JOURNAL OF ADVANCED TRANSPORTATION
J. Adv. Transp. 2016; **50**:1770–1784
Published online 18 November 2016 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/atr.1428

A GIS-based method to identify cost-effective routes for rural deviated fixed route transit

Hongtai Yang[1], Christopher R. Cherry[2], Russell Zaretzki[3], Megan S. Ryerson[4], Xiaobo Liu[1]* and Zhijian Fu[1]

[1]*National United Engineering Laboratory of Integrated and Intelligent Transportation, School of Transportation and Logistics, Southwest Jiaotong University, Chengdu, Sichuan, China*
[2]*Department of Civil and Environmental Engineering, University of Tennessee-Knoxville, Knoxville, TN, USA*
[3]*Department of Business Analytics and Statistics, University of Tennessee-Knoxville, Knoxville, TN, USA*
[4]*Department of City and Regional Planning, Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA, USA*

SUMMARY

Deviated fixed route transit (DFRT) service connecting rural and urban areas is a growing transportation mode in the USA. Little research has been done to develop frameworks for route design. A methodology to explore the most cost-effective DFRT route is presented in this paper. The inputs include potential DFRT demand distribution and a road network. A heuristic is used to build possible routes by starting at urban cores and extending in all network directions in certain length increments. All the DFRT routes falling in the length range desired by the users are selected. The cost effectiveness of those routes, defined by operating cost per passenger trip, is compared. The most cost-effective route is selected and presented in a GIS map. A case study illustrates the methodology in several Tennessee metropolitan regions. The most cost-effective route length is case specific; some routes (e.g. those out of our Nashville case) are most cost effective when short, while others (e.g. those out of Memphis) are most cost effective when long. Government agencies could use the method to identify routes with the lowest operating cost per passenger given a route length or an operating cost budget. Copyright © 2016 John Wiley & Sons, Ltd.

KEY WORDS: deviated fixed route transit; geographic information system; transit route design; rural to urban transit; demand responsive transit

1. INTRODUCTION

The lack of public transport opportunities to access necessary basic service outlets and amenities located in distant urban centers has always been a problem for rural residents [1, 2]. The traditional public transport services operated in rural areas are fixed route, fixed schedule transit and demand responsive transit (DRT) [3]. Over the last 50 years, flexible transportation services that combine the features of fixed route, fixed schedule transit and purely demand responsive transit have been proposed by practitioners [4, 5]. Koffman [6] and Potts [7] report operating strategies of existing flexible public transportation services provided by transit agencies. Among the many types of flexible public transportation services, route deviation service or deviated fixed route transit (DFRT) is most widely used [6, 8].

DFRT can be referred to as mobility allowance shuttle transit (MAST) [9–14], route deviation transit [7, 15–17], demand adaptive transit systems [18–21] or flex-route transit [22, 23] in different studies. It operates along a route with fixed stops at generally fixed times, “but may deviate from the route alignment to collect or drop off passengers who have requested the deviation” [24]. The only

*Correspondence to: Xiaobo Liu, National United Engineering Laboratory of Integrated and Intelligent Transportation, School of Transportation and Logistics, Southwest Jiaotong University, 01116, No.111, First Section, North of Second Ring Road, Chengdu, Sichuan 610031, China. E-mail: xiaobo.liu@swjtu.cn

restriction on flexibility is that the deviations must be within a predetermined distance from the fixed route [10]. Point to point pick-up and drop-off within a certain area can be provided using DFRT services. Its flexibility could make it more suitable to serve rural areas than fixed route counterparts because of dispersed rural population. In late 1980s, the deregulation of the intercity bus (ICB) industry resulted in abandonment of many ICB (fixed route, fixed schedule) routes, especially in rural areas, because of the low revenue generation from rural riders [25, 26]. Service cuts negatively impacted the mobility of bus-riding rural residents. DFRT has been playing an increasingly important role in meeting rural residents' transportation needs. According to the National Transit Database (NTD), in 2010 there were 302 rural DFRT operators in the US, providing around 22 million passenger trips with total revenue miles around $65 million. Because the service area of DFRT is a buffer along the route, which is different from the service area of the fixed route, fixed schedule transit, a buffer around the bus stop [27], the route planning method for regular bus is not applicable to DFRT planning. Although the DFRT services have been implemented across US, there are few standards to which transit service providers can turn when designing service: Koffman [6] summarized transit agencies' experiences with flexible transit services and provides a good summary of practices that were commonly used.

The existing literature related to DFRT focuses on optimizing various service parameters of DFRT while constraining other parameters. No studies have focused on the route planning of DFRT using existing networks and heterogeneous demand. Recognizing the lack of literature on empirical route design, this study proposes a methodology to design efficient DFRT routes.

This study is organized as follows. First, we review the relevant literature and gain insight on methods and applications of DFRT service planning. Second, we propose a method to develop efficient DFRT routes with empirical network and population data. Using this method, we present a case study and follow this with suggestions on how our methodology can be used by planners and decision makers. Although this methodology is for DFRT, it is applicable to any flexible transportation services whose service area is a buffer along the route.

2. LITERATURE REVIEW

Relatively few existing studies focus on DFRT DFRT. Errico et al. [28] classified existing studies on flexible transportation services (called "semi-flexible transit" in his study) into two major areas: "one describing practical experiences, with few or no methodological contributions", and "the other proposing methodologies to address a number aspects of the planning process". They point out that "there is no or very limited use of optimization in the planning activities" in the actual implementation of flexible transportation systems. This observation is also noted by Potts et al. [7] and Scott [29]. Our paper proposes a methodology to select the most cost-effective DFRT routes based on actual demand. Therefore, it contributes to the methodological literature on route planning and optimization.

In methodological studies directly related to DFRT, the common theme is the development of relationships between various design parameters of DFRT, which help transit planners consider the trade-offs between parameters or optimize one specific parameter. They are usually proposed under an ideal or hypothetical operating environment, such as a grid network [10, 22, 30].

Quadrifoglio et al. [9, 12–14] studied the upper and lower bounds of DFRT service on the maximum longitudinal velocity [13], scheduling algorithms [9, 12] and the sensitivity to the shape of the service area [14]. The relationships obtained between the main parameters of service, such as slack time and headway, can be very helpful in the design of DFRT systems. Fu [22] developed a model for the relationship between various system variables, such as the number of feasible deviations, slack time, zone size and dwell time, based on a grid network. Smith and Demetsky [30] explored the relationship between service zone size, which is the area between fixed stops where deviations are permitted, and slack time distribution among zones. Zhao and Dessouky [10] analyzed the relationship between service cycle time, and the length and width of the service area. Those studies serve as a good foundation for DFRT parameter optimization, yet provide little practical or empirical approaches to service planning or route design. Alshalalfah and Shalaby [31, 32], and Qiu et al. [33, 34] explored the issues related to replacing fixed route transit service by flex-route service and the impacts on stakeholders.

Other articles do not study DFRT solely but focus on flexible transportation services that include "all types of hybrid services that are not pure DRT or fixed-route service" [6]. Most of the flexible

DOI: 10.1002/atr

transportation services are implemented in rural and small urban areas [16]. Koffman [6] studied the flexible transportation services operated by 24 transit systems and acknowledged that there are few standards that transit service providers can use when designing service. Potts et al. [7] analyzed features of small, medium and large urban and rural transit agencies and described the appropriate flexible transportation services strategies for each of them. Nourbakhsh and Ouyang [35] compared the performance of the flexible route service to fixed-route and taxi service under hybrid network structure, and showed that each system is advantageous under certain passenger demand levels. Mulley and Daniels [36] proposed a method to use flexible transport services to improve access to public transport in low density areas. Qiu et al. [37] proposed a new transit service named *demi*-flexible transit, which falls between flexible transit and fixed route systems by not providing door-to-door services but offering flexibility to transit customers, and proved its advantages of serving low demand areas. In summary, most relevant studies state that there is a need to look at the planning and operation of flexible transportation services. The existing methodological DFRT studies focus on operations of existing or hypothetical DFRT service, but do not provide guidance on route planning. Important challenges of where to implement the deviated fixed routes remain unanswered. This paper will address these challenges by proposing a method to plan DFRT routes in a study area.

3. METHODOLOGY

In this section, we present a methodology to select the most cost-effective routes meeting the users' route length criteria. Because rural to urban DFRT usually traverse long distances, we eliminate the use of local streets as main bus routes, limiting mainline service to interstates and state highways. The proposed method generates an exhaustive set of routes of specified length. These routes are compared by their cost effectiveness (operating cost divided by potential ridership). The method selects the route with the highest cost effectiveness as the leading route.

The method proposed here relies on geographic information systems (GIS). It is organized into four steps: (i) define model inputs including urban centers, road network and travel demand; (ii) specify parameters for DFRT service and route extension; (iii) execute route extension and inventory qualified routes; and (iv) evaluate and compare cost effectiveness of developed routes. The details of each step are presented below.

3.1. Define model inputs including urban centers, road network and travel demand

Because the main role of rural DFRT is to transport rural residents to meaningful destinations (such as hospitals, airports, shopping malls) in urban centers, we assume that urban centers are the destinations of DFRT riders and could serve as the starting point (or ending point) of DFRT routes. The road network in the study area is needed and included in the GIS map. In terms of travel demand, we assume that DFRT trips are generated from places outside of urban centers and attracted to urban centers. A DFRT demand map (e.g. population distribution fitting DFRT rider profiles) is needed as an input to facilitate ridership calculation and route comparison. Because rural transit services are usually used by captive transit riders, the DFRT service demand is different from car-based travel demand. Because of the limited data on DFRT, little research has been conducted to estimate DFRT services' demand. Fravel et al. [25] developed a toolkit to estimate demand for rural intercity bus services, which can serve as a reference to estimate demand for DFRT services. Because DRT services are usually widely available to rural residents, the DRT ridership data could be used to estimate its demand distribution, which could serve a proxy for DFRT services' demand distribution because DFRT riders have similar characteristics as DRT riders [38] and could be attracted to DFRT [39]. An example of using method to estimate DFRT demand is shown in the case study section.

3.2. Specify parameters for DFRT service and route extension

The service area of each route is a buffer around the road network. Because the DFRT could deviate on both sides of the route, the width of the buffer is twice the allowed maximum deviation distance L_d specified by users. TCRP Synthesis 53 [6] states that most DFRT service providers allow a maximum deviation distance of 0.75 miles. Here, 0.75 miles is used as a commonly allowed maximum deviation

DOI: 10.1002/atr

distance, while each service provider could have their own choice based on policy or service design. The second parameter required by the algorithm is the desired route length range. All of the routes falling into that range are inventoried along with their comparative cost effectiveness. The third parameter is service frequency. Frequency of DFRT is required in order to calculate the operating cost.

All interstates and highways outside of the urban areas are divided into L_s length segments by the algorithm, where L_s is the fourth parameter specified by users. Some of the segments are shorter than L_s because the length of interstates and highways may not be integer multiples of L_s. By ordering all routes by their length, L_s will be the maximum length difference between two neighboring routes. When L_s is short, the heuristic generates a larger number of qualified route alternatives and more cost-effective solutions, but increases computational time. Choice of L_s requires consideration of the trade-off between computational time and result quality. However, the length of the segment should not be shorter than the shortest length of the unit demand area, such as a census tract, on the demand map because the demand distribution for shorter segments cannot be accurately estimated when demand distribution is averaged over a larger area.

By laying the $2L_d$-wide buffer on each L_s-long segment on the travel demand map, the DFRT demand served (or rider catchment area and potential ridership) for each segment is calculated. The column names (variables) in the segments table are represented below:

- Segment ID
- Segment length (mi.)
- Coordinates (longitude and latitude) of segment ends
- Connected to an urban county: defined by at least one end of the segment on the boundary of an urban county. Yes: 1; no: 0.

3.3. Execute route extension and inventory qualified routes

A route segment starts with one end on the boundary of an urban county and the other end outside of the county. It radiates out by connecting to a neighboring segment and adds that segment to the route. An example of the route development is shown in Figure 1.

Here, there are four segments connecting to urban county A, segment 1, 2, 5 and 8. They will be the first four routes and will be added to the dataset of developed routes. These four routes will extend out by connecting to neighboring segments. For instance, route 1 (composed of segment 1 only) will connect to segment 2 to develop the fifth route and connect to segment 3 to develop the sixth route. The first four routes perform this task together. After this round of route extension, seven routes are added to the table of developed routes. They are routes composed by segments 1 and 2, segments 1 and 3, segments 2 and 3, segments 5 and 4, segments 8 and 6, segments 8 and 7, and segments 8 and 9. Now the table of developed routes contains 11 routes. This process will continue until one of the following four criteria is met:

(1) A route reaches an urban county boundary, such as the case of routes composed by segments 1 and 2, and segments 4 and 5.
(2) The route length is equal or longer than the upper bound of the user defined route length range (dictated by allowable budget for instance).

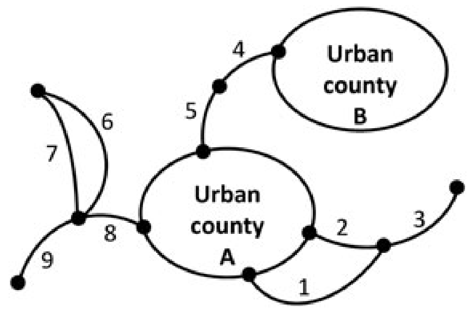

Figure 1. Route development illustration.

DOI: 10.1002/atr

(3) The next connecting segment is already in the route. For instance, the route composed of segments 8, 7 and 6 can only extend out by connecting to segment 9, because segments 6, 7 and 8 are already in the route.
(4) No more connecting segments are available.

The route development process will stop when all the routes meet at least one of the exit criteria. Routes that meet the user defined length criteria will be selected from the dataset for analysis. The column names (variables) of the developed routes dataset are presented below.

- Route ID
- Route length (mi.)
- ID of segments in the route
- Coordinates (longitude and latitude) for the starting point and ending point of the route. The coordinates of the ending point will be used to determine the next connecting segment.

The number of columns is not fixed. As the routes extend out, the table will include more segments. The number of rows is also not fixed. As the routes extend out, additional rows will be included in the table.

3.4. Evaluate and compare cost effectiveness of developed routes

It is assumed that potential riders who live in the buffered DFRT route service area will use DFRT while those who live outside will not use DFRT. It is important to note that the demand estimates are most likely accurate relative to other census tracts, rather than absolute. The ridership estimates are assumed to be completely determined by DFRT travel demand (trip generation) map generated from socioeconomic and population distribution variables and the service area of the route. As such, the cost effectiveness estimates are appropriate for ranking and comparison of routes rather than forecasting the absolute ridership. In this study, the operating cost is assumed to be determined by the annual vehicle miles provided by the service. The operating cost of each route per year is calculated as follows:

$$\begin{aligned} \mathit{Operating\ cost} = {} & \mathit{Route\ length} \times \mathit{Service\ days\ in\ a\ year} \times \mathit{Service\ frequency} \\ & (\mathit{number\ of\ round\ trips\ per\ day}) \times 2\ (\mathit{two\ trips\ per\ round\ trip}) \\ & \times \mathit{Average\ operating\ cost\ per\ annual\ vehicle\ mile}. \end{aligned}$$

The operating cost is assumed to be based on route length and service frequency. Although many variables such as vehicle size affect the operating cost, it is not considered here. The travel distance of deviation is not considered in the determination of operating cost. So comparison of cost effectiveness should be made among routes that have the same allowed maximum deviation distance (buffer width). The average operating cost per annual vehicle mile of DFRT service is obtained from the National Transit Database (NTD) [24]. The mean of 324 systems in 50 states is \$2.88 per vehicle mile traveled.

Thus, each developed route has an associated cost (based on vehicle miles) and ridership (based on demand estimates) yielding a cost effectiveness value. The cost effectiveness changes as routes extend; operating costs increase and potential ridership simultaneously increases. Because the operating cost is calculated, the cost effectiveness, defined as operating cost per passenger trip served, is calculated. Routes with the lowest operating cost per passenger trip served are defined as the most cost-effective routes for that class. The most cost-effective route has the following quality: the route has the lowest operating cost per passenger, or, equivalently, the highest potential ridership per route mile. To evaluate the effectiveness of this method, in general, one can divide the percentage of DFRT demand covered by the percentage of area covered. If drawing a random line across the study area, the average ratio of DFRT demand covered to area covered should be one. However, because populations are generally located near transportation infrastructure, all of the developed routes follow the transportation infrastructure and cover more population than a random line across the state. We can compare

DOI: 10.1002/atr

the most cost-effective and least cost-effective routes by comparing these population coverage ratios. A higher ratio indicates a better solution (route).

4. CASE STUDY

The DFRT route planning method is applied to the entire state of Tennessee. The DFRT trip generation map is first developed as an input for DFRT route planning. The four steps of the methodology are followed and the most cost-effective routes and results are discussed.

4.1. Trip generation modeling

This section describes a method to estimate DFRT trip generation rate (trip generation divided by the area) of all census tracts in Tennessee using DRT OD data. Yang and Cherry [38] found that most of DRT trips originate in rural areas and terminate in urban centers, similar to DFRT trips. Balog [39] stated that DRT demand distribution could be a suitable proxy for DFRT demand distribution because DRT riders have the same demographics as DFRT riders and could be attracted to DFRT services. Thus, we use DRT trip generation as a reasonable proxy for DFRT trip generation. If the trip generation of the study area is readily available, this step could be skipped. When the scope of the planning study is large, such as at a state level, DRT services are usually provided by multiple agencies. If all the agencies keep well maintained records of the DRT trips provided, the OD data can be easily transformed into trip generation. But when some agencies do not keep records of the DRT trips, the trip generation of the areas served by those agencies is more difficult to estimate and a model needs to be constructed to determine the relationship between trip generation and the demographics of the census tract. There are 10 human resource agencies in Tennessee. They all provide DRT services in their service areas. However, only one agency made their DRT ridership data available as the trip generation model is constructed based solely on this ridership data. The model shows that there is a relationship between the trip generation and the demographics of the census tract. This model is applied to areas without ridership data to estimate their trip generation.

4.1.1. Data description

Mid-Cumberland Human Resource Agency (MCHRA) keeps a record of all of the demand responsive trips it provides. Because the DRT services provided by MCHRA "is available to anyone regardless of age or income", it does not confine eligible users; thus, its users should be a representative sample of the DFRT users. Second, because of the nature of the DRT service, the trips can originate and terminate anywhere within the service area and can cross city and county boundaries. Its service area includes 12 counties: Cheatham, Dickson, Houston, Humphreys, Montgomery, Robertson, Rutherford, Stewart, Sumner, Trousdale, Williamson and Wilson. The DRT does not operate in Davidson County, meaning it does not pick up passengers from Davidson County. The map of the service area is shown in Figure 2. The service hours run from 6 a.m. to 6 p.m. Monday through Friday. A two-year dataset (July 1, 2009–June 30, 2011; fiscal year 2009–2010) was obtained and analyzed. The dataset contains an anonymous but unique passenger ID number, trip date, pick up address, pick up time, drop off address, drop off time, trip mileage, fare, passenger age, gender and trip purpose for all trips. There were 169 112 trips in fiscal year 2009 and 180 488 trips in fiscal year 2010. This dataset is used to construct a model to estimate trip generation based on known census data (e.g. population density, no-vehicle household density), which can then be applied to estimate demand in other places in Tennessee where actual data was not available. The trip generation of a census tract is calculated by counting how many trips originate in the census tract. Because the trips in the dataset are unlinked passenger trips, where a trip is originally produced (i.e. home) is not identified. To address this, trips are linked into trip tours and assigned to the origin of the original trip link in the rural area.

4.1.2. Trip linking

In each trip tour, there could be many trips. Only the origin of the first trip is the trip producing location. The origins of other trips in this trip tour are actually intermediate destinations of the traveler. Confusing them could lead to incorrect estimation of trip generation. An examination of the data

DOI: 10.1002/atr

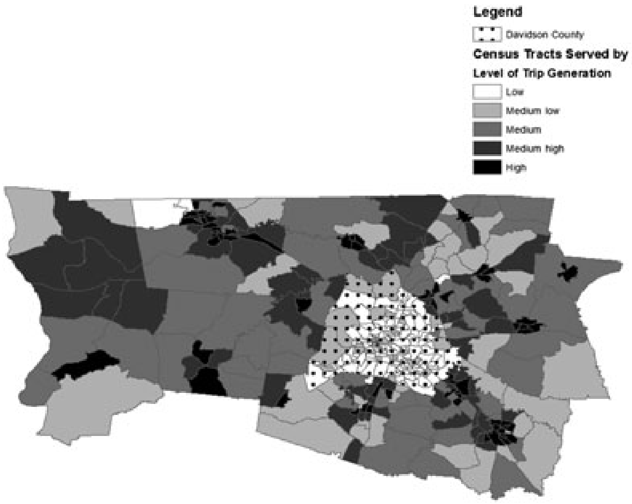

Figure 2. Trip generation of census tracts in MCHRA service area.

reveals that the connecting trips in a trip tour are usually performed within ten hours while two different trip tours are usually more than ten hours apart. A criterion to link trips into trip tours was developed: trips performed by the same person (indicated by matching rider ID) within 10 hours, with the drop off address of one leg matching the pick-up address of the following leg are linked together. The trip linking was performed using SAS PROC SQL.

In total, 349 600 trips in the dataset are linked into 190 914 trip tours. Among the 190 914 trip tours, 148 500 tours are composed of two trips, 4134 tours are composed of three trips and 965 tours are composed of four trips. The number of trip tours generating from each census tract were counted as inputs in the trip generation model.

4.1.3. Model selection

Because the trip tour count data are only available in the MCHRA service area (other service providers either do not keep a record of individual trips they provided or are reluctant to offer that information), a model was constructed to estimate the trip generation of census tracts in Tennessee outside of MCHRA's service area relying on relationships between trip generation and demographics. While it is acknowledged that the ridership of the transit is dependent on the quality of the service, cost and accessibility, only demographics of a census tract are used to estimate the trip generation so that the model could be easily applied to other areas in Tennessee. First, the number of trip tours generating from each census tract were counted from the last section. The trip generation rate (unit is trip tours per square mile) of each census tract is used as the dependent variable of the trip generation model. The trip generation rates of census tracts in MCHRA service area are shown in Figure 2. As noted above, Davidson County (Nashville) has very low trip generation because MCHRA does not pick up passengers there. Instead, those residents can rely on urban transit, paratransit and taxi service [40]. When applying the model, it is reasonable to assume that other big urban areas also have very low trip generation because they have much better transit services than rural areas. This is important when developing the model. In the figure, Davidson County is highlighted as the political boundary of the urban area (Nashville). For illustration, trip generation rates are divided into five quintile levels (0–0.17, 0.17–0.66, 0.66–6.03, 6.03–20.87, 20.87–449.70 trip tours per square mile per year). This model is applied to areas outside MCHRA's service area to estimate trip generation in all of Tennessee's census tracts.

The trip generation rate is represented as count data (number of passenger trip tours) divided by the area of the census tract. Poisson regression and negative binomial regression models are usually used to model count data. The less restrictive NB model is more widely used than Poisson regression because it does not assume the variance of the responsive variable is equal to the mean. In this dataset, the variance of the responsive variable is larger than the mean so the NB model is used. Often when analyzing count data, the number of exogenous zeroes is larger than would be expected under the

DOI: 10.1002/atr

NB model leading to a poor fit. This is the case in Davidson County. Zero-inflated negative binomial (ZINB) model that allows for excess exogenous zeros is more appropriate for this data.

In this case, a ZINB model [41] assumes that the outcome is a mixture of two different processes: (i) an urban county process that generates no trips; and (ii) a rural county process that generates trips that follow a count distribution which is a function of demographic variables. The urban/rural mixture component of the ZINB model is captured with a binary logit model while the count distribution follows a standard NB model. The expected count is a function of the two processes. In this study, the expected trip generation rate is defined as follows:

$$\begin{aligned} E(\mathit{trip\ generation\ rate}) = {} & P(\mathit{trip\ generation\ rate\ is\ zero}) \times 0 \\ & + P(\mathit{trip\ generation\ rate\ is\ not\ zero}) \\ & \times E(\mathit{trip\ generation\ rate} \mid \mathit{trip\ generation\ rate\ is\ not\ zero}). \end{aligned}$$

SAS 9.3 was used to construct the model. There are 397 census tracts in the MCHRA service area. A total of 300 observations were randomly chosen as the training sample, and the other 97 observations are used as validation sample. The data were tested to show that there are excess zeros in the data with p-value less than 0.01, which means that ZINB is required rather than standard NB. The density of total population, population of age 16 and over, population age 65 and over, people without employment, not in labor force, females, unemployed, total households, households with no vehicles available and households with income below poverty level served as the predictors of the trip generation model. A dummy variable indicating whether the census tract is located within the five urban counties also served as an independent variable. A stepwise variable selection procedure was used to select the best model.

Before settling on the best model, several different specifications were attempted. Trip generation (defined as the number of trip tours generating from each census tract) and the number of trip tours per thousand people were tested as dependent variables. The model using trip generation rate (the number of trip tours per square mile per year) as the dependent variable provided the model with the most predictive accuracy.

4.1.4. Model results

The best model has an R-square of 0.54 for the training sample and R-square of 0.42 for the validation sample. The model contains three variables: a dummy variable whether the census tract is in the large urban area county boundary, the log transformation of the density of no vehicle households in the census tract and the log transformation of the density of population aged 16 and above in the census tract. Then, all the observations including both training sample and validation sample were used to estimate the parameters of the three variables and the intercept. The results are shown in Table I.

The fitted model was used to estimate the trip generation of other census tracts in Tennessee. The trip generation map is shown in Figure 3 (a). The trip generation rates from low to high are 0.24–8.15, 8.15–17.06, 17.06–53.36, 53.36–899.40 and 899.40–8909.40 trip tours per square mile. For the most part, trip generation rates are high surrounding the large urban areas (labeled in white)

Table I. Trip generation model results.

Negative binomial regression part			Logistic regression part		
Variable	Coefficient	S.E.	Variable	Coefficient	S.E.
Intercept	−0.18	0.31	Intercept	−24.68*	0.28
If in the urban county	−5.56*	0.29	If in the urban county	25.28	0.00
Log of no vehicle household density	0.97*	0.12	—	—	—
Log of population aged 16 and above density	1.37*	0.29	—	—	—

[1]Dependent variable is the log of trip tours per square mile.
[2]* indicates significant at 0.05 level.

DOI: 10.1002/atr

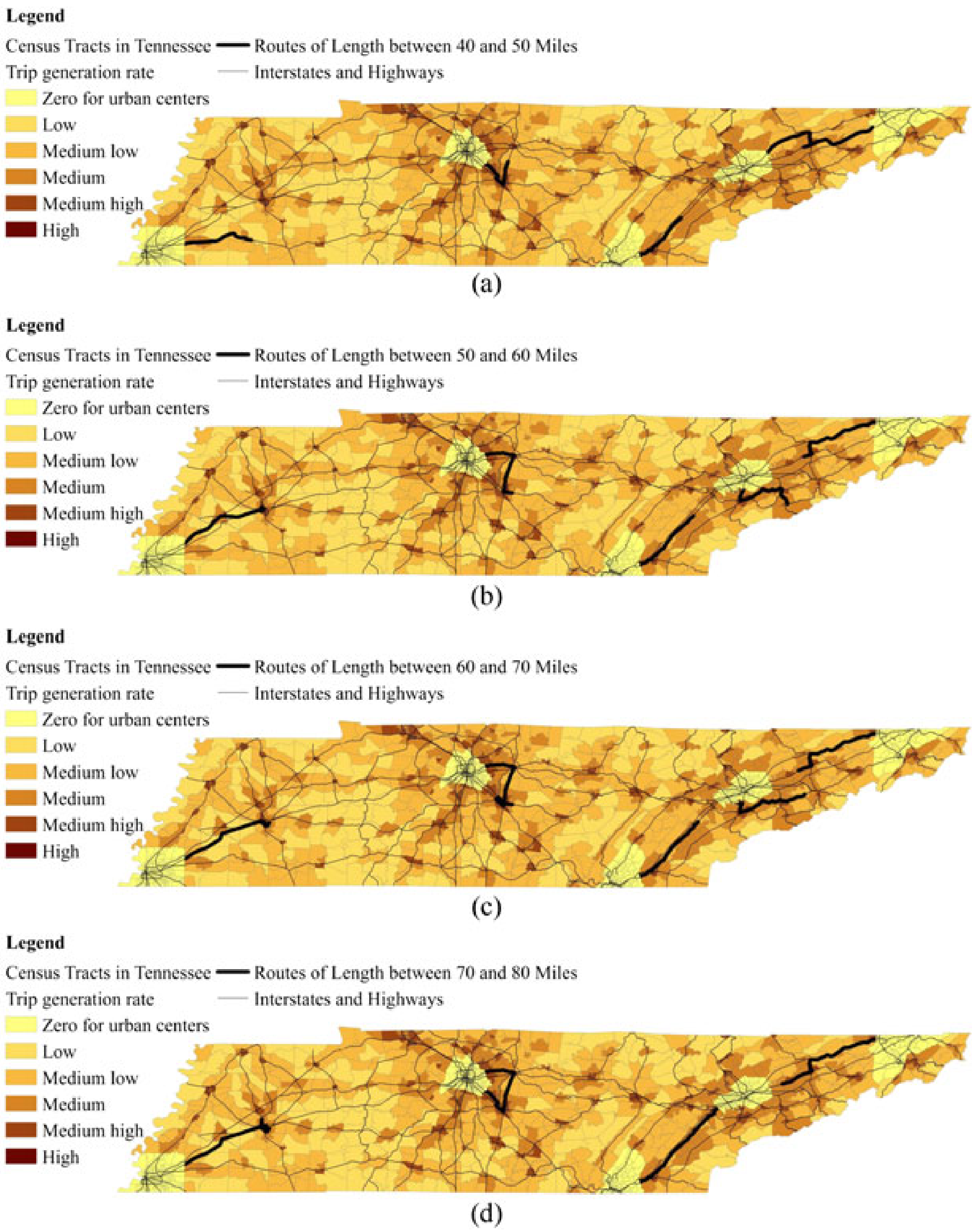

Figure 3. Cost-effective routes based on 0.75-mile buffer (trip generation rates from low to high are 0.24–8.15, 8.15– 17.06, 17.06–53.36, 53.36–899.40 and 899.40–8909.40).

primarily because of higher densities. The color of a census tract indicates the trip generation rate range where darker colors represent higher trip generation potential (i.e. transit compatible demographics and/or high population) and lighter colors represents low trip generation (i.e. transit incompatible demographics and/or low population). The trip generation rate represents the average number of DRT passenger trips originating from a unit area in a year, with units of passenger trips per square mile per year. It is calculated by aggregating the expected trips (trip generation) in each census tract and dividing the number of trips by the area of the census tract. All the generated trips are assumed to be attracted to urban centers [38]. The trip generation rate ranges from 0.01 to 889.40 trips per square mile per year, which are divided into five levels for visual distinction with each level representing the same number of census tracts. Again we note that census tracts in the urban areas are excluded from the DFRT service because people living in urban areas are generally served by urban transit.

4.2. Cost-effective routes selection

Five large urban area counties are identified as destinations, shown in Figure 3 (a) (five light color areas), because they contain major activity centers (e.g. hospitals) and operate urban public transport systems. From left to right, the urban counties and associated metropolitan areas are: Shelby County (Memphis), Davidson County (Nashville), Hamilton County (Chattanooga), Knox County (Knoxville) and Washington-and-Sullivan County (Kingsport).

Because most of the existing rural to urban transit routes in Tennessee are between 40 and 80 miles [42], the route length range is set to be between 40 and 80 miles. We choose one mile as the segment length (L_s) because the shortest length of a census tract in the study area is around one mile.

According to Koffman [6], 1.5 mile is sometimes used as the maximum deviation distance in addition to the 0.75-mile deviation. To understand how different width of the service area could influence the route locations and the percent of riders that could be served under different service area assumptions, both, 0.75-mile and 1.5-mile buffers are made on each side of a route to represent the two service areas. Those buffers are overlaid on the trip generation map to assess the trip generation potential. So, each route has two trip generation potentials corresponding to the two buffers.

4.3. Results discussion

The cost effectiveness of all routes is calculated and the best route connecting each urban county is picked for each of four distance categories (40–50 miles, 50–60 miles, 60–70 miles and 70–80 miles), which reflects budget availability. The best routes based on 0.75-mile buffer and 1.5-mile buffer are shown in Figures 3 and 4, respectively. Both figures reveal that the best routes of different lengths are sometimes serving different locations, indicating the best route location is sensitive to its design length and the demand density distribution served. Even given a demand density map, it could be hard for a transit planner to choose corridors with the most potential ridership. These maps illustrate the value of this route-finding methodology to find the best routes for different distance (budget) categories.

Comparing these best routes in Figures 3 and 4, their locations are different in some cases, suggesting that best route locations can be influenced by the width of the buffer. The strength of influence is determined by the demand distribution and the width of the service area.

Table II shows the normalized operating cost per trip, where Nashville's shortest route has the lowest operating cost per unit of demand and is set to 1.0. The range is high, with Memphis' worst performing route length being the shorter 40 to 50-mile route (8.98 times worse than Nashville's 40 to 50-mile route). Because of the relatively low rural population immediately outside Memphis, DFRT service must travel long distances to reach rural population centers. Thus, Memphis's cost effectiveness increases with longer routes, eventually reaching high trip generating areas. There is not a simple pattern between route length and cost effectiveness, because of the different distribution of high trip generation areas surrounding each city. For Chattanooga, the highest cost effectiveness lies in the 40 to 50-mile category. For Memphis, the highest cost effectiveness lies in the 60 to 70-mile category.

In general, Nashville routes have the lowest cost per passenger, followed by routes serving Knoxville and Kingsport. Routes serving Memphis have the highest cost per passenger because of the sparsely populated rural areas and dispersed transportation infrastructure.

To test the sensitivity of segment length (L_s), we tested segment lengths ranging from 1 to 10 miles in increments of one mile. Consistent with Table II, the length range of routes is between 40 and 80 miles. The most cost-effective route connecting to each urban county is selected. Figure 5 shows the normalized operating cost per trip for each route, where again Nashville's shortest route constructed by 1-mile segments has the lowest operating cost per unit of demand and is set to 1.0. Generally when segment length is short, the normalized operating cost per trip of routes is low. The 1-mile segment outperformed other segment lengths. If the computational capability is powerful, or the road network is simple, choosing a shorter segment length will result in a more precise and cost-effective solution.

The percent of Tennessee's service area covered and the estimated demand served by proposed routes is calculated and shown in Table III. The column "Trip Gen" shows the percentages of the estimated trip generation covered by the service area. The column "Area" shows the percentage of Tennessee area covered by the proposed route service. It should be noted that some of the best routes of different length ranges are close to each other or overlap. When their buffers overlap, the total buffer

DOI: 10.1002/atr

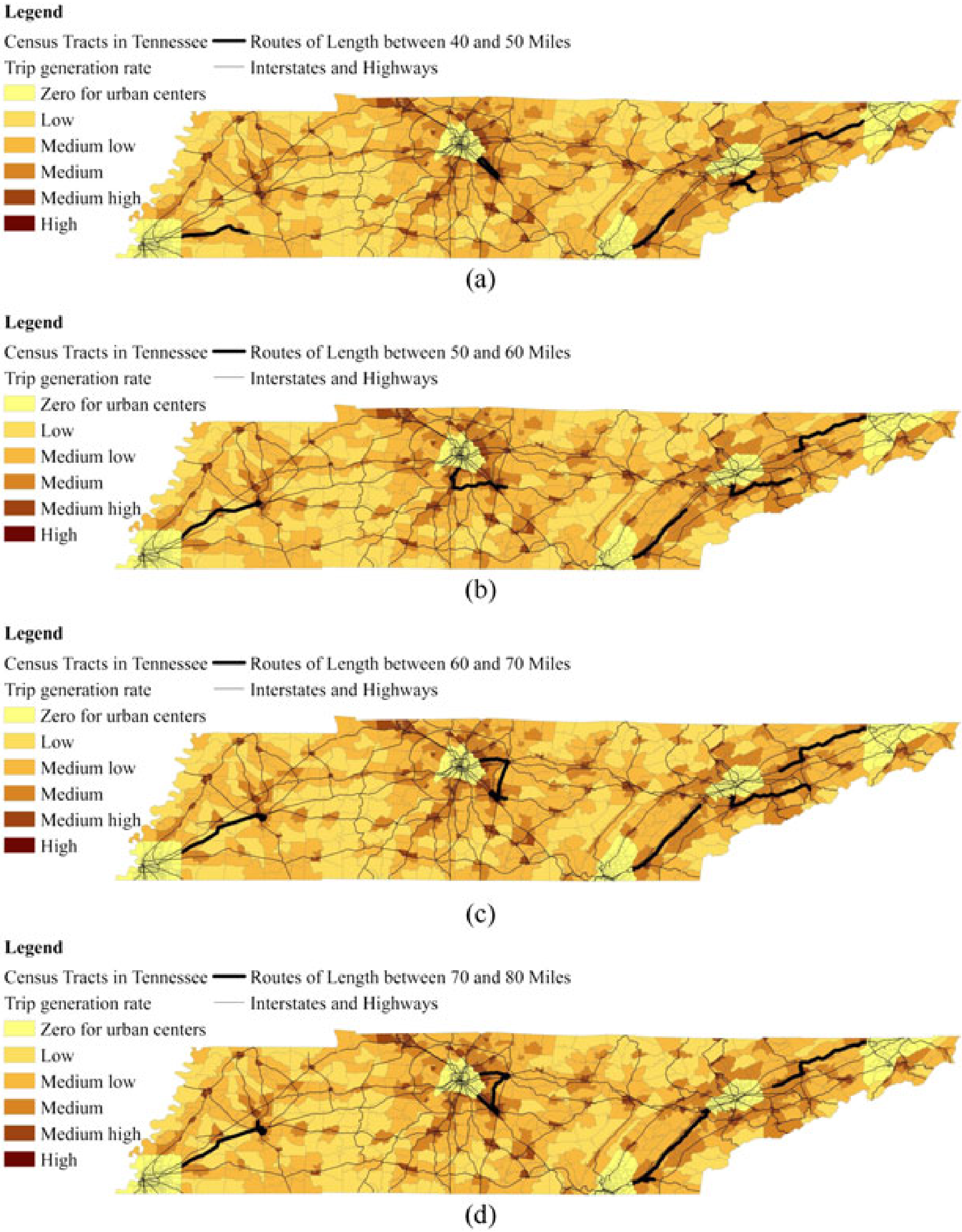

Figure 4. Cost-effective routes based on 1.5-mile buffer (trip generation rates from low to high are 0.24–8.15, 8.15– 17.06, 17.06–53.36, 53.36–899.40 and 899.40–8909.40).

area in the state is smaller. This is why the percent of area and estimated demand covered by routes between 70 and 80 miles is lower than those covered by routes between 60 and 70 miles. The column "Ratio" is calculated by dividing trip generation percentage by area percentage, indicating the percent of trip generation covered per unit percent of area. Because populations are generally located near transportation infrastructure, all of the developed routes follow the transportation infrastructure and cover more population than a random line across the state. We compare the most cost-effective and least cost-effective routes by comparing these population coverage ratios. All ratio values in the table are above one, with the highest value above 4, indicating that with 1% of the area coverage the proposed routes cover more than 4% of the trip generation potential. The ratios of most cost-effective routes are several times larger than the least cost-effective routes. It shows that this method can reveal high demand areas and cost-effective routes. Another application of this ratio is that it can be compared

DOI: 10.1002/atr

Table II. Lowest normalized operating cost per trip (0.75-mile deviation).

Range (miles)	Chattanooga	Kingsport	Knoxville	Memphis	Nashville
40–50	2.20	2.58	1.41	8.98	1.00
50–60	2.36	2.44	1.42	3.63	1.28
60–70	2.77	1.94	1.31	2.33	1.04
70–80	2.55	2.23	1.79	2.63	1.40

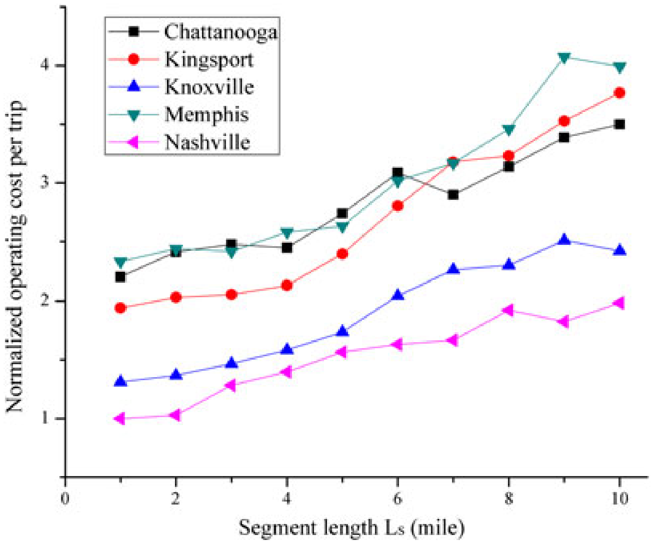

Figure 5. Relationship between segment length and lowest normalized operating cost per trip of routes between 40 and 80 miles (0.75-mile deviation).

Table III. Percent of area and estimated trip generation covered by routes based on 0.75-mile buffer.

Dist. range (mile)	Most cost-effective routes			Least cost-effective routes		
	Trip gen	Area	Ratio	Trip gen	Area	Ratio
40–50	2.32%	0.61%	3.80	0.83%	0.52%	1.60
50–60	3.56%	0.84%	4.24	1.26%	0.68%	1.86
60–70	4.20%	0.93%	4.52	1.58%	0.89%	1.78
70–80	3.21%	0.79%	4.06	1.94%	1.12%	1.73

to other ratios to reveal which routes are more effective in terms of covering more potential demand within a unit service area. For example, a policy objective could be to provide the highest ratio of demand coverage to area coverage, regardless of which city is served. This policy would likely eliminate the Memphis routes in favor of more or longer Nashville routes. For routes proposed, based on 0.75-mile buffer, the routes in the 60 to 70-mile category and 70 to 80-mile category have the highest ratios suggesting that those routes could serve more trips per unit service area. While not considering budget constraints, longer routes are preferred to shorter routes because they have higher ratio of demand coverage to area coverage.

5. CONCLUSION

In this study, a heuristic method of inventorying all possible DFRT routes and choosing the best route based on the lowest operating cost per passenger is proposed. Government agencies could use the methods and results of this study to determine the most cost-effective routes under budget constraints or policy objectives. The private sector could use these results for service planning and revenue estimation.

DOI: 10.1002/atr

In this study, we focus on identifying the highest potential demand per unit of service (a mile of travel). This parameter was extended by applying an average cost of service (cost per mile) to identify cost-effective routes, because this objective is highly valued by government agencies and service providers [7]. However, there are many possible ways to choose near-optimum routes, depending the administrator's objectives. Those include cost effectiveness (ridership per unit cost) [7], social equity considerations (number of carless households reached) [43], and political considerations (number of jurisdictions covered) [44], among other potential goals. The selection of the objective function is based on the needs of the transit agency. The proposed method is quite general and could be modified to handle any of the above mentioned objectives. Furthermore, a ranking could be developed for each of the three objectives and the routes selected can be those that perform well on all three rankings. Alternatively, the three objectives could be combined into a more comprehensive objective using a weighted summation. The route selection could then be based on this more comprehensive objective.

There are several assumptions and approximations used in this study. First, because the DFRT transports rural residents to urban centers, all proposed routes start from the urban centers. The DFRT usually connects to the meaningful destinations in urban centers, such as hospitals, airports, shopping malls and so on. But how the DFRT connects to the destinations in the urban centers is not within the scope of this study. There are many ways to define the boundary of the urban centers. In this study, a simple approach of using the county border where the urban center located is used to define the boundary. Other urbanized area boundaries could be used. Alternatively, the urban transit network coverage could be used to define this boundary. This approach would assist in the development of a multimodal transportation system and avoid redundant service. However, in the scope of a statewide study, we assume that the main county boundaries are appropriate approximations of urban cores. Next, because the input data for the trip generation model is the DRT ridership data, we assume that the same profile of rider who uses the DRT service will also use rural DFRT service. It is unclear if DFRT will have higher ridership because it introduces some schedule constraints but also requires less advance planning (i.e. scheduling pickups in advance). Although the ridership is dependent on the quality of the transit service and cost, they are not taken into consideration in estimating the trip generation. Our trip generation model only identifies potential demand based on constant and consistent service parameters across alternative routes. The model should only be used to assist in relative rank and prioritization. Because this study focuses on initial route planning, the allowed maximum deviation time is not considered in the route development process. Different sized deviations could be used with this method and the relative merits of those deviations can be compared to costs. The results should be used cautiously and are best applied as a relative ranking tool rather than an absolute ridership estimation tool. Once ranked, more specific, route-level analysis could be conducted to assess feasibility of individual routes. Because the purpose of this study is to provide a methodology to select the cost-effective DFRT route, the capital cost is also not considered in this study. Despite its limitations, this study provides much new information and proposes methodologies to evaluate and design rural transit. Future automated sketch-planning software tools can use the methods proposed here to quickly identify best routes for rural DFRT service.

ACKNOWLEDGEMENTS

This study was funded by Tennessee Department of Transportation, Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University and the Fundamental Research Funds for the Central Universities (Grant No. 2682015BR012). The authors appreciate the support of TDOT staff, Dironna Belton, Paula Shaw, Sherri Carroll, John Reed, Stephanie Vincent and Sandra Hoff. The authors also appreciate the support from MCHRA.

REFERENCES

1. Nutley S. Indicators of transport and accessibility problems in rural Australia. *Journal of Transport Geography* 2003; **11**(1): 55–71. DOI:10.1016/S0966-6923(02)00052-2.
2. Kamruzzaman M, Hine J. Participation index: a measure to identify rural transport disadvantage? *Journal of Transport Geography* 2011; **19**(4): 882–899. DOI:10.1016/j.jtrangeo.2010.11.004.

3. Wilson H, Sussman I, Wang H Higonnet B. *Scheduling Algorithms for Dial-a-Ride Systems*. Urban Systems Laboratory: MIT: Cambridge, MA, 1971.
4. Arrillaga B, Mouchahoir GE, Demand-responsive transportation system planning guidelines, in MITRE technical report. Urban Mass Transportation Administration: Washington, DC, 1974.
5. Cole LM, Tomorrow's transportation: new systems for the urban future. 1968.
6. Koffman D. *TCRP Synthesis 53: Operational Experiences with Flexible Transit Services—A Synthesis of Transit Practice*. Transportation Research Board: Washington, DC, 2004.
7. Potts JF, Marshall MA, Crockett EC, Washington J, A guide for planning and operating flexible public transportation services. 2010.
8. Qiu F, Li W Zhang J. A dynamic station strategy to improve the performance of flex-route transit services. *Transportation Research Part C: Emerging Technologies* 2014; **48**: 229–240. DOI:10.1016/j.trc.2014.09.003.
9. Quadrifoglio L, Dessouky MM Ordóñez F. Mobility allowance shuttle transit (MAST) services: MIP formulation and strengthening with logic constraints. *European Journal of Operational Research* 2008; **185**(2): 481–494. DOI:10.1016/j.ejor.2006.12.030.
10. Zhao J, Dessouky M. Service capacity design problems for mobility allowance shuttle transit systems. *Transportation Research Part B: Methodological* 2008; **42**(2): 135–146. DOI:10.1016/j.trb.2007.07.002.
11. Ceder A. Integrated smart feeder/shuttle transit service: simulation of new routing strategies. *Journal of Advanced Transportation* 2013; **47**(6): 595–618. DOI:10.1002/atr.185.
12. Quadrifoglio L, Dessouky MM Palmer K. An insertion heuristic for scheduling mobility allowance shuttle transit (MAST) services. *Journal of Scheduling* 2007; **10**(1): 25–40.
13. Quadrifoglio L, Hall RW Dessouky MM. Performance and design of mobility allowance shuttle transit services: bounds on the maximum longitudinal velocity. *Transportation Science* 2006; **40**(3): 351–363.
14. Quadrifoglio L, Dessouky MM. Sensitivity analyses over the service area for mobility allowance shuttle transit (MAST) services, in Computer-aided Systems in Public Transport. 2008, Springer. p. 419-432.
15. Durvasula PK, Smith BL, Turochy RE, Brich SC, Demetsky MJ, Peninsula Transportation District Commission Route Deviation Feasibility Study. 1998, Charlottesville, Virginia.
16. Rosenbloom S, Service routes, route deviation, and general public paratransit in urban, suburban, and rural transit systems. 1996.
17. Lu X, Yu J, Yang X, Pan S Zou N. Flexible feeder transit route design to enhance service accessibility in urban area. *Journal of Advanced Transportation* 2015. DOI:10.1002/atr.1357.
18. Crainic TG, Malucelli F, Nonato M. A demand responsive feeder bus system. in CD-ROM of 7th World Congress on Intelligent Transport Systems. 7WC-ITS, Torino, Italia. 2000.
19. Crainic TG, Malucelli F, Nonato M Guertin F. Meta-heuristics for a class of demand-responsive transit systems. *INFORMS Journal on Computing* 2005; **17**(1): 10–24.
20. Malucelli F, Nonato M, Crainic TG, Guertin F, Adaptive memory programming for a class of demand responsive transit systems. 2001: Springer.
21. Errico F, Crainic TG, Malucelli F, Nonato M, The design problem for single-line demand-adaptive transit systems. 2011: CIRRELT.
22. Fu L. Planning and design of flex-route transit services. *Transportation Research Record: Journal of the Transportation Research Board* 2002; **1791**(1): 59–66. DOI:10.3141/1791-09.
23. Broome K, Worrall L, Fleming J Boldy D. Evaluation of flexible route bus transport for older people. *Transport Policy* 2012; **21**: 85–91.
24. Federal Transit Administration. National Transit Database. 2012 [cited 2015 October 24]; Available from: http://www.ntdprogram.gov/ntdprogram/.
25. Fravel FD, Barboza R, Quan J, Sartori J, TCRP report 147: toolkit for estimating demand for rural intercity bus service. 2011, Washington, DC: Transportation Research Board. 64, 124 p.
26. Fravel F, Tauchen H Gilbert G. Regulatory policy and economies of scale in the U.S. intercity bus industry. *Transportation* 1982; **11**(2): 173–187. DOI:10.1007/BF00167930.
27. Yang H, Cherry C. Statewide rural–urban bus travel demand and network evaluation: an application in Tennessee. *Journal of Public Transportation* 2012; **15**(3): 97–111. DOI:10.5038/2375-0901.15.3.6.
28. Errico F, Crainic TG, Malucelli F Nonato M. A survey on planning semi-flexible transit systems: Methodological issues and a unifying framework. *Transportation Research Part C: Emerging Technologies* 2013; **36**(0): 324–338. DOI:10.1016/j.trc.2013.08.010.
29. Scott RA, Demand responsive passenger transport in low-demand situations, N.Z.T. Agency, Editor. 2010.
30. Smith BL, Demetsky MJ Durvasula PK. A multiobjective optimization model for flexroute transit service design. *Journal of Public Transportation* 2003; **6**(1): 81–100.
31. Alshalalfah B, Shalaby A. Development of important relationships for the planning of flex-route transit services. in Transportation Research Borad 89th Annual Meeting. 2010. Washington, DC.
32. Alshalalfah B, Shalaby A. Feasibility of flex-route as a feeder transit service to rail stations in the suburbs: case study in Toronto. *Journal of Urban Planning and Development* 2011; **138**(1): 90–100.
33. Qiu F, Li W Haghani A. A methodology for choosing between fixed-route and flex-route policies for transit services. *Journal of Advanced Transportation* 2015; **49**(3): 496–509. DOI:10.1002/atr.1289.
34. Qiu F, Li W Haghani A. An exploration of the demand limit for flex-route as feeder transit services: a case study in Salt Lake City. *Public Transport* 2015; **7**(2): 259–276. DOI:10.1007/s12469-014-0097-9.

35. Nourbakhsh SM, Ouyang Y. A structured flexible transit system for low demand areas. *Transportation Research Part B: Methodological* 2012; **46**(1): 204–216. DOI:10.1016/j.trb.2011.07.014.
36. Mulley C, Daniels R. Quantifying the role of a flexible transport service in reducing the accessibility gap in low density areas: a case-study in north-west Sydney. *Research in Transportation Business & Management* 2012; **3** (0): 12–23. DOI:10.1016/j.rtbm.2012.04.006.
37. Qiu F, Shen J, Zhang X An C. Demi-flexible operating policies to promote the performance of public transit in low-demand areas. *Transportation Research Part A: Policy and Practice* 2015; **80**: 215–230. DOI:10.1016/j.tra.2015.08.003.
38. Yang H, Cherry CR, Use characteristics and demographics of rural transit riders: a case study in Tennessee. Transportation Planning and Technology, 2016. Accepted.
39. Balog JN. *Guidebook for Attracting Paratransit Patrons to Fixed-Route Services*, vol. **24**, Transportation Research Board: Washington, D.C., 1997.
40. Mid-Cumberland Human Resource Agency. Mid-cumberland public transit. 2013 [cited 2015.
41. Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* 1992; **34**(1): 1–14. DOI:10.2307/1269547.
42. Yang H, Rural to urban intercity transit user characteristics analysis, Demand Estimation and Network Design. 2013.
43. Foth N, Manaugh K El-Geneidy AM. Towards equitable transit: examining transit accessibility and social need in Toronto, Canada, 1996–2006. *Journal of Transport Geography* 2013; **29**: 1–10. DOI:10.1016/j.jtrangeo.2012.12.008.
44. Stommes ES, Brown DM. Transportation in rural America: issues for the 21st century. *Community Transportation* 2002; **20**(4): 2–10.

DOI: 10.1002/atr