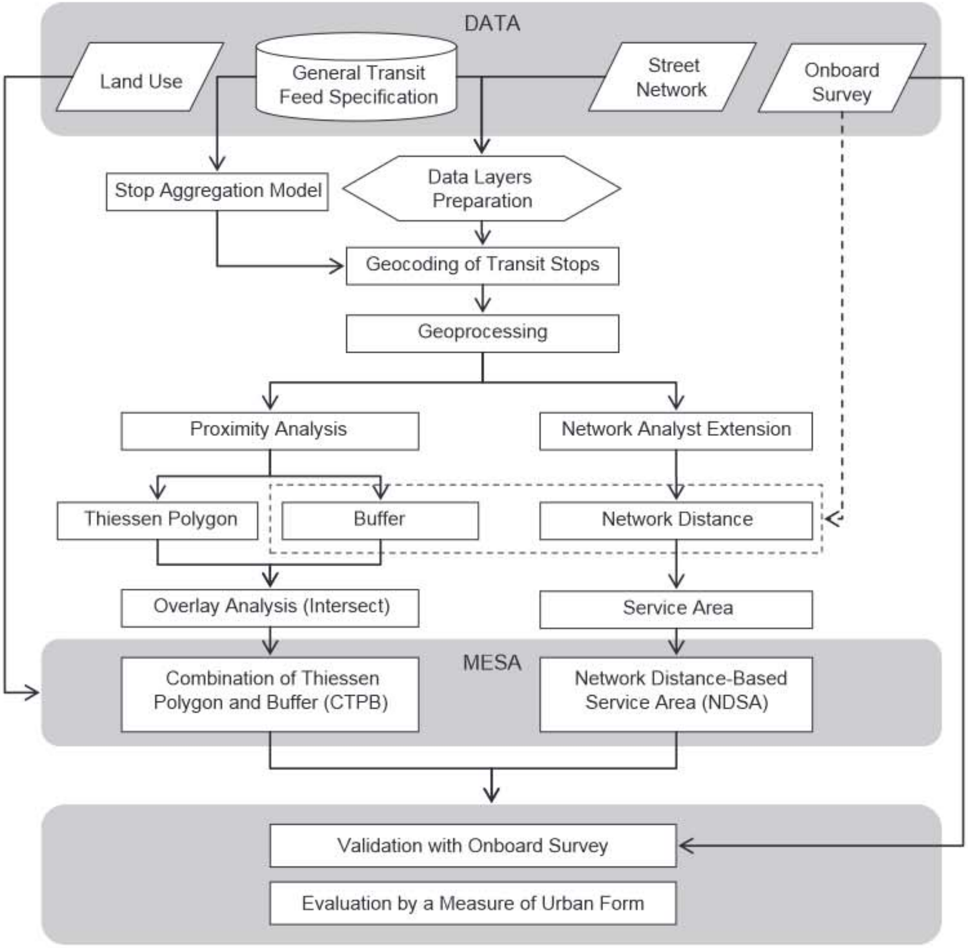


FIGURE 1 Research methodology.

implement the CBSAM methodology. The result with this route indicates that stop groups (155 stop groups in both directions) are better in representing stops that have an identical or similar transit service areas. These stop groups could be represented as a single node in the network representation (*27*). In the end, CTPB and NDSA are used to generate a set of catchment areas for the same bus stop data set with these 155 bus stop groups.

Combination of Thiessen Polygon and Buffer

Thiessen polygons are mutually exclusive, but they are not bounded by any specific distance from the points that are used to generate them. So the CTPB is introduced to generate mutually exclusive service areas for train stations (*21*). The CTPB can be constructed as follows:

1. All pseudonodes in CBSAM are triangulated into a triangulated irregular network.
2. On the basis of these pseudonodes, create (*a*) Thiessen polygons and (*b*) buffer polygons with a specific distance threshold.
3. Thiessen polygons are intersected by the buffer polygons that were created.

The resultant CTPB polygons are both mutually exclusive and bounded by a fixed radius from the bus stops.

Network Distance-Based Service Area

The network distance-based method, which delineates transit service areas that are equally distant from a stop along all available network paths, is popular in GIS applications (*8, 26*). Procedures to determine service areas can be implemented using the network analyst extension in ArcGIS (*20*). The detailed procedure is as follows:

1. All pseudonodes in CBSAM are candidates for facilities with the service area function in ArcGIS.
2. On the basis of these pseudonodes, choose "not overlapping polygons per facility" and create polygons that are closest to each facility.
3. NDSA polygons are both mutually exclusive and accessible along the network path.

Validation with Onboard Survey Data

MESAs were generated corresponding to both CTPB and NDSA for 155 bus stop groups defined by CBSAM. A spider diagram is developed to display bus stop locations and the transit passenger origins simultaneously, identifying the coverage for each bus stop or stop group. In this way, all MESAs generated by the two methods are validated using the onboard survey data. The spider diagram in GIS is implemented in Visual Basic, so one can clearly observe passenger movements both from trip origin to boarding stop and from alighting stop to destination. More details are described later in the section on data.

Evaluation Through Measure of Urban Form

Once MESAs are generated around transit stops, an evaluation procedure is necessary for direct transit demand modeling. One way of looking at this is to examine land uses, represented by either the number of land use types or proportion of each land use area, for each MESA. A particular land use type within each MESA can be used to quantify land use indicators by the total land area and numerical proportions of each land use type (*27*). Conceptually, these indicators can be associated with two classical urban form measures: the Shannon–Weaver diversity index (SWDI) based on the number of observations of a certain land use type, and the land use mix diversity (LUMD) based on the area of a certain land use type. SWDI is a commonly used diversity index in the ecological literature; it is originally derived by Shannon (*29*). SWDI can be used to measure the urban form as follows:

$$H' = -\sum p_i \ln(p_i)$$

where H' is the Shannon–Weaver diversity index expressed as $e^{H'}$, and p_i is the numerical proportion of land use type i in the catchment area.

LUMD, proposed by Bhat and Gossen (*30*), is formulated in this study as follows:

$$\text{LUMD} = 1 - \left[\frac{\left|\frac{\text{res}}{T}-\frac{1}{7}\right| + \left|\frac{\text{com}}{T}-\frac{1}{7}\right| + \left|\frac{\text{ins}}{T}-\frac{1}{7}\right| + \left|\frac{\text{ind}}{T}-\frac{1}{7}\right| + \left|\frac{\text{rec}}{T}-\frac{1}{7}\right| + \left|\frac{\text{vac}}{T}-\frac{1}{7}\right| + \left|\frac{\text{oth}}{T}-\frac{1}{7}\right|}{\frac{7}{4}}\right]$$

where

res = acres in residential use,
com = acres in commercial use,
ins = acres in institutional use,
ind = acres in industrial use,
rec = acres in recreational use,
vac = acres in vacant use,
oth = acres in other land uses, and
T = res + com + ins + ind + rec + vac + oth.

Through these two indexes, the difference of urban form using the two MESA analyses can be examined. A paired t-test is conducted to test whether the two MESA methods generate statistically different measures of urban form.

DATA

Google's General Transit Feed Specification

The GTFS is an open format updated by hundreds of transit agencies in the United States and used by Google to incorporate transit information (e.g., routes, stops, and schedules) into applications such as Google Maps (*31, 32*). From the GTFS data, the detailed location of individual stops (14,601 stops in stops.txt) from November 2008 is used. This location information was geocoded using latitude and longitude in World Geodata System 1984 (WGS84).

Onboard Survey Data

Onboard survey data from Metro Transit in 2005 is analyzed to derive access and egress patterns for this research. In the survey, responses indicate not only location information of trip origins and destinations, and boarding and alighting stops, but also access and egress modes. For example, access to and egress from a transit stop by walking, in conjunction with trip origins and destinations, indicate walking distance. More specifically, after there was geocoding of the trip origin to the boarding stop, and the alighting stop to the destination, the Euclidean distance among the origin–boarding pairs or alighting–destination pairs was calculated. These distances can be used to investigate willing-to-walk distance, as well as validation with service areas so generated. Table 2 summarizes the onboard survey data, all records available, all records in service areas along Route 6, and records by Route 6 only. For an increased accuracy of the validation, records by Route 6 only are processed with a distance threshold and spatial constraints to remove any possible outliers.

Street Network Data

The street network data (i.e., TIGER/Line files) were downloaded from the U.S. Bureau of the Census website (*33*). Street network data include all roadways and local streets.

RESULTS

How Far People Are Willing to Walk

To measure the walking distance, onboard survey data are analyzed at three levels. The records taken by Route 6 only are preprocessed to remove any possible outliers. However, data in all records available and records in service areas around Route 6 may have many potential outliers (at least in a spatial dimension). These outliers are clearly illustrated in box plots for both access and egress. The median values over three categories (i.e., line within the box) range from 200 to 300 m. The median lines seem not to be equidistant from

TABLE 2 Onboard Survey Data and Result of Validation

Record	Access	Egress	Total
All records available	8,507	4,312	12,819
All records in service areas along Route 6	910	595	1,505
Records by Route 6 only	206	241	447
Missing x-y coordinates	34	166	200
Outliers			
Distance = 0	5	4	9
Distance > 1,600 m (1 mi)	4	7	11
Records available	163	64	227
Validation with CTPB	126	29	155
Validation with NDSA	122	33	155
Hybrid model (downtown: NDSA; others: CTPB)	127	35	162

the quartiles, meaning that these data are skewed to the right (median values extend to the lower hinge).

Histograms for walking distances of access and egress are used to examine various aspects of the distribution qualitatively. More specifically, the walking distance distribution looks similar to an exponential distribution, meaning that the walking distances in the data are clustered within even smaller distances. The exception seems to be egress distances for records taken by Route 6 only.

Another way of looking at the willing-to-walk distance is to use a cumulative frequency graph. In the data set, the cumulative proportion of observations is on the vertical axis, and the walking distance is on the horizontal axis. The 95th percentile value is given for all three categories (access from 590 to 640 m and egress from 610 to 835 m). The distribution might suggest that the Euclidean distance threshold range from 600 to 800 m. On the basis of findings and insights, the selection was 600 m for CTPB and 800 m for NDSA as different distance thresholds.

Results of Two Methods

With the distance thresholds selected, Figure 2 illustrates the results of CTPB and NDSA along Route 6. Overall, the total spatial coverage is similar (CTPB: 9,162 acres, NDSA: 8,806 acres). However, some individual service areas are significantly different in the coverage areas and shapes. More details are discussed later.

Spider Diagram with Onboard Survey Data

Through the spider diagram of both access and egress, the generated MESAs are examined for proper delineation. Figure 3 illustrates two selected locations, Downtown Minneapolis and a suburban area. Validation results are also presented in Table 2. Results show that overall, for both methods, 68.3% of diagram pairs (155 of 227) are adequately situated within the generated MESAs.

An interesting observation from the spider diagram is that the two approaches have their own strengths. As shown in Figure 3, service areas of NDSA in Downtown Minneapolis, where service areas have small spacing and are generated perpendicularly to roadways, capture the access more accurately. On the contrary, CTPB applies better to locations where there are shortcuts, parking lots, or open spaces. This suggests that the two approaches complement each other and their combined use is possible. For example, a hybrid model, which applies NDSA in downtown areas and CTPB in other areas, is also tested. The validation indicates that 71.4% (162 of 227) of diagram pairs adequately correspond to the generated MESAs (Table 2).

Results of Paired *t*-Test

Once the layer of MESAs is overlaid with the layer of parcel-level land uses, land use patterns within MESAs are examined using the two urban form measures. Table 3 summarizes the comparison results and shows that for SWDI, the mean value of CTPB is a little higher than that of NDSA; and for LUMD, the mean value of CTPB is much lower than that of NDSA. A hypothesis test is conducted to determine whether the difference of the urban form measures within the MESAs is statistically significant.

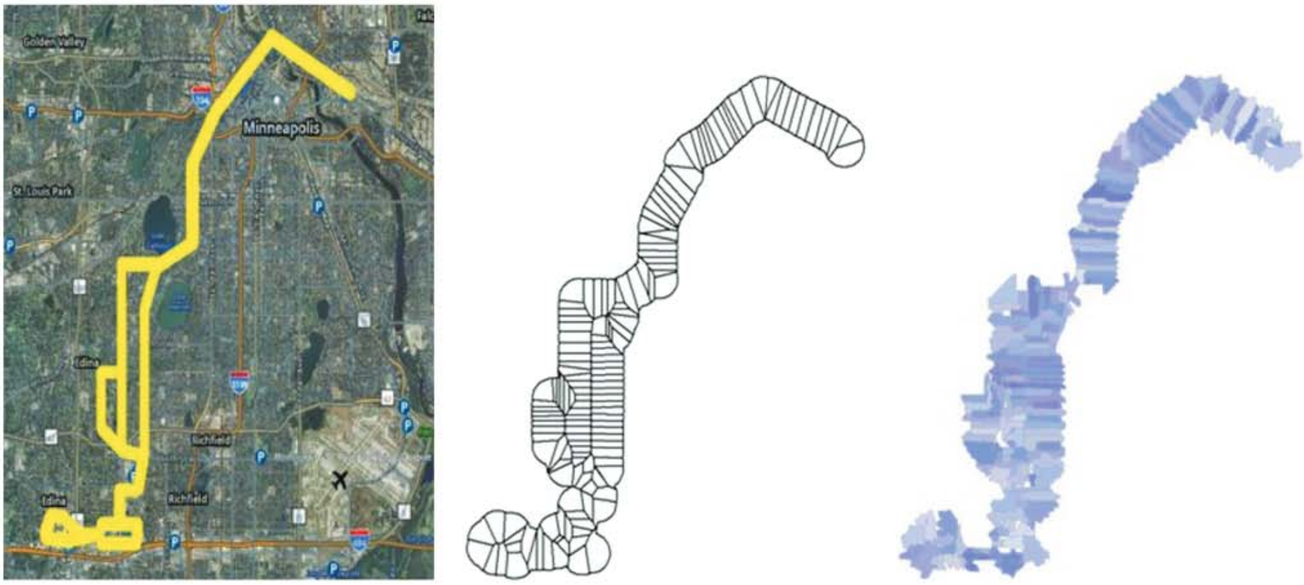
The null hypothesis (H_0) is that the urban form measure is the same regardless of the MESA generation method. The alternative hypothesis (H_A) is that the urban form measure depends on the MESA generation method.

Let $\bar{d}$ be the difference of the urban form measure for MESAs generated by the two methods with the null hypothesis at a 95% confidence interval ($\alpha = .05$), then

H_0: $\bar{d} = 0$

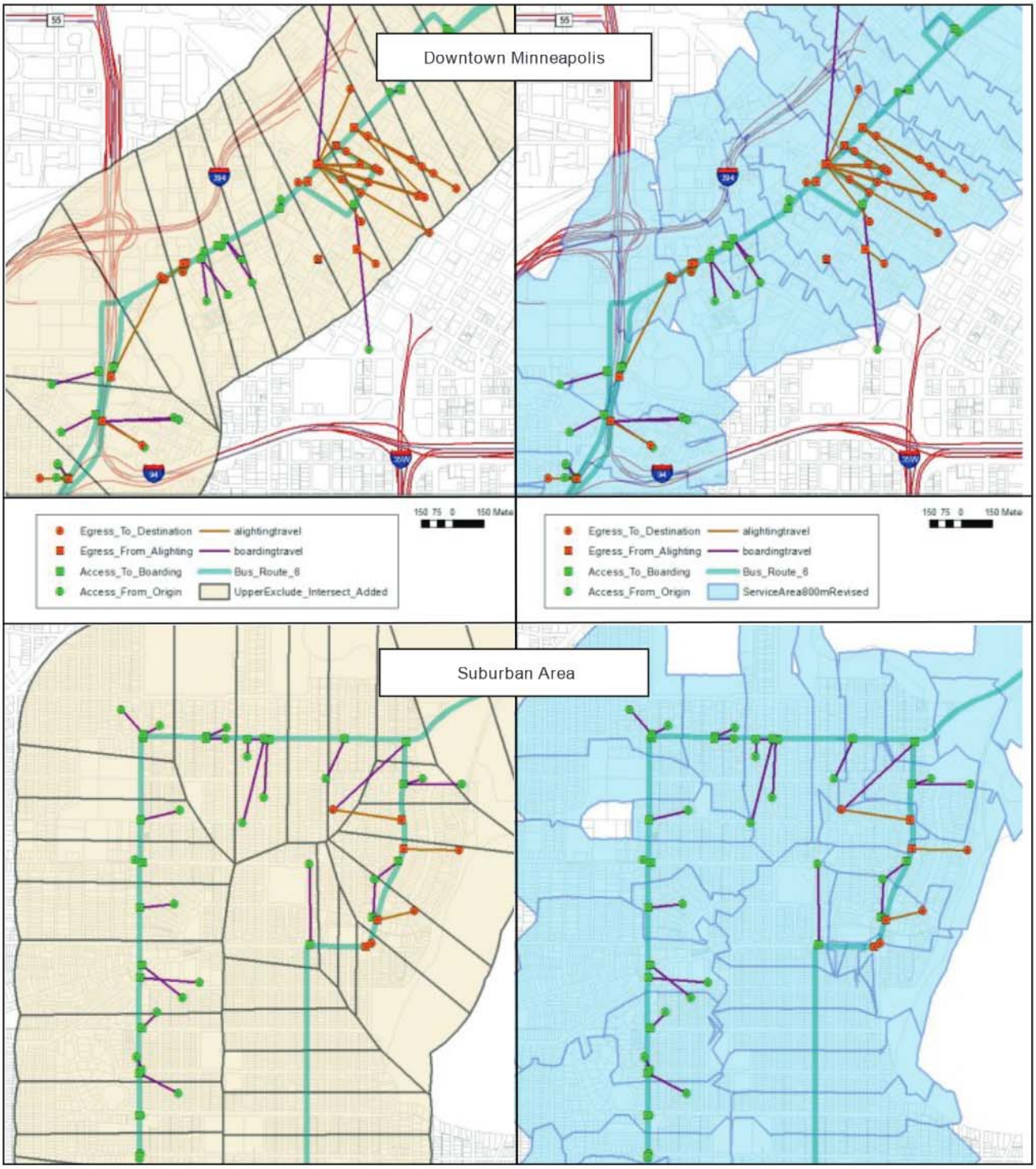
and

H_A: $\bar{d} \neq 0$



(a) (b) (c)

FIGURE 2 Results of CTPB and NDSA along Route 6: (a) Route 6, (b) CTPB, and (c) NDSA.



(a) (b)

FIGURE 3 Spider diagram using onboard survey data for validation: (a) CTPB and (b) NDSA.

TABLE 3 Paired Samples Statistics and Test

	Measure of Urban Form by Method			
	SWDI[a]		LUMD[b]	
Statistic	CTPB	NDSA	CTPB	NDSA
Mean	1.619	1.566	0.172	0.420
Variance	0.575	0.425	0.017	0.015
Standard deviation	0.758	0.652	0.130	0.122
Standard error means	0.061	0.052	0.010	0.010

NOTE: For SWDI and LUMD: N = 155; hypothesized mean difference = 0; df = 154; t critical one-tail = 1.655; t critical two-tail = 1.975. df = degrees of freedom.

[a] Pearson correlation = .870; t-statistic = 1.777; P (T ≤ t) one-tail = .039, two-tail = .078.

[b] Pearson correlation = - .168; t-statistic = - 16.049; P (T ≤ t) one-tail = .000, two-tail = .000.

Table 3 also includes results of the paired t-test. For the SWDI measure, there is no significant difference between CTPB and NDSA methods based on the p-value (.078) > .05. For the LUMD measure, on the contrary, there is rejection of the null hypothesis with the p-value (.000) < .05. If LUMD is used for measuring urban form, at the α = .05 level of significance, there exists enough evidence to conclude that there is a difference in the LUMD measure for the MESAs generated by the two methods. Results of the paired t-test indicate that either CTPB or NDSA methods can be used to generate MESAs if the numerical proportion of land use types in each MESA is considered as the basic unit of analysis. Finally, MESAs by CTPB and NDSA methods may significantly differ in area proportion, and additional research effort should investigate this difference when LUMD is constructed on the basis of area proportions.

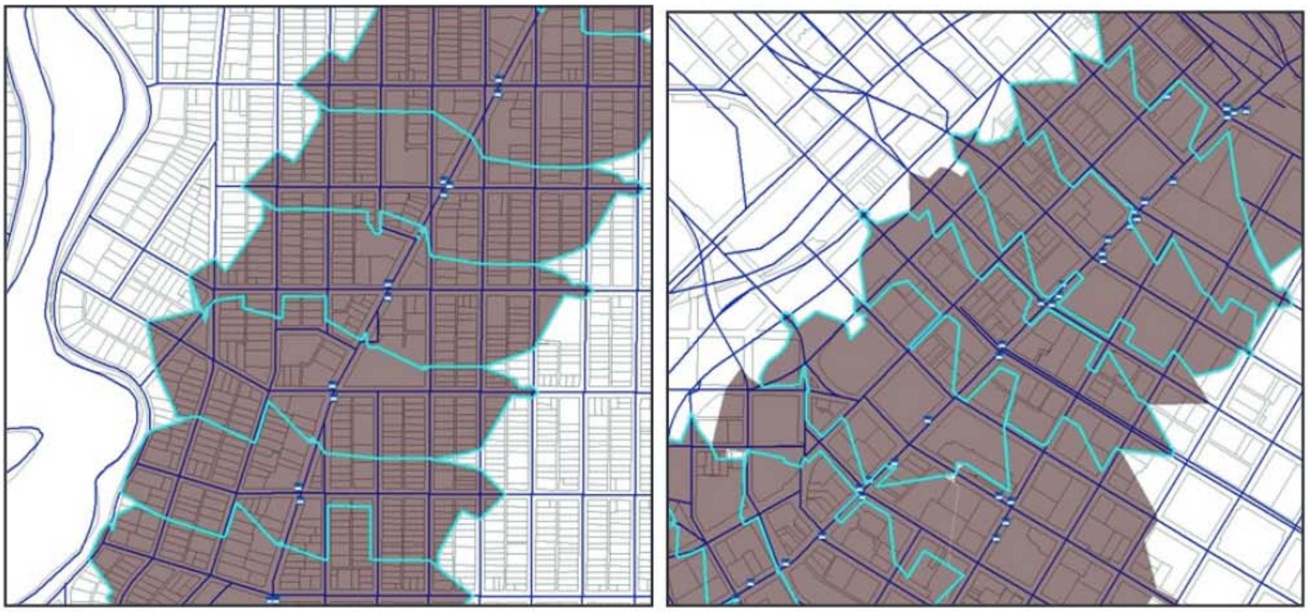
DISCUSSION OF RESULTS

An advantage of NDSA is clearly a higher degree of accuracy compared with CTPB in some cases. This includes a better spatial representation, which is perpendicularly generated to an oblique roadway, as shown in Figure 4a. However, one barrier for NDSA is the lack of a complete street-level network. The other is the problematic implementation of NDSA when creating service areas. As shown in Figure 4b, service areas generated by NDSA are unrealistic and often overestimated and underestimated. This stems from the basic assumption that service areas (polygons) are generated along available paths only, a complicated situation when spacing between facilities is very narrow. Furthermore, compared with traditional neighborhoods with grid street patterns, suburban neighborhoods in which streets often are curvilinear with cul-de-sacs make it almost impossible to have separate (nonoverlapping) service areas (27).

Although the NDSA has been a popular GIS method for service area analysis, this study indicates that the method does not yield a more meaningful strategy for generating transit catchment areas, especially when spacing between stops is very small. Furthermore, a simple approach, the CTPB method, seems to be a better alternative for generating certain mutually exclusive service areas when using consecutive transit stop areas. This is also consistent with the fact that the majority of direct demand models rely on a circular buffer method, measured around a single transit stop. For this reason, the CTPB method can be applied to determine catchment areas along a single route in future research.

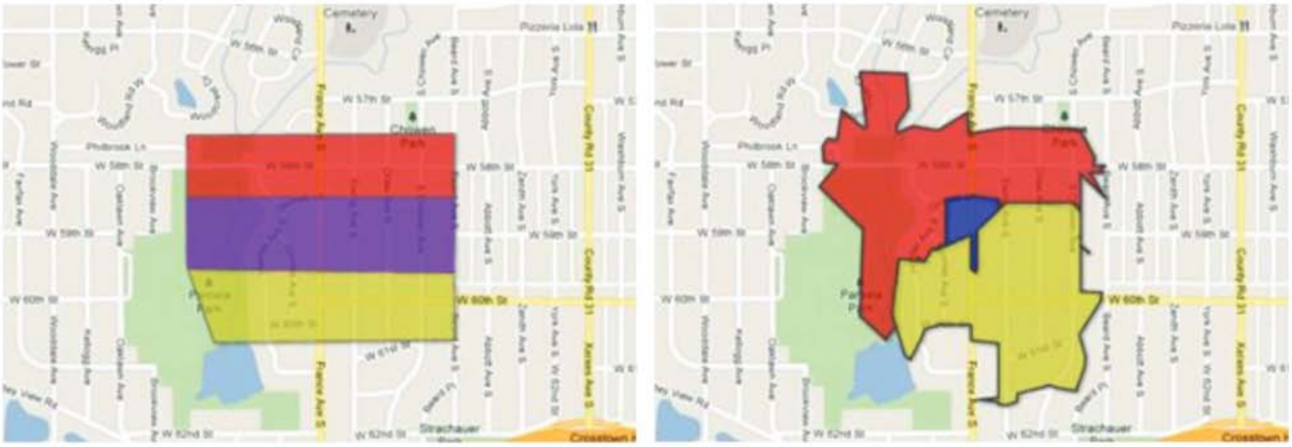
CONCLUSIONS

This study investigates the willing-to-walk distance in spatial accessibility using onboard survey data and describes an effort to generate mutually exclusive transit service areas at the route level. More spe-



(a)

FIGURE 4 Discussion on generated MESAs: (a) sensitive to oblique roadways in NDSA of two locations in Minneapolis–Saint Paul. (continued)



(b)

FIGURE 4 (continued) Discussion on generated MESAs: (b) three neighboring MESAs generated from CTPB (left) and from NDSA (right) with Google's Fusion Table.

cifically, with a walking distance threshold, the CTPB and the NDSA methods are used to highlight the generation of mutually exclusive service areas for accurate transit access and egress estimates. Then, results of two methods are validated with the onboard survey by using a spider diagram, and the land use patterns within each service area are also statistically tested by using a paired t-test.

The overall conclusion of this paper is that the NDSA approach is not necessarily a better strategy for generating certain catchment areas, especially for MESAs at the route level. Furthermore, the CTPB approach may be a better alternative for generating certain MESAs when using consecutive transit stops. These findings may be used to help transit agencies improve the capability of spatial data integration in direct demand modeling. This capability is being pursued in current research.

Analyzing other routes in Metro Transit and other transit agencies will be of additional benefit to validate the two methods. A proposed extension to this study combines two methods with their own strategic advantages, as shown in the hybrid method tested here. Also, there are plans to create a user-friendly interface and systematic procedure, to allow people to generate mutually exclusive service areas around transit lines or stops.

ACKNOWLEDGMENTS

The authors thank Metro Transit and MetroGIS for providing the data analyzed in the study. Specific recognition goes to Paul Krech and John Levin of the scheduling department at Metro Transit for generating the GTFS data set from the November 2008 schedules.

REFERENCES

1. Shekhar, S., and S. Chawla. Spatial Databases: A Tour. Prentice Hall, Upper Saddle River, N.J., 2003.
2. Peng, Z., and K. J. Dueker. Spatial Data Integration in Route-Level Transit Demand Modeling. Journal of Urban and Regional Information Systems Association, Vol. 7, No. 1, 1995, pp. 26–37.
3. Lee, S. G., M. Hickman, and D. Tong. Stop Aggregation Model: Development and Application. In Transportation Research Record: Journal of the Transportation Research Board, No. 2276, Transportation Research Board of the National Academies, Washington, D.C., 2012, pp. 38–47.
4. Kuby, M., B. Anthony, and C. Upchurch. Factors Influencing Light-Rail Station Boardings in the United States. Transportation Research Part A: Policy and Practice, Vol. 38, No. 3, 2004, pp. 223–247.
5. Usvyat, L., L. Meckel, M. DiCarlantonio, and C. Lane. Sketch Model to Forecast Heavy-Rail Ridership. Presented at 88th Annual Meeting of the Transportation Research Board, Washington, D.C., 2009.
6. Guerra, E., R. Cervero, and D. Tischler. Half-Mile Circle: Does It Best Represent Transit Station Catchments? In Transportation Research Record: Journal of the Transportation Research Board, No. 2276, Transportation Research Board of the National Academies, Washington, D.C., 2012, pp. 101–109.
7. Zhao, F., L.-F. Chow, M.-T. Li, I. Ubaka, and A. Gan. Forecasting Transit Walk Accessibility: Regression Model Alternative to Buffer Method. In Transportation Research Record: Journal of the Transportation Research Board, No. 1835, Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 34–41.
8. Kimpel, T. J., K. J. Dueker, and A. M. El-Geneidy. Using GIS to Measure the Effect of Overlapping Service Areas on Passenger Boardings at Bus Stops. Journal of Urban and Regional Information Systems Association, Vol. 19, No. 1, 2007, pp. 5–11.
9. Lam, W., and J. Morrall. Bus Passenger Walking Distances and Waiting Times: A Summer-Winter Comparison. Transportation Quarterly, Vol. 36, No. 3, 1982, pp. 407–421.
10. O' Sullivan, S., and J. Morrall. Walking Distances to and from Light-Rail Transit Stations. In Transportation Research Record 1538, TRB, National Research Council, Washington, D.C., 1996, pp. 19–26.
11. Hsiao, S., J. Lu, J. Sterling, and M. Weatherford. Use of Geographic Information System for Analysis of Transit Pedestrian Access. In Transportation Research Record 1604, TRB, National Research Council, Washington, D.C., 1997, pp. 50–59.
12. Polzin, S. E., R. M. Pendyala, and S. Navari. Development of Time-of-Day–Based Transit Accessibility Analysis Tool. In Transportation Research Record: Journal of the Transportation Research Board, No. 1799, Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 35–41.
13. Kittelson and Associates, Inc. TCRP Report 88: A Guidebook for Developing a Transit Performance-Measurement System. Transportation Research Board of the National Academies, Washington, D.C., 2003.
14. Alshalalfah, B., and A. Shalaby. Relationship Between Walk Access Distance to Transit and Socioeconomic, Demographic, and Transit Service Characteristics. Presented at 84th Annual Meeting of the Transportation Research Board, Washington, D.C., 2005.
15. Utsunomiya, M., J. Attanucci, and N. H. Wilson. Potential Uses of Transit Smart Card Registration and Transaction Data to Improve Transit

Planning. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1971,* Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 119–126.

16. Alshalalfah, B., and A. Shalaby. Case Study: Relationship of Walk Access Distance to Transit with Service, Travel, and Personal Characteristics. *Journal of Urban Planning and Development,* Vol. 133, No. 2, 2007, pp. 114–118.
17. Hoback, A., S. Anderson, and U. Dutta. True Walking Distance to Transit. *Journal of Transportation Planning and Technology,* Vol. 31, No. 6, 2008, pp. 681–692.
18. Foda, M. A., and A. O. Osman. Using GIS for Measuring Transit Stop Accessibility Considering Actual Pedestrian Road Network. *Journal of Public Transportation,* Vol. 13, No. 4, 2010, pp. 23–40.
19. Daniels, R., and C. Mulley. Explaining Walking Distance to Public Transport: The Dominance of Public Transport Supply. *Proc., World Symposium on Transport and Land Use Research,* Whistler, British Columbia, Canada, 2011.
20. ArcView Network Analyst: Optimum Routing Closest Facility and Service Area Analysis. Environmental Systems Research Institute, Redlands, Calif., 1996.
21. Upchurch, C., M. Kuby, M. Zoldak, and A. Barranda. Using GIS to Generate Mutually Exclusive Service Areas Linking Travel On and Off a Network. *Journal of Transport Geography,* Vol. 12, No. 1, 2004, pp. 23–33.
22. O'Neill, W. A., R. D. Ramsey, and J. Chou. Analysis of Transit Service Areas Using Geographic Information Systems. In *Transportation Research Record 1364,* TRB, National Research Council, Washington, D.C., 1992, pp. 131–138.
23. Horner, M. W., and A. T. Murray. Spatial Representation and Scale Impacts in Transit Service Assessment. *Environment and Planning B: Planning and Design,* Vol. 31, No. 5, 2004, pp. 785–797.
24. Landex, A., and S. Hansen. Examining the Potential Travelers in Catchment Areas for Public Transport. *Proc., ESRI International User Conference,* San Diego, Calif., Environmental Systems Research Institute, Redlands, Calif., 2006.
25. Gutiérrez, J., and J. C. García-Palomares. Distance-Measure Impacts on the Calculation of Transport Service Areas Using GIS. *Environment and Planning B: Planning and Design,* Vol. 35, No. 6, 2008, pp. 480–503.
26. Cheng, C., and A. W. Agrawal. TTSAT: A New Approach to Mapping Transit Accessibility. *Journal of Public Transportation,* Vol. 13, No. 1, 2010, pp. 55–72.
27. Lee, S., M. Hickman, and D. Tong. Development of a Temporal and Spatial Linkage Between Transit Demand and Land Use Patterns. *Journal of Transport and Land Use,* Vol. 6, No. 2, 2013, pp. 33–46.
28. Venables, W. N., D. M. Smith, and the R Core Team. *An Introduction to R.* http://www.r-project.org/. Accessed Dec., 2010.
29. Shannon, C. E. A Mathematical Theory of Communication. *Bell System Technical Journal,* Vol. 27, No. 3, 1948, pp. 379–423 and pp. 623–656.
30. Bhat, C. R., and R. Gossen. A Mixed Multinomial Logit Model Analysis of Weekend Recreational Episode Type Choice. *Transportation Research B: Methodological,* Vol. 38, No. 9, 2004, pp. 767–787.
31. Google General Transit Feed Specification. http://code.google.com/transit/spec/transit_feed_specification.html. Accessed March 4, 2011.
32. GTFS Data Exchange. http://www.gtfs-data-exchange.com. Accessed Dec. 2010.
33. U.S. Bureau of the Census, U.S. Department of Labor. TIGER files. http://www.census.gov/geo/www/tiger/. Accessed Dec. 2010.

All errors remain with the authors.

The Public Transportation Planning and Development Committee peer-reviewed this paper.