International Journal of Geographical Information Science
Vol. 24, No. 3, March 2010, 347–364

Taylor & Francis
Taylor & Francis Group

A new method for determining the population with walking access to transit

S. Biba[a], K.M. Curtin[b]* and G. Manca[b]

[a]*Department of Geography, University of Texas at Dallas, Richardson, TX, USA;* [b]*Department of Geography and GeoInformation Science, George Mason University, Fairfax, VA, USA*

(*Received 6 September 2008; final version received 20 November 2008*)

The use of geographic information systems in determining transit service areas has not progressed far beyond simple buffering operations even though there is widespread capability to analyze network walking distances in conjunction with demographic, cadastral, and land-use data sets. This article presents a method for determining the population with walking access to bus stop locations using the spatial and aspatial attributes of parcels and the network distances from parcels to bus stop locations. This parcel-network method avoids the well-known and unrealistic assumptions associated with the existing methods and reduces overestimation of the population with access to transit, resulting in improved spatial precision and superior inputs to transit service decision-making processes. Comparisons of the parcel-network method, the buffer method, and the network-ratio method are made in a study area within the Dallas metropolitan area. The novel integration of cadastral data with network analysis in our method holds promise for research in many areas of geographic information science.

Keywords: walking distance; transit; transportation; parcel; network

1. Introduction

The use of geographic information systems (GIS) in determining transit service areas has not progressed far beyond simple buffering operations even though GIS have the capability to analyze network walking distances in conjunction with demographic, cadastral, and land-use data sets. It has long been recognized that GIS are a significant tool for transportation modeling due to their ability to realistically model linear and network features within such systems (Miller and Shaw 2001, Curtin 2007, 2008). There has been a call in the literature to give greater attention to the digital spatial representation (Goodchild *et al.* 2007) of both transit-network elements (Kwan *et al.* 2003) and the demand for transit service (Horner and Murray 2004).

In the context of transportation, GIS have been employed in vehicle routing (Chang *et al.* 1997, Osegueda *et al.* 1999, Ioannou *et al.* 2002, Tarantilis *et al.* 2004), vehicle scheduling (Weigel and Cao 1999, Calvo *et al.* 2004), travel demand modeling (Miller and Storm 1996), complex trip planning (Karimi *et al.* 2004), transportation hazard analysis (Lepofsky *et al.* 1993), the determination of transit coverage areas (O'Sullivan *et al.* 2000), and for many other types of transportation-related analyses. Included among these analyses is the use of GIS to determine the service areas – and the associated populations within those areas – for

*Corresponding author. Email: curtin@gmu.edu

ISSN 1365-8816 print/ISSN 1362-3087 online

DOI: 10.1080/13658810802646679
http://www.informaworld.com

transit routes in order to perform planning and operational functions (Dallas Area Rapid Transit 1998, Dallas Area Rapid Transit 2002, Federal Highway Administration 2002, Wu and Hine 2003). Unfortunately, since the methods used within the GIS consistently overestimate the population with access to transit (O'Neill *et al.* 1992, Zhao *et al.* 2003), planners and decision makers lose confidence in the ability of GIS to provide reliable answers to policy questions, particularly those questions concerning the changes in accessibility that will be achieved by changes in service.

In order to address this concern, this research presents an improved method for determining the population with walking access to transit facilities – the parcel-network method. This method employs high-quality cadastral data with network functionality to determine walking distances to transit facilities in order to more precisely estimate the population with access to transit. As is shown below, this method provides more conservative population estimates than the methods currently in use, it produces summary statistics and output data sets that cannot be generated through the other methods, and it provides a more flexible basis for further refinement of the transit forecasting and planning process. Although our method is designed for a transportation application, the novel integration of cadastral data with network analysis in our method holds promise for research in many areas of geographic information science. Section 2 reviews the literature in this area, followed by a description of the steps involved in building and using the parcel-network method. This is followed by an application of this new method using a set of data from the north Dallas region. A quantitative comparison of the results is made for the buffer, network-ratio, and parcel-network methods. Additional research opportunities are discussed in Section 6.

2. Literature review

The literature concerning the determination of transit service areas falls into two main areas. First, there is research concerning the existing methods for determining transit service areas, specifically the buffer (or area-ratio) method and the network-ratio method. Second, all methods for determining transit service areas depend on estimating walking distance for transit users. We review both of these research areas below.

2.1. Existing methods for determining transit service areas

It is well-recognized among transportation planners that access to transit service is an important measure of service performance (Murray 2001) and that improvements in access can lead to significant changes in urban form (Murray *et al.* 1998). Historically, transit service area determination has been implemented within GIS by creating a distance buffer around the transit route, or stops along that route (O'Neill *et al.* 1992, Hsiao *et al.* 1997, Peng *et al.* 1997, Ayvalik and Khisty 2002), and by estimating the population within that buffer based on an overlay of census polygons (Peng and Dueker 1995). The buffer method (or area-ratio method) assumes a uniform distribution of population within the census polygon, which may not be true. Moreover, this method implies that the entire population within the buffer has walking access to the transit route. The errors associated with polygon overlays of demographic data are well documented (MacDougall 1975, Newcomer and Szajgin 1984, Chrisman 1987, Veregin 1989) and must be considered when buffers of transit facilities are overlain on population summary polygons. Moreover, the buffer method has been shown to consistently overestimate the population within the service area since the actual walking distances within the buffer are greater than the Euclidean distances used to generate the buffer (O'Neill *et al.* 1992). A comparison of Euclidean versus network distances between census polygon centroids and transit features produced much smaller estimates of population coverage with the network distance measure (Horner and Murray 2004). Additionally, the magnitude of the overestimation is a function of the size of the analysis zone. That is, there are scale effects related to the size of the analysis zone that influence the size of the population that will be assumed to have access to the transit facilities within that zone. It has been shown that scale and choice of digital representation can dramatically alter the results of GIS-based transit studies (Horner 2004), and it has been recognized that parcel-based methods would provide the most disaggregated (and presumably the most accurate) source for transit demand data if population levels were associated with the parcels (Handy and Niemeier 1997, Horner and Murray 2004, Lee 2005). It has long been hoped that the ability to perform such disaggregate data analysis in a GIS would lead to advances in transportation analysis (Hanson and Schwab 1995, Nyerges 1995).

O'Neill *et al.* (1992) used a refined method – termed the network-ratio method – to more accurately measure accessibility to transit services. This method considered the total length of the street network within analysis zones surrounding a transit route and the length of the streets within those zones that are also within a specified network distance from the transit stops. The formula for computing population with access to transit with the network-ratio is as follows:

$$A_i = \frac{W_i}{M_i} \times P_i$$

where:

A_i = The population in analysis zone i with access to transit
M_i = The total length of street network in analysis zone i
W_i = The length of the street network within walking distance to transit in zone i
P_i = The total population of zone i.

Although the network-ratio method eliminates the error associated with the assumption of uniform population distribution over census polygons, it does assume that population along a street is proportional to street length and that there is a uniform distribution of population on every street. This assumption is particularly weak in mixed residential zones or zones with retail, industrial, and recreational activities, which are precisely the kinds of areas that are likely to have transit routes. Moreover, when streets serve as the boundaries of analysis zones, it is unclear to which zone they should be assigned for the computation of the network-ratio values. This is always the case when census polygons are used as the analysis zones, since census blocks are required to have at least one side defined by a road feature, and many census polygons at every level of the hierarchy are defined by roads on all sides (US Census Bureau 2005). Lastly, the network-ratio method includes all roads within walking distance of a transit facility in the computation of the population with access. When highways and their associated frontage roads and off-ramps are included, there can be substantial error in the population estimates. In fact, research has shown that these limitations of the network-ratio method lead to consistent overestimation of the population with walking access to transit although the errors are not as large as those seen with the buffer method (Zhao *et al.* 2003).

2.2. *The importance of walking distance*

It is recognized that the problem of population overestimation in the context of transit access is a function of the inability to accurately model the population that is within a reasonable walking distance to the transit feature. Research has found that the spatial accessibility to the transit feature is the primary determinant of transit use and only in the presence of such accessibility will a user consider other factors such as cost, comfort, security, or other factors (Beimborn *et al.* 2003). One travel survey in Orange County, California, found that 80% of bus riders walked from their trip origins to bus stops and 90% of the riders walked to their destinations after exiting the bus (Hsiao *et al.* 1997). Among transit planners and researchers, it is nearly universally accepted that a 400-m (¼-mile) walking distance is the maximum distance that most potential users are willing to travel to reach their nearest transit stop (O'Neill *et al.* 1992, Hsiao *et al.* 1997, Phillips and Edwards 2002). Research has found that transit use diminishes quickly beyond 91 m (300 ft) and vanishes after 580 m (0.36 miles) (Zhao *et al.* 2003). Other studies have found that for every additional 500 m (1640 ft) from a station, the probability that an individual will walk to transit decreases by 50% (Loutzenheiser 1997), and similarly every 10% increase in walking distance results in a 10% decline in transit use (Dill 2003). Still another study finds that walking is the dominant mode up to 1600 m (eight blocks) in urban settings (Ayvalik and Khisty 2002). While particular measures of the impact of walking distance on transit use may differ, as will the results of traffic surveys in different areas, there is no dissent in the belief that walking distance is a primary factor in the decision to use transit. Only with the presence of a transit facility within walking distance can characteristics of the riders themselves, such as gender, race, age (Burkhardt 2003), vehicle availability, and employment type (Dill 2003), or environmental factors such as transit type (O'Sullivan and Morrall 1996), the availability of sidewalks, street crossings, and transit amenities (Evans *et al.* 1997), be investigated for their influence on transit use. It is important to distinguish between research designed to determine the level of use of transit facilities and research designed to determine access to transit facilities. This article is primarily concerned with the latter, as it is clear that good estimates of the population with access to transit are an important input to estimates of transit use.

The ability to precisely measure walking distances to transit facilities has been elusive given the large number of possible walking paths for the population and given the quality of network data available for analysis. It has been recognized in the literature that an increased focus on pedestrian accessibility (Evans *et al.* 1997) and analysis of improvements to sidewalks or street patterns could assist in the search for transit improvement strategies (Hsiao *et al.* 1997). In the absence of such data, several works (Kuby *et al.* 2004, Upchurch *et al.* 2004) have included walking distance via a raster-based approach. The walking distance is determined from each cell that is not part of the network to its closest network node. This gives essentially a Euclidean distance measure that does not take pedestrian barriers into account.

The use of Euclidean distance as a surrogate for walking distance is inappropriate given the clear difference between that measure and network distance. Additionally, barriers (fences, walls, creeks, buildings, large or busy roadways, etc.) frequently exist on the landscape that prevent walking along the Euclidean shortest path to a transit facility (Dill 2003). Although information on these barriers is not included in network databases and the time and expense required to collect such information is prohibitive, some way of recognizing the increased walking distances created by such barriers must be included in a model designed to measure transit accessibility. A reasonable assumption is that no significant barriers to pedestrian travel will exist along roads designed to be used by pedestrians to access transit. Zhao et. al. (2003) took advantage of this by utilizing the addressed street as the point of access to the walking network, and in the parcel-network model we follow that example.

In summary, there have been clear calls in the literature for the use of disaggregate population data (specifically parcel-level data) for transit demand modeling, there have been documented difficulties (particularly with overestimation of population) with the traditional methods of measuring population with access to transit, and it is uncontested that walking distance is the appropriate metric to use in these studies. To the best of our knowledge, no existing research has simultaneously addressed these issues. Perhaps most importantly, this review of the literature demonstrates that the research questions addressed in this article are not germane only to the topic of transit ridership or transportation geography more generally. Rather, this research deals with a broad range of topics across the breadth of GIScience, including the effects of changing scale, issues of fundamental spatial representation, concerns regarding areal interpolation, and challenges in network analysis and routing behavior. This research addresses these issues through the definition and implementation of the parcel network method.

3. The parcel-network method

The parcel-network method exploits both the superior precision of parcel (or cadastral) databases and the network functionality that has been increasingly incorporated into GIS. With the acceptance of GIS among local agencies, the locations of parcels have become increasingly accurate due to the desire of local governments to improve their ability to process land information and to simplify the processes of land transfer and taxation (Reeve 1984, Werle 1984, Wilcox 1984). Moreover, cadastral databases also include attribute information such as street addresses and land use or zoning categories (residential, commercial, etc.). These newly developing databases hold the potential to greatly enhance research efforts in areas such as land use planning (Walker *et al.* 2007, Chapin *et al.* 2008), conservation (Strager and Rosenberger 2007, Wallace *et al.* 2008), areal interpolation of population (Xie 2006), address geocoding (Christen *et al.* 2006, Zandbergen 2008), environmental health (Miranda *et al.* 2002), transportation applications (Furth *et al.* 2007), and in many other applications where disaggregate spatial samples of population must be selected (Lee *et al.* 2006). For the purpose of this research, parcels provide a level of spatial accuracy regarding the population that has not previously been exploited for transit access studies and that allows detailed modeling of walking access to transit facilities. The parcel-network method differs from traditional methods of determining access to transit in that it looks outward from the population locations (the parcels) to the transit features rather than looking out from the transit features and making assumptions about what population lies within some distance of those features.

Implementing the parcel-network method requires four major steps, some of which have several sequential processes (Figure 1). The first task is to apply demographic characteristics to the parcels; the second task creates a walking network from parcels to the transit facilities; the third task is to compute network walking distances from each parcel centroid to its nearest bus stop; and finally analysis is performed to assess the population that can access the transit facility across the walking network. These tasks represent a novel combination of parcel-based attribute imputation with network analytic techniques. The following subsections describe each of these four tasks in more detail.

3.1. Demographic attribute transfer

If the building footprints are not included in the cadastral database (as was the case in our data), there is no way to know the exact orientation of the structures or position of the

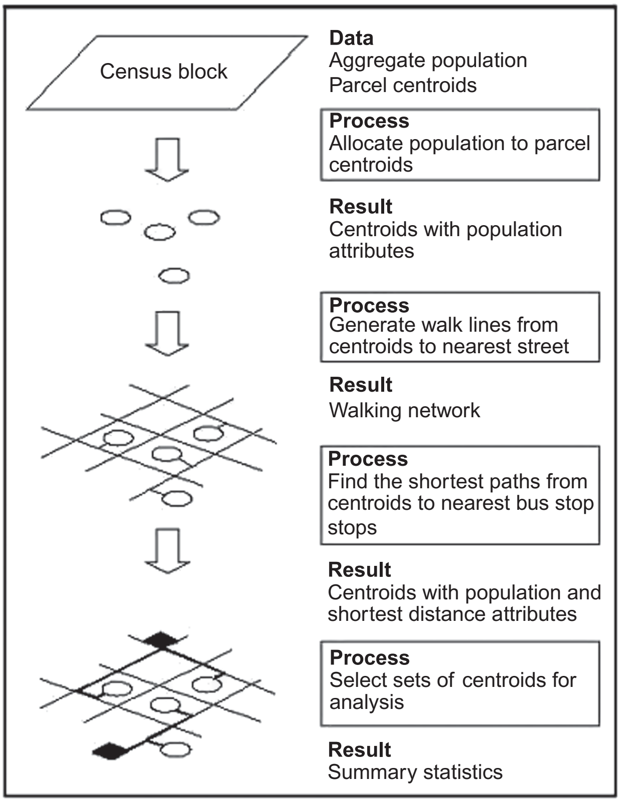

Figure 1. The parcel-network method.

building entrances within the parcel. Some point must be chosen as the starting point from which a pedestrian would leave the parcel to walk to the bus route. The centroid of the parcel is a reasonable representation of that starting point. Thus, the first step in the parcel-network method is to generate centroid points for all parcel polygons in the study area. Although the centroid could fall outside of the parcel polygon if the shape is irregular, in practice there are very few residential polygons where this is the case. Pertinent attribute information (based on the census blocks or other supplementary data) can then be associated with the parcel centroids. It is imperative to note that this step highlights the flexibility of the parcel-network method in that it allows the user a great deal of latitude in choosing how they model the characteristics of the population under consideration. That is, the population of employees at commercial or industrial parcels could be considered, the population of students at parcels containing educational institutions could be an input, or the residential population could be of primary interest.

Although any attribute information could be used to analyze the characteristics of the population with (or without) access to transit, in the research presented here, the object was to identify the total residential population with access to transit. Therefore the following steps were taken: (1) the appropriate census block identifier was associated with each parcel centroid in order to access total population numbers, (2) the total number of dwelling units per census block was computed, (3) the population per dwelling was computed (census block population/dwelling units) for the census block, (4) the population per dwelling unit was transferred back to the parcel centroids, and (5) the estimated population was computed for each parcel. It is important to note that each parcel (and its centroid) is associated with a variable number of dwelling units. A parcel containing one single-family home would have only one dwelling unit associated with it. A parcel with multiple dwelling units (such as an apartment building) would be assigned a population in proportion to the total number of dwelling units in the census block. The result is a set of parcel centroids that now have reasonable population attributes. Although the total population associated with the census blocks was the only demographic information used in this research, a wealth of other data could be associated with parcel centroids in this way (income, car ownership, etc.) if a particular transit research project demanded it.

The parcel-network method increases spatial precision because it relies on a discrete rather than a continuous allocation of attributes, unlike both the buffer method and the network-ratio method. The network-ratio method allocates along linear length rather than area and therefore results in less overestimation than the buffer method. The parcel-network method improves even further by allocating population to discrete points (parcel centroids), and no assumptions about population distribution over street length or analysis zone area are made. The parcel-network method allocates population uniformly among dwelling units at each parcel centroid. Although we recognize that the assumptions inherent in our method would be unreasonable in areas where there were substantial numbers of vacant parcels, or where homes had been subdivided in ways that are not captured in the cadastral database, we believe the population allocation method we present could only be improved by using actual counts obtained through expensive and time-consuming household surveys.

3.2. Creating a walking network

Since the street network does not connect to the parcel centroid points that now have associated attribute information, new links in that network must be generated in order to find the walking paths that pedestrians would follow. In order to do so, an automated process (a script) was employed within the GIS to select each centroid, identify the point on the street network closest to that centroid, and generate a new network link connecting those two points. The result is a set of parcel centroids with population attributes and network connections to the walking network. Due to odd-shaped parcels and corner parcels, a centroid can occasionally be closer to a street that is not the addressed street for that parcel. That is, when residential parcels back up to an arterial street, the centroid may be closer in terms of Euclidean distance to the arterial even though there is no direct pedestrian access to the arterial due to walls or other barriers (Figure 2). In such cases, the walk line was forced to connect with the addressed street. If one uses the addressed street to generate the walking path from the centroid, it is virtually certain there will not be a barrier for a pedestrian.

3.3. Computing walking distances

With the parcel centroids connected to the street network, an origin for a walking trip to transit has been defined. Still undetermined are the destinations of those trips and the paths along which the walking trip will take place. For the purposes of this research, these destinations are potential bus stops along the existing bus routes. In our study area, potential bus stops are located at all intersections along a bus route. If the parcel-network method is to be used for new route planning, the entire route may initially represent all points of access, with bus stops then strategically located in an optimization process such as that described by Murray (2001).

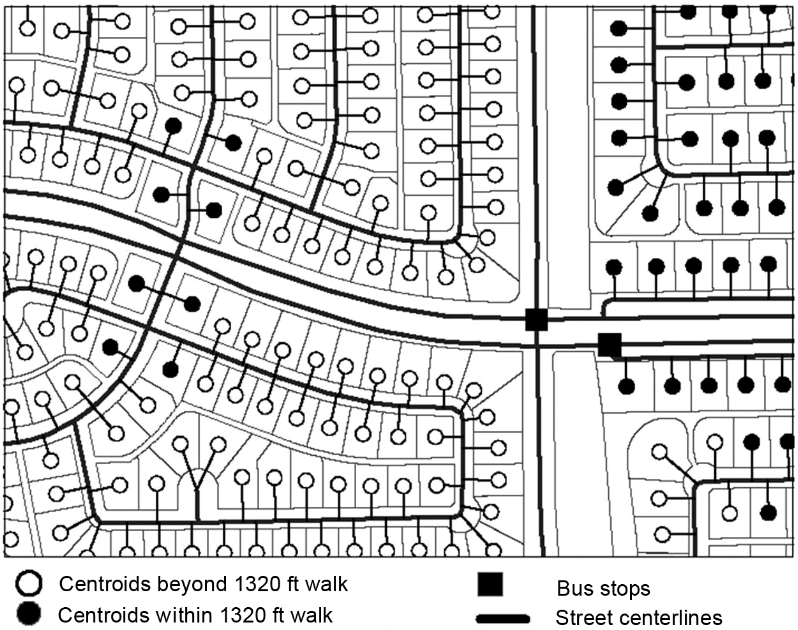

Figure 2. Parcel centroids and the walking network.

With a complete and connected walking network using parcel centroids as origins and bus stop locations as destinations, the next step in the parcel-network method is to generate walking network distances between all of the origins and destinations and then choose the smallest of those distances for each parcel centroid. There are extremely efficient shortest path algorithms, implemented in both industry standard GIS software products and in software designed for travel demand modeling that can be used to populate such origin–destination matrices.

Once the shortest path to a bus stop is determined, the distance traveled and the closest bus stop can be transferred as attributes to the parcel centroids. The shortest paths themselves can also be generated and saved as separate GIS features. At the conclusion of the first three stages of the parcel-network method, the result is a set of parcel centroids with population, the distance to the nearest bus stop, and the identifier for that bus stop.

3.4. Determine population with transit access

With the generation of a complete walking network from each parcel centroid and the association of demographic attributes with those centroids, the fourth and final step in the parcel-network method is to use these data sources to determine the population with access to transit. We can now estimate population in any distance band from a transit facility. In this study, we chose to select those centroids where the walking distance is 400 m (¼ mile) or less to the bus stops of each distinct route. Based on this selection, a summary of the population associated with each route can be made. Additional demographic information for this population can be employed to generate a model of transit use. Several such models have appeared in the literature (Loutzenheiser 1997, Zhao *et al.* 2003) although they depend on initial base populations that are universally believed to be overestimations.

4. Data and study area

The study area for this research project consists of an approximately 160,000 m^2 (100 square mile) section of the Dallas Area Rapid Transit (DART) service area to the north of the city of Dallas in the communities of Richardson and Plano, TX. The inputs to the parcel-network method include (1) a street centerline file maintained by DART which is similar to many publicly available street network databases in its topology and attributes, (2) cadastral (parcel) databases for Dallas and Collin counties, which include spatial descriptions of parcel boundaries, and both spatial and aspatial attributes, most importantly the street address of the parcel, (3) land use data from the North Central Texas Council of Governments which includes data on the types of activity (single-family residential, multi-family residential, commercial, etc.) at parcel locations, (4) a development monitoring point file describing the number of multifamily units at parcel locations, (5) census block area boundary files with associated demographic data, and (6) digital line files describing the paths of existing bus routes along the street network.

Although DART maintains 119 bus routes (at the time of this research) in the Dallas area, this project examines in detail six routes in the study area described above, chosen in order to represent the different types of routes that DART manages. These include a dense multi-family corridor route (463); a transit express route (564); a university, multifamily, and commercial route (562); and several mixed multifamily and single-family neighborhood routes (358, 573, and 361) (Figure 3). All of these routes serve at least one rail station and one transit center with the exception of Routes 573 and 562. Note that Routes 358, 463, and 573 share the same initial portion of their routes through a dense multifamily neighborhood. The percentages of residential and nonresidential land uses within a 400-m (¼-mile) buffer of each route, the percentage of the population living in single-family and multifamily housing within that buffer, and summary statistics for the routes are given in Table 1. In particular, note the high concentration of nonresidential land uses surrounding Route 564 and the high level of multifamily development in the vicinity of Route 463. These values can significantly influence the overestimation given by the existing (buffer and network-ratio) methods.

Several data quality issues exist. First the parcel database is regularly updated while the census data is from 2000. In the intervening years, it is known that more housing has been

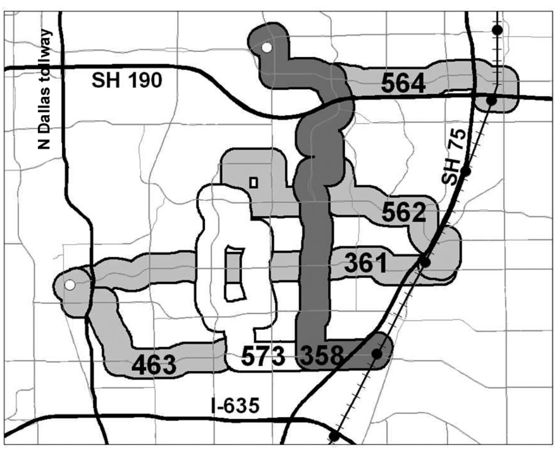

Figure 3. Study area bus routes.

Table 1. Land use and population statistics for six bus routes.

		Route 358	Route 361	Route 463	Route 562	Route 564	Route 573
Land use area %	All residential types	30.3	40.7	47.3	32.7	14.9	57.8
	Commercial, undeveloped, and infrastructure	69.7	59.2	52.8	67.3	85.1	42.2
Population percentages	Single-family	40.2	50.1	14.3	49.5	65.4	47.4
	Multifamily and residential care	59.8	49.9	85.7	50.5	34.6	52.6
General characteristics	Population (buffer method)	17,338	16,062	27,203	13,843	4270	28,006
	Route length (miles; 1609 m per mile)	8.8	6.8	6.6	7.8	5.4	8.8
	¼-mile buffer area (square miles)	4.5	3.6	3.5	3.7	2.8	4.2
	Mean population density (per square mile)	3955	4846	7570	3690	1491	6602

All data computed for the area within a 1320-ft. buffer of the bus routes.

built in the study area. The result of this data set mismatch is that a smaller population from 2000 is allocated to more dwelling units than were present in 2000. Therefore, the authors believe that this results in conservative estimates of the population with access to transit. Additionally, there are cases where a multifamily parcel (such as a large apartment complex) overlays more than one census block. In such cases, the dwelling count for that apartment complex must be appropriately allocated to two or more census blocks. The positional inaccuracy of census block boundaries has been thoroughly discussed elsewhere, and a few alterations in boundary positions were necessary to capture the proper parcel centroids. Lastly, there were a number of cases where census attribute data were known to be incorrect based on field examinations. However, these were left unchanged in order to preserve the ability to replicate the results of this research.

For this study, parcel centroids were reclassified as single-family, multifamily, residential care facilities, or commercial parcels. Commercial centroids were included in creating the walking network, but they were assigned a zero dwelling count and therefore did not participate in the population allocation or contribute to the summary statistics. Residential care facilities, nursing homes and group quarters were given a dwelling count of one if they only had a bed count. If they had a dwelling count they were treated as any other multifamily facility. In this study, the centroids were reclassified and the number of dwelling units transferred using the land use polygon data layer, the development monitoring point file, aerial photography, various reference sources, and occasionally verification with field examinations.

5. Results

The objective of this research is to present a new method that more precisely determines the population within a service area of a transit route and compare its results with the buffer and network-ratio methods.

5.1. Comparison of methods

The results are summarized in Table 2. As expected, both the buffer and network-ratio methods give larger population estimates than the parcel-network method in every case. More specifically, the buffer method always gives the largest estimate for each route. The average increase in estimate when comparing the buffer method to the parcel-network method is nearly 71% although this is clearly affected by a single outlier with an overestimation of 184% (this outlier on Route 564 will be discussed in detail below). The elimination of this outlier still results in an average overestimation of 48%.

The network-ratio method gives a smaller population estimate, but still uniformly results in a higher population value than the parcel-network method. The average increase in estimate of 32% is also affected by the same outlier route that influenced the buffer method, and the elimination of this route results in an average increase in estimate of 14.6%.

The overestimations given by the current methods can be dramatic. To demonstrate this, consider that in the study area being examined here, there are numerous cases where network links are close to transit facilities but have no associated population along those links (Figure 4). In particular, these are highway segments and major arterials that are dominated by commercial or industrial land uses. The network-ratio method assigns population to these links even though no such population exists, resulting in an overestimation of the population with transit access.

The problem is very apparent for Route 564 shown in detail in Figure 4. The focus is on a single analysis zone (a census block group) associated with Route 564. The buffer method (also termed the area-ratio method) compares the area within a buffer [in this case a buffer of 400 m (¼-mile) around Route 564] with the total area of the analysis zone. This ratio is then applied to the population of that analysis zone. In this case, the area within the buffer zone is 16.5% of the total area of the census block group. Therefore, the buffer method suggests that 16.5% of the population of the block group has access to transit (369 people). Similarly, the network-ratio method compares the length of the streets within walking distance of the route to the total length of streets in the analysis zone. For the case being examined here, 16.8% of the streets in the analysis zone are within a 400-m (¼-mile) walking distance of Route 564. This would suggest that 376 people have access to transit.

As shown by the residential parcel centroids (represented with triangles on Figure 4), the parcel-network methods estimates that only a small population (32 people) actually lives within the 400-m (¼-mile) buffer in this census block. The other parcel centroids are commercial or industrial with no associated population. Of the 32 people, none has access to Route 564 within a 400-m (¼-mile) walking distance. Therefore, the estimate from the parcel-network method for this census block group is that none have walking access to this transit facility. This example both highlights the severity of the overestimation that is possible using the buffer and network-ratio methods and highlights the strength of the parcel-network method to accurately spatially identify the locations of potential transit users.

Table 2. Population estimates.

Route numbers	358	361	463	562	564	573
Buffer method	17,338	16,062	27,203	13,843	4270	28,006
Overestimation (%)	51	53	42	54	184	41
Network-ratio method	12,762	12,235	23,277	10,075	3264	22,241
Overestimation (%)	11	16	22	12	117	12
Parcel-network method	11,482	10,507	19,131	8978	1505	19,797

All data computed within a 1320-ft. buffer.

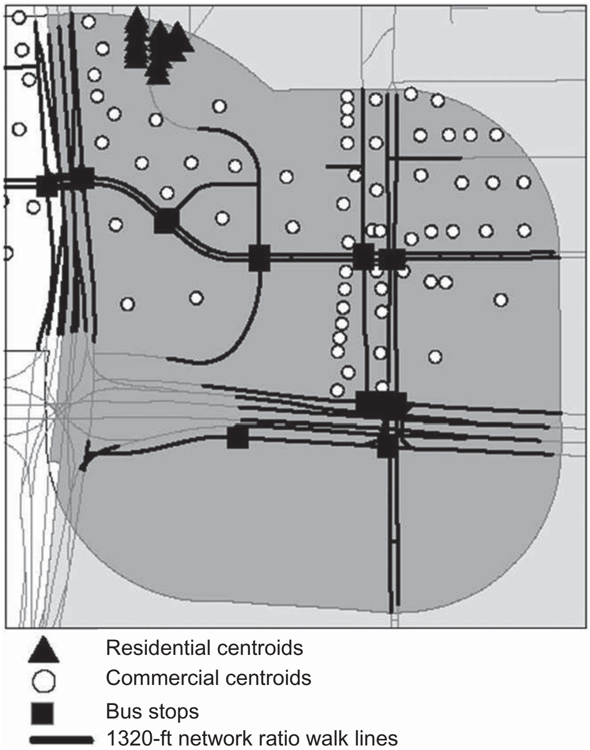

Figure 4. Detail of Route 564.

5.2. New analysis possibilities with the parcel-network method

Beyond the improvement in spatial precision provided, there are a number of analysis options that become available with an application of the parcel-network method. For example, the knowledge of the walking distances from each parcel centroid to its nearest transit facility (e.g. bus stop) allows the researcher to know the average walking distance to transit, a basic summary statistical measure that is unavailable with either of the other methods. This measure is provided in Table 3 for the north Dallas study area routes. This data confirms that the use of the buffer method would capture so many parcels that are outside the 400-m (¼-mile) walking distance, that the average walking distance can be greater than 400 m (¼-mile). In the past, this information could only be determined through expensive and time-consuming survey data collection procedures.

Table 3. Transit accessibility.

Route numbers	358	361	463	562	564	573
Population with accessibility	11,482	10,507	19,131	8978	1505	19,797
Single-family only (%)	36	47	11	37	43	47
Multifamily and residential care only (%)	64	53	89	63	57	53
Mean walking distance all residential types within buffer	1257	1228	1251	1321	1752	1035
Standard deviation	795	577	652	610	567	544

All data computed within a 1320-ft. buffer.

Perhaps more importantly, Table 3 shows a breakdown of the percentage of the population with access to transit by type of residence (multifamily vs. single-family). The data shows that multifamily and residential care parcels tend to have a higher percentage of access [using the 400-m (¼-mile) buffer] than single-family residences. Transit providers make concerted efforts to target these multifamily developments in order to better serve the population. Since these multifamily residences are already well served by transit, this suggests that further increases in transit use could be achieved by targeting routes to the underserved single-family residential areas.

6. Conclusions and future research

It has been shown that the parcel-network method, introduced in this article, can address well-documented issues in determining the population with access to a transit route, specifically the need to implement disaggregated population data (parcel data) and to improve on existing methods for determining the population with access to transit facilities through modeling walking networks over which that population will travel. Although data availability may still be an issue in some areas, this method can be implemented using increasingly available cadastral data in conjunction with GIS and network analytic techniques. The method eliminates the demonstrated overestimation that plagued previous methods due to their allocation schemes. By associating aggregate population data with parcel centroid points (rather than with areas or lines) and by creating a walking network that extends from the centroid of each parcel, the parcel-network method provides a degree of spatial precision that had not previously been possible without the use of data collected through survey research. Since decisions regarding the provision of transit service and the locations of routes depend on estimates of the number of potential riders, the parcel-network method offers a more robust tool for transit ridership forecasting. In the future, this method could be expanded to explore not only the physical access to transit but many other demographic characteristics (car ownership, income, age distribution, etc.) that could influence ridership patterns and therefore route optimality. Once the parcel-network method is implemented for a particular study area, the relative permanence of the parcel and street network data may allow for efficient update when new demographic data become available.

In addition to these improvements, the parcel-network method allows for analysis options that had not previously been available. Not only do we know the distance to the nearest transit facility but we can also generate the exact path that each (cost minimizing) transit user would follow. This path information could be used in several ways to improve transit accessibility. First, it could help in the study of neighborhood design in order to make improvements in existing pedestrian facilities through sidewalk improvements, elimination of barriers, or adding amenities. Second, the path information could be used to model changes to transit routes that would reduce total (or average) walking distances to the route for a chosen population.

Immediate implications of the parcel-network method for future research include the need to test new, hypothetical routes in order to determine if changes in route geometry would provide access to transit for larger populations. Since there are an astronomically large number of possible route geometries, the findings given above, regarding the residential types that are currently well (or poorly) served by transit, lead to suggestions for narrowing the set of new route geometries to those that are likely to improve on transit accessibility. Measuring access to different types of destinations (e.g. work, retail, and commercial) could help identify areas of greater transit need (Lee 2005). With a move beyond simple buffers, the types of spatial inter-relationships among routes [such as overlapping service areas (Peng and Dueker 1995, Peng *et al.* 1997, Kimpel *et al.* 2007)] need to be re-examined, and the effects those interrelationships have on competition between – or cooperation among – routes should be measured. Lastly, if the parcel-network method were repeated for several metropolitan areas or parts of metro areas, the relationship between the values generated by the buffer method and the parcel-network method could be formalized as a heuristic, which could then be used by planners who lacked the time or resources to complete the data preparation necessary for the parcel-network method. That is, we could determine an expected overestimation of the buffer method when compared to the parcel-network method for different types of transit routes.

The parcel-network method can also provide a more accurate starting point for application in other transit use models. In this study, the outer limit of accessible walking distance was chosen to be 400 m (¼-mile). Other models used to predict transit use may consider the quality of transit service and walking distances as far as 1.6 km (1 mile). A data set prepared using the parcel-network method could be easily applied in transit use models that employ a distance decay formula in the estimation of transit use rather than transit accessibility. Similarly, models including more complex transit access modes (e.g. park-and-ride or kiss-and-ride) could be explored using driving paths rather than walking paths. A data set prepared using the parcel-network method also has the identifier of the nearest bus stop for each centroid, data that are unavailable in other methods. By summarizing attributes based upon the bus stop identifier, the population served by each bus stop can easily be determined. Such information could be useful in determining bus stop placement and amenities. Additional transportation applications that could benefit from this method include those concerned with studies of routine activities (Stopher *et al.* 1996), pedestrian spatial behavior (Shoval and Isaacson 2006), microsimulation-based evacuation scenarios (Cova and Johnson 2002, 2003, Chen 2008), demand-responsive transit systems (Cortes and Jayakrishnan 2002) or para-transit services (Fu 2002, Aldaihani and Dessouky 2003), congestion re-routing (Kwan and Casas 2006), or emergency response (Huang and Pan 2007).

Perhaps most importantly, this article highlights the potential for the combination of highly detailed spatial information (in the form of cadastral databases) with network analysis techniques to advance research across GIScience. For instance, in spatial-criminological research, the journeys to and from crime (Lu 2003) and neighborhood risk for crime (Malczewski and Poetz 2005) could be informed by the parcel-network method, as could investigations of work and residential segregation in urban areas if employment and racial information are associated with the parcel centroids (Ellis *et al.* 2004). The role of transport networks in changing urban form (Torrens 2006, Stevens *et al.* 2007) could be more clearly understood with the more detailed associations delimited by the parcel network method. Much of the rapidly growing subdiscipline of health geographics is concerned with access to services by different populations in need and calls in the literature for more disaggregate data have appeared (Langford and Higgs 2006). Research where GIS is used to explore laborsheds (Sultana and Weber 2007), spatial access for the disabled (Church and Marston 2003, Casas 2007), voting geography (Gimpel *et al.* 2008), real estate geography (Zeng and Zhou 2001), and retail geography (Larsen and Gilliland 2008) could all benefit from the advances and flexibility that the parcel-network method affords. We intend to encourage this advancement across GIScience and across Geography.

Acknowledgments

The authors thank several anonymous referees who provided constructive critiques on this article. Any remaining errors or omissions are solely the responsibility of the authors.

References

Aldaihani, M. and Dessouky, M.M., 2003. Hybrid scheduling methods for paratransit operations. *Computers and Industrial Engineering*, 45, 75–96.

Ayvalik, C.K. and Khisty, C.J., 2002. Heuristic analysis of impacts of commuter rail station consolidation on pedestrian access. *Transportation Research Record*, 1793, 47–54.

Beimborn, E.A., Greenwald, M.J., and Jin, X., 2003. Accessibility, connectivity, and captivity: impacts on transit choice. *Transportation Research Record*, 1835, 1–9.

Burkhardt, J.E., 2003. Critical measures of transit service quality in the eyes of older travelers. *Transportation Research Record,* 1835, 84–92.

Calvo, R.W., *et al*., 2004. A distributed geographic information system for the daily car pooling problem. *Computers and Operations Research*, 31, 2263–2278.

Casas, I., 2007. Social exclusion and the disabled: an accessibility approach. *The Professional Geographer*, 59, 463–477.

Chang, N.B., Lu, H.Y., and Wei, Y.L., 1997. GIS technology for vehicle routing and scheduling in solid waste collection systems. *Journal of Environmental Engineering-Asce*, 123, 901–910.

Chapin, T.S., Deyle, R.E., and Baker, E.J., 2008. A parcel-based GIS method for evaluating conformance of local land-use planning with a state mandate to reduce exposure to hurricane flooding. *Environment and Planning B-Planning & Design*, 35, 261–279.

Chen, X., 2008. Microsimulation of hurricane evacuation strategies of Galveston Island. *The Professional Geographer*, 60, 160–173.

Chrisman, N.R., 1987. The accuracy of map overlays: a reassessment. *Landscape and Urban Planning*, 14, 427–439.

Christen, P., Willmore, A., and Churches, T., 2006. A Probabilistic geocoding system utilising a parcel based address file. *In*: *Data mining: theory, methodology, techniques, and applications, ed.* G.J. Williams and S.J. Simoff Berlin: Springer, 130–145.

Church, R.L. and Marston, J.R., 2003. Measuring accessibility for people with a disability. *Geographical Analysis*, 35, 85–96.

Cortes, C.E. and Jayakrishnan, R., 2002. Design and operational concepts of high-coverage point-to-point transit system. *Transportation Research Record,* 1783, 178–187.

Cova, T.J. and Johnson, J.P., 2002. Microsimulation of neighborhood evacuations in the urban – wildland interface. *Environment and Planning A*, 34, 2211–2229.

Cova, T.J. and Johnson, J.P., 2003. A network flow model for lane-based evacuation routing. *Transportation Research Part A-Policy and Practice*, 37, 579–604.

Curtin, K.M., 2007. Network analysis in geographic information science: review, assessment, and projections. *Cartography and Geographic Information Systems*, 34, 103–111.

Curtin, K.M., 2008. Network data structures. *In*: K.K. Kemp, ed. *The encyclopedia of geographic information science*. Los Angeles: Sage Publications, 314–317.

Dallas Area Rapid Transit, 1998. *DART service standards: policies and guidelines*. Dallas, TX, p. 22.

Dallas Area Rapid Transit, 2002. *Service planning five year action plan*. Dallas, TX, p. 69.

Dill, J., 2003. Transit use and proximity to rail: results from large employment sites in the San Francisco, California, bay area. *Transportation Research Record*, 1835, 19–24.

Ellis, M., Wright, R., and Parks, V., 2004. Work together, live apart? Geogrpahies of racial and ethnic segregation at home and at work. *Annals of the Association of American Geographers*, 94, 620–637.

Evans, J.E., Perincherry, V., and Douglas, G.B.I., 1997. Transit friendliness factor: approach to quantifying transit access environment in a transportation planning model. *Transportation Research Record*, 1604, 32–39.

Federal Highway Administration, 2002. *Toolbox for regional policy analysis*. Washington, DC: US Department of Transportation.

Fu, L.P., 2002. Planning and design of flex-route transit services. *Transportation Research Record*, 1791, 59–66.

Furth, P.G., Mekuria, M.C., and SanClemente, J.L., 2007. Stop spacing analysis using geographic information system tools with parcel and street network data. *Transportation Research Record*, 2034, 73–81.

Gimpel, J.G., *et al*., 2008. Distance-decay in the political geography of friends-and-neighbors voting. *Political Geography*, 27, 231–252.

Goodchild, M.F., Yuan, M., and Cova, T.J., 2007. Towards a general theory of geographic representation in GIS. *International Journal of Geographical Information Science*, 21, 239–260.

Handy, S.L. and Niemeier, D.A., 1997. Measuring accessibility: an exploration of issues and alternatives. *Environment and Planning A*, 29, 1175–1194.
Hanson, S. and Schwab, M., 1995. Describing disaggregate flows: individual and household activity patterns. *In*: S. Hanson, ed. *The Geography of Urban Transportation*. New York: The Guilford Press, 166–187.
Horner, M.W., 2004. Spatial dimensions of urban commuting: a review of major issues and their implications for future research. *The Professional Geographer*, 56, 160–173.
Horner, M.W. and Murray, A.T., 2004. Spatial representation and scale impacts in transit service assessment. *Environment and Planning B-Planning & Design*, 31, 785–797.
Hsiao, S., *et al*., 1997. Use of geographic information system for analysis of transit pedestrian access. *Transportation Research Record*, 1604, 50–59.
Huang, B. and Pan, X.H., 2007. GIS coupled with traffic simulation and optimization for incident response. *Computers Environment and Urban Systems*, 31, 116–132.
Ioannou, G., Kritikos, M.N., and Prastacos, G.P., 2002. Map-route: a GIS-based decision support system for intra-city vehicle routing with time windows. *Journal of the Operational Research Society*, 53, 842–854.
Karimi, H.A., Peachachavanich, R., and Peng, J., 2004. Finding optimal bus service routes: internet-based methodology to serve transit patrons. *Journal of Computing in Civil Engineering*, 18, 83–91.
Kimpel, T.J., Dueker, K.J., and El-Geneidy, A.M., 2007. Using GIS to measure the effect of overlapping service areas on passenger boardings at bus stops. *URISA Journal*, 19, 5–11.
Kuby, M., Barranda, A., and Upchurch, C., 2004. Factors influencing light-rail station boardings in the United States. *Transportation Research Part A – Policy and Practice*, 38, 223–247.
Kwan, M.P. and Casas, I., 2006. GABRIEL: GIS activity-based travel simulator. Activity scheduling in the presence of real-time information. *Geoinformatica*, 10, 469–493.
Kwan, M.P., *et al*., 2003. Recent advances in accessibility research: representation, methodology and applications. *Journal of Geographical Systems*, 5, 129–138.
Langford, M. and Higgs, G., 2006. Measuring potential access to primary healthcare services: the influence of alternative spatial representations of population. *The Professional Geographer*, 58, 294–306.
Larsen, K. and Gilliland, J., 2008. Mapping the evolution of 'food deserts' in a Canadian city: supermarket accessibility in London, Ontario, 1961–2005. *International Journal of Health Geographics*, 7, 16.
Lee, B.H., 2005. Parcel-level measure of public transit accessibility to destinations. *In: Access to Destinations*, Ed. D.M. Levinson and K.J. Krizek. Amsterdam: Elsevier.
Lee, C., Moudon, A.V., and Courbois, J.Y.P., 2006. Built environment and behavior: spatial sampling using parcel data. *Annals of Epidemiology*, 16, 387–394.
Lepofsky, M., Abkowitz, M., and Cheng, P., 1993. Transportation hazard analysis in integrated GIS environment. *Journal of Transportation Engineering-Asce*, 119, 239–254.
Loutzenheiser, D.R., 1997. Pedestrian access to transit: model of walk trips and their design and urban form determinants around bar area rapid transit stations. *Transportation Research Record*, 1604, 40–49.
Lu, Y., 2003. Getting away with the stolen vehicle: an investigation of journey-after-crime. *Professional Geographer*, 35, 415–433.
MacDougall, E.B., 1975. The accuracy of map overlays. *Landscape and Planning*, 2, 23–30.
Malczewski, J. and Poetz, A., 2005. Residential burglaries and neighborhood socioeconomic context in London, Ontario: global and local regression analysis. *The Professional Geographer*, 57, 516–529.
Miller, H.J. and Shaw, S., 2001. *Geographic information systems for transportation*. New York: Oxford University Press.
Miller, H.J. and Storm, J.D., 1996. Geographic information system design for network equilibrium-based travel demand models. *Transportation Research Part C – Emerging Technologies*, 4, 373–389.
Miranda, M.L., Dolinoy, D.C., and Overstreet, M.A., 2002. Mapping for prevention: GIS models for directing childhood lead poisoning prevention programs. *Environmental Health Perspectives*, 110, 947–953.
Murray, A.T., 2001. Strategic analysis of public transport coverage. *Socio-Economic Planning Sciences*, 35, 175–188.
Murray, A.T., *et al*., 1998. Public transportation access. *Transporation Research – Part D*, 3, 319–328.

Newcomer, J.A. and Szajgin, J., 1984. Accumulation of thematic map errors in digital overlay analysis. *The American Cartographer*, 11, 58–62.

Nyerges, T.L., 1995. Geographical information system support for urban/regional transportation analysis. *In*: S. Hanson, ed. *The geography of urban transportation*. New York: The Guilford Press, 240–265.

O'Neill, W.A., Ramsey, R.D., and Chou, J., 1992. Analysis of transit service areas using geographic information systems. *Transportation Research Record*, 1364, 131–138.

O'Sullivan, S. and Morrall, J., 1996. Walking distances to and from light-rail transit stations. *Transportation Research Record*, 1538, 19–26.

O'Sullivan, D., Morrison, A., and Shearer, J., 2000. Using desktop GIS for the investigation of accessibility by public transport: an isochrone approach. *International Journal of Geographical Information Science*, 14, 85–104.

Osegueda, R., *et al*., 1999. GIS-based network routing procedures for overweight and oversized vehicles. *Journal of Transportation Engineering-Asce*, 125, 324–331.

Peng, Z. and Dueker, K.J., 1995. Spatial data integration in route-level transit demand modeling. *Journal of the Urban and Regional Information Systems Association*, 7, 26–37.

Peng, Z., *et al*., 1997. A simultaneous route-level transit patronage model: demand, supply, and inter-route relationship. *Transportation*, 24, 159–181.

Phillips, C.G. and Edwards, H.R., 2002. Socioeconomic, community-based approach for developing integrated mass transit systems application to city of Baltimore, Maryland. *Transportation Research Record*, 1797, 71–79.

Reeve, R.E., 1984. Computerizing land title, tax and valuation records in orange county, NC. *Computers, Environment, and Urban Systems*, 9, 133–141.

Shoval, N. and Isaacson, M., 2006. Application of tracking technologies to the study of pedestrian spatial behavior. *The Professional Geographer*, 58, 172–183.

Stevens, D., Dragicevic, S., and Rothley, K., 2007. iCity: a GIS-CA modelling tool for urban planning and decision making. *Environmental Modelling and Software*, 22, 761–773.

Stopher, P.R., Hartgen, D.T., and Li, Y.J., 1996. SMART: simulation model for activities, resources and travel. *Transportation*, 23, 293–312.

Strager, M.P. and Rosenberger, R.S., 2007. Aggregating high-priority landscape areas to the parcel level: an easement implementation tool. *Journal of Environmental Management*, 82, 290–298.

Sultana, S. and Weber, J., 2007. Journey-to-work patterns in the age of sprawl: evidence from two midsize southern metropolitan areas. *The Professional Geographer*, 59, 193–207.

Tarantilis, C.D., Diakoulaki, D., and Kiranoudis, C.T., 2004. Combination of geographical information system and efficient routing algorithms for real life distribution operations. *European Journal of Operational Research*, 152, 437–453.

Torrens, P.M., 2006. Simulating sprawl. *Annals of the Association of American Geographers*, 96, 248–275.

Upchurch, C., *et al*., 2004. Using GIS to generate mutually exclusive service areas linking travel on and off a network. *Journal of Transport Geography*, 12, 23–33.

US Census Bureau, 2005. Geographic areas reference manual. *In*: *Chapter 11 – census blocks and block groups*.

Veregin, H., 1989. Developing and testing of an error propagation model for GIS overlay operations. *International Journal of Geographical Information Systems*, 9, 595–619.

Walker, W.T., Gao, S., and Johnston, R.A., 2007. UPlan – geographic information system as framework for integrated land use planning model. *Transportation Research Record*, 1994 117–127.

Wallace, G.N., *et al*., 2008. Assessing the ecological and social benefits of private land conservation in Colorado. *Conservation Biology*, 22, 284–296.

Weigel, D. and Cao, B.Y., 1999. Applying GIS and OR techniques to solve sears technician-dispatching and home-delivery problems. *Interfaces*, 29, 112–130.

Werle, J.W., 1984. Problems in automating traditional land records data. *Computers, Environment, and Urban Systems*, 9, 199–202.

Wilcox, D.J., 1984. Proposed map accuracy standards for a multipurpose cadastre. *Computers, Environment, and Urban Systems*, 9, 203–207.

Wu, B.M. and Hine, J.P., 2003. A PTAL approach to measuring changes in bus service accessibility. *Transport Policy*, 10, 307–320.

Xie, Z.X., 2006. A framework for interpolating the population surface at the residential-housing-unit level. *GIScience & Remote Sensing*, 43, 233–251.

Zandbergen, P.A., 2008. A comparison of address point, parcel and street geocoding techniques. *Computers Environment and Urban Systems*, 32, 214–232.
Zeng, T.Q. and Zhou, Q.M., 2001. Optimal spatial decision making using GIS: a prototype of a real estate geographical information system (REGIS). *International Journal of Geographical Information Science*, 15, 307–321.
Zhao, F., *et al*., 2003. Forecasting transit walk accessibility – regression model alternative to buffer method. *Transportation Research Record*, 1835, 34–41.