Journal of Transport Geography 115 (2024) 103808

Contents lists available at ScienceDirect

Journal of Transport Geography

journal homepage: www.elsevier.com/locate/jtrangeo

Measuring geographic accessibility in data poor rural areas by augmenting the road network with a triangular irregular network – A case study in the O.R. Tambo District Municipality of the Eastern Cape, South Africa

Lourens Snyman, Serena Coetzee *

Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa

ARTICLE INFO

Keywords:
Geographic accessibility
Rural
Triangular irregular network
South Africa

ABSTRACT

Travelling long distances to access public service facilities is costly and time consuming, especially to those who suffer the burden of poverty and deprivation. Assessing accessibility based on route-based distances does not work well in rural areas of Africa where people often travel along footpaths and dirt roads (in and out of vehicles) that are not mapped, not necessarily well-defined, could vary based on the season or change over short periods of time. The South African government therefore recommended that geographic accessibility be assessed based on a road network that is augmented by connecting communities to the main road network via a triangulated irregular network (TIN). While this method is recommended, a study about its suitability has not been published. We do this by describing an implementation, analysing the results and comparing these to other techniques. In this paper we present and compare the recommended method to geographic accessibility results of two other frequently used distance measuring techniques: 1) straight-line distances; 2) route-based distances in the road network. We applied the techniques to South African Police Service stations in the O.R. Tambo District Municipality, Eastern Cape, South Africa, which is characterized by rural villages scattered across the landscape, not connected to the road network and where data for footpaths and dirt roads does not exist. The results confirm that in the absence of a completely mapped travel network, the recommended method is a suitable alternative for measuring geographic accessibility in rural areas of Africa without having to spend time and money on mapping footpaths and dirt roads that keep on changing anyhow. However, the results reveal sensitivity to the threshold distance and significant local variation depending on distance from urban areas and proximity to natural barriers such as rivers. The latter must be considered when deciding whether to apply the recommended distance measurement technique in rural accessibility studies. Further research could compare the techniques for other study areas and also refine the method, e.g., by experimenting with different algorithms for centroids of populated areas and finding ways to estimate travel speeds on TIN arcs.

1. Introduction

Despite rapid urbanization in Africa, large parts of the continent are still characterized by sparsely populated non-urban areas with limited or poor road infrastructure. Rural access is essential for reducing poverty and supporting inclusive economic growth. Yet, in many African countries, farmers are disconnected from markets due to poor transport connectivity (Transport and ICT Global Practice, 2016) and more generally, rural populations are often disconnected from public services. Travelling long distances to reach these facilities is costly and time consuming, especially to those who suffer the burden of poverty and deprivation. Knowing where the population demand is in relation to public service facilities provides decision makers with valuable information for connecting the rural population to public services.

Accessibility measures how reachable a set of locations is by a particular group of people, and is the most widely used metric for access to public services (Church and Murray, 2009). It can be measured in two ways: geographically, by considering distances to be travelled by people to reach the closest location, or socially, by considering people's behaviour (Kwan, 2009; Miller, 2007; Neutens et al., 2010). In a rural context, where service providers are far apart, physical distances are adequate for assessing geographic accessibility (Neutens, 2015).

---

* Corresponding author at: Department of Geography, Geoinformatics and Meteorology, Geography Building, Hatfield campus, Lynnwood Road, University of Pretoria, 0083, South Africa.

E-mail address: serena.coetzee@up.ac.za (S. Coetzee).

https://doi.org/10.1016/j.jtrangeo.2024.103808
Received 28 November 2022; Received in revised form 30 October 2023; Accepted 16 January 2024

However, the travel network data for rural or non-urban areas is often not available and it takes much time and effort to collect and prepare for accessibility studies (Blanford et al., 2012; Ilie et al., 2019). Inaccurate accessibility measurements based on limited data may lead to poor return on investment when public service facilities are added or removed with the aim of improving accessibility and could also lead to public dissatisfaction with service delivery (Masiya et al., 2019).

Large parts of rural South Africa also have transport connectivity challenges. Despite post-apartheid reform initiatives aimed at improving equitable access to public services in South Africa, the struggle to achieve this is not yet over (Van Rensburg, 2014; Masiya et al., 2019). To this day, there are structural differences between the former white 'platteland', rural areas where mostly commercial farming takes place, and former homelands, rural areas with mostly subsistence farming on tribal land. The former homelands are characterized by backlogs in infrastructure such as streets, water, electricity and sanitation, and they are also still poorly mapped. The Rural Transport Strategy noted that "rural people have vastly inferior access to basic social services and the economic mainstream" (Department of Transport, 2007). Many people travel by foot between rural settlements or to public transport pickup points. In some cases, they have to walk several kilometers before reaching the road network and travel a further distance in vehicles on unmapped dirt roads before reaching the government maintained roads. High occurrence of rural poverty, remoteness and low population densities make it difficult to develop and integrate these areas (Department of Transport, 2007).

The ultimate goal would be to improve rural accessibility through a road network, as specified in indicator 9.1.1 of the United Nations Sustainable Development Goal (SDG) global indicator framework, which measures the "proportion of the rural population who live within 2 km of an all-season road" (United Nations Statistical Commission, 2016). For the time being, we are left with no other choice than to consider walking routes and dirt roads when assessing accessibility. In this context, accurate measurement of accessibility is difficult because footpaths or walking trails and dirt roads are usually not included in datasets of government maintained road networks, and often they are not even mapped at all. Moreover, footpaths and dirt roads may vary depending on the season and identifying them on aerial imagery (e.g. for remote digitizing via OpenStreetMap) can be difficult.

Consequently, the Department of Public Service Administration (DPSA) guidelines for measuring geographic accessibility aimed at improving geographic access to government service points in South Africa recommend that the road network data be improved by using triangulated irregular networks (TIN) to connect communities and service points to the main road network (DPSA, 2012). In data poor areas, the TIN serves as proxy for footpaths and dirt roads, as an alternative to the time intensive work of preparing a travel network that may anyhow vary depending on the season. While government recommends this approach, a study of its suitability in rural areas of South Africa, description of its implementation, and analysis and comparison to other methods has not been published to date. We therefore assessed the suitability of the recommended approach by implementing it for a rural part of South Africa and compared the accessibility results to that of two other frequently used methods, with reference to threshold distances for police stations specified in the South African guidelines for provision of social facilities (Green and Argue, 2012).

In this paper we describe how we implemented the DPSA recommended approach by combining the government maintained road network for in-vehicle travel with a TIN that represents unmapped footpaths and dirt roads for both in-vehicle and out-of-vehicle travel. We compared the geographic accessibility results of three distance measuring techniques: 1) straight-line distances; 2) route-based distances in the road network; and 3) route-based distances in an augmented travel network. Since the South African guidelines specify threshold distances for police stations (Green and Argue, 2012), we only considered distance, not travel speeds on different kinds of roads. The next section provides background about accessibility, how it can be measured and the particular knowledge gap that is addressed in this paper. In Section 3, we describe the study area, explain which data and software were used, and how we compared the accessibility results. In Section 4, we explain how the road network was augmented and compare the accessibility results for the three distance measurement techniques. In Section 5, we discuss results with specific reference to the suitability of this method in rural areas. Section 6 concludes.

2. Literature review and knowledge gap

Accessibility measures how reachable a set of locations is by a particular group of people, and is the most widely used metric for access to public services (Church and Murray, 2009). Accessibility is multifaceted and can be measured in different ways. For example, health care service accessibility has been measured by assessing the costs of health care utilization (affordability), health service compliance and satisfaction (acceptability), adequacy of health service provision (availability), travel impedance between patients and providers (geographic accessibility) and appropriateness and suitability of health services (accommodation) (Neutens, 2015). Researchers also distinguish between place-based and people-based measures of accessibility (Miller, 2007; Neutens et al., 2010; Kwan, 2009). Place-based measures consider the closeness between a key location in an individual's life (e.g., home or workplace) and desired locations (e.g., a service provider location), while people-based measures consider the activities of individuals in space and time, and how they use places in the real and in the virtual world (Miller, 2007).

In this paper, the focus is on geographic accessibility, a form of place-based accessibility, also sometimes referred to as spatial or physical accessibility. Geographic accessibility is "the degree to which transport systems enable people to reach desired activity locations" (Neutens, 2015) or "the relative ease by which the locations of activities, such as work, shopping, and health care, can be reached from a given location" (Luo and Wang, 2003). The DPSA guidelines define accessibility as "the supply and demand of facilities within a defined area by making use of a movement network" and recognizes the "combined effects of various factors in determining the optimal size, number and location of facilities to meet the demand of citizens within reach of them".

Geographic accessibility can be measured in different ways. The simpler metrics are generally favoured because they are easy to implement and directly interpretable in absolute units, which make them easy to understand. Examples are the ratio between potential users and the number of service locations or the distance or travel time between potential users of the service and locations where the service is provided. Travel time depends on the mode of transport, e.g. walking vs driving, and therefore requires more information than 'just' the line segments of a road network. In more complex gravity-based metrics there is a trade-off between the size and/or quality of health care facilities and travel impedance (Neutens, 2015). Researchers have confirmed that in developing countries, physical separation between users and providers is the most important resource constraint in achieving equitable access to services (Hundt et al., 2012; Kahabuka et al., 2011; Müller et al., 1998; Ndirangu et al., 2009; Noor et al., 2003; Poku-Boansi et al., 2010; Stock, 1983; Tanser et al., 2001). In a rural context where service providers are far apart, physical distances are adequate for assessing geographic accessibility (Neutens, 2015).

In the public sector 'acceptable proximity' usually specified, meaning that if a service is located within a specified threshold (maximum) distance, the service location is 'acceptable' and it is accepted that the user (or population) is covered (Drezner and Hamacher, 2002). Along this line, the guidelines for the provision of social facilities in South African settlements acknowledge that geographic accessibility can be specified as distance or travel time, but except for fire stations, only distances are specified. Threshold distances are also specified because the population density varies significantly across the country and it is therefore difficult to specify a population threshold (Green and Argue, 2012). This study therefore considers geographic accessibility based on distances.

Distance-based accessibility can be measured either from the population perspective or from the service provider perspective. Accessibility by the population is indicated by the shortest distance between the population and a service provider location; a shorter distance implies better accessibility by the population. Accessibility of the service provider is indicated by the size of the population within a specified threshold distance from a service provider location; a larger population implies better accessibility of the service provider location. Accordingly, the widely used two-step floating catchment area (2SFCA) method (Luo and Wang, 2003) measures geographic accessibility in terms of both the size of the population catchment (service provider locations that fall within a threshold distance) and the size of the service catchment (population within a threshold distance) (Mcgrail and Humphreys, 2009).

There are different ways of measuring distance, e.g., by calculating the distances between locations based on straight-line distances on a sphere or spheroid, by calculating shortest routes between locations via a road network, or by calculating distances on a contiguous surface, e.g., divided into grid cells, so that accessibility is a function of the spatial structure of the surface (Rodrigue et al., 2009). Accessibility based on straight-line distances is simpler to calculate, e.g., as Euclidean distance or Manhattan distance, but ignores the travel network and natural barriers (Boscoe et al., 2012). This may cause deviation from actual travel distances (Salonen et al., 2012; Polo et al., 2013) and may overestimate the population within a certain threshold distance from a location (Mavoa et al., 2012). Furthermore, accessibility based on straight-line distances ignores the mode of transport that is used. However, this technique is used most frequently because it is so simple to calculate (Boscoe et al., 2012). Calculating distances in a network provides a more accurate reflection of travel distances, but requires network data and processing of routing algorithms. Contiguous accessibility is useful in areas where road network data is not available and/or where people travel on foot (see e.g. Blanford et al., 2012; Pozzi and Robinson, 2008; Tanser, 2006; Tanser et al., 2006).

Several researchers have compared geographical accessibility based on distance types and found that accessibility varies significantly depending on the type of distance used. See for example, Apparicio et al. (2008, 2017) for health services; La Rosa (2013) and Higgs et al. (2012) for greenspaces and Polo et al. (2013) for rabies vaccination sites. Also, there are often localized variations due to the specificities of an area (Boscoe et al., 2012). This points to the importance of carefully choosing not only a distance type appropriate for the task at hand, but also carefully considering the impact of a specific threshold distance in a study area.

The lack of geospatial data is reported as a challenge in many accessibility studies (e.g., Al-Taiar et al., 2010; Ndirangu et al., 2009; Müller et al., 1998). As a result, some researchers spend much time and money on collecting and improving data (Blanford et al., 2012; Noor et al., 2006; Tanser et al., 2006) or simply revert to using straight-line distances (Al-Taiar et al., 2010; Feikin et al., 2009; Noor et al., 2003; Schoeps et al., 2011; Tanser et al., 2001). Noor et al. (2006) laments that sophisticated access models cannot be scaled up to the national level with precision if the paucity of geospatial data is not addressed. The lack of data calls for “creative GIS methodologies” (Neutens, 2015), of which the method presented in this paper is a good example.

The DPSA’s recommendation to augment travel networks with a TIN is an example of a scalable solution in the case of data poor rural areas. While this is recommended, to date, suitability through an assessment of such an implementation and analysis of the accessibility results has not been published. Accessibility studies for South Africa reported in literature, e.g., Baloyi et al. (2017) for public ambulances in the City of Tshwane, Mokgalaka et al. (2013) for primary health care demand in the City of Cape Town and Green et al. (2014) for fire stations across South Africa, have not applied the government recommended approach and/or were conducted for metropolitan areas.

3. Method

3.1. Study area

The O.R. Tambo District Municipality is in the Eastern Cape Province, located in the south-east of South Africa, see Fig. 1. We chose this district municipality for our study because of its large population living in a rural setting far away from roads that are maintained by the government; government datasets usually include only those roads maintained by them.

Among district and metropolitan municipalities in the Eastern Cape, the O.R. Tambo District Municipality has the largest population. Of the 1.3 million people, only 13.9% reside in urban areas, based on the definition for an urban area provided by Statistics South Africa for the 2011 Census (Statistics South Africa, 2012). Many areas of South Africa, including significant parts of the Eastern Cape, were never proclaimed as urban areas, and can only be classified as such by aerial photographs. Because they were not proclaimed, there was little or no service delivery (e.g., roads, electricity, waste removal, addressing) in these areas (Statistics South Africa, 2019). Many of these areas, even today, are governed by traditional authorities, and not municipalities, therefore service delivery (incl. Transportation systems), spatial planning and land use management are often not implemented optimally (du Plessis, 2018). Based on 2016 data, the Eastern Cape is the province in which the highest proportion of people live in poverty in all its dimensions (12.7%) (Statistics South Africa, 2021). Understanding and improving accessibility is one way of alleviating poverty in rural villages (Lucas, 2011). The CSIR guidelines for the provision of social facilities in settlements specify recommended maximum travel distances to social facilities, aimed at improving access to social facilities in South Africa. For police stations, the following is specified: “8km urban/metro; 15km peri-urban; 24km rural and settlement type E (small towns/ isolated regional service centres)” (Green and Argue, 2012). People visit police stations for services, ranging from reporting a crime or missing person, to the certification of documents and police clearance certificates (SAPS, 2023).

The map in Fig. 2 shows the urban areas, 19 police stations (three of them in Mthatha) and road network in the study area, while the inset map confirms that many people live in areas with limited or no access to the road network maintained by government. It shows the total road length (km) per person in a small area (blank or white areas on the map are unpopulated). The satellite image in Fig. 3 reveals the unmapped dirt roads and footpaths in a non-urban area of O. R. Tambo District Municipality.

3.2. Accessibility comparison

As a first step, we prepared the augmented travel network by connecting the centroids of small areas (smallest available Census area, see 3.3) to the road network and then augmenting this network with a TIN, as described in 4.1.

As a second step, we calculated three cost matrices for the distances between police stations and centroids of non-urban small areas, one each for the three distance measurement techniques:

A. Straight-line distance
B. Shortest route in the road network
C. Shortest route in the augmented travel network

A cost matrix specifies the costs between two sets of elements – in this case, a set of origins and a set of destinations. Cost refers to the distance between origin (small area centroid) and destination (police station). With the exception of fire stations, South African guidelines specify

Fig. 1. O. R. Tambo District Municipality in the Eastern Cape, South Africa (Data sources: Municipal Demarcation Board, OpenStreetMap, Statistics South Africa and NAVTEQ).

accessibility thresholds in terms of distance (Green and Argue, 2012) and we therefore considered only distance, not travel speeds on different kinds of roads. From these matrices, the nearest police station (destination) for each small area centroid (origin) was identified. Based on this, we computed the geographic accessibility by the population of each small area centroid (distance to the closest police station) and geographic accessibility of the police station (the sum of populations of small areas for which a particular station was identified as the closest station). Accessibility results for the three distance measurement techniques are presented in 4.2.

We compared the accessibility by population (4.2.1) and the accessibility of the police station (4.2.2). Next, we compared the geographic accessibility results for the three distance measurement techniques statistically, based on the variance between the three measurement techniques (4.2.3), total squared error (TSE) and mean squared error (MSE) between the augmented travel network and the other two techniques respectively (4.2.4), and the error rate per locality (4.2.5). For the latter, we also used a local indicator of spatial association (LISA) statistic to identify and visualize spatial patterns of high and low error rates between the different techniques (Anselin, 1995). Lastly, we plotted a Lorenz curve and calculated the Gini coefficient to measure equality of accessibility for each technique respectively (4.2.6). Although the Lorenz curve was originally created to measure the level of inequality based on the concentration (or distribution) of wealth (Lorenz, 1905), it is also widely used to measure inequality with regards to geographic accessibility (Cromley, 2019),(Azmoodeh et al., 2021), (Rofé et al., 2015; Guzman et al., 2017).

3.3. Data and software

For our study, we used a road dataset for the Eastern Cape Province provided by AfriGIS (https://www.afrigis.co.za/). Population data is available by Statistics South Africa for the most recent Census in 2011 at different levels of a spatial hierarchy: province, municipality, main-place (~towns), sub-place (suburbs or villages) and small area. A small area is created by combining all the enumeration areas with a population of <500 with adjacent enumeration areas within the same sub-place. The centroid of a small area polygon was used to represent the origin of a trip to a police station (destination). The locations of South African Police Service (SAPS) stations were obtained from the Department of Public Service and Administration. Major dams and rivers were identified from the 1:50000 topographic data from the Chief Directorate: National Geo-spatial Information (http://www.cdngiportal.co.za/cdngiportal/) and steep mountainous areas from the 30-m digital elevation model (DEM) from the Shuttle Radar Topography Mission (SRTM) (https://earthexplorer.usgs.gov/). Satellite imagery from Google Maps was used throughout the study for visualization, orientation and general route verification. Preparation of the travel network and distance calculations were done in TransCAD 4 (https://www.caliper.com/tcovu.htm). The LISA statistic was calculated in GeoDa (https://geodacenter.github.io/index.html). Maps included in this paper were prepared in Maptitude (https://www.caliper.com/maptovu.htm) and GeoDa.

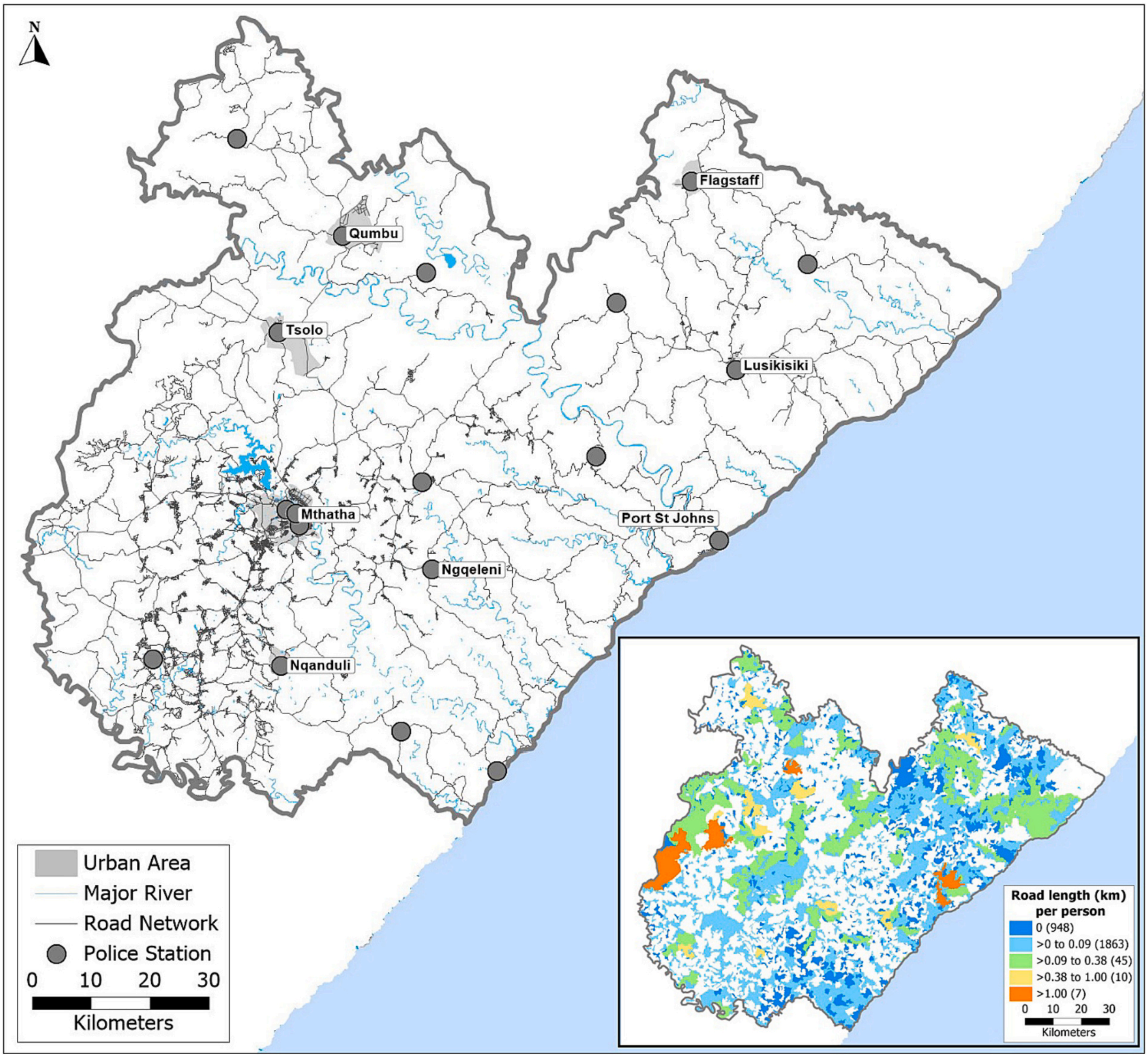


Fig. 2. Urban areas, police stations and the road network in O. R. Tambo District Municipality. The inset map indicates road density per person for each small area (Data sources: see 3.3).

4. Results

4.1. *Preparing the augmented travel network*

The first step was to create a centroid for each non-urban small area in the Census data. These centroids represent the origin of a trip to a police station by anyone living in that small area. Next, each centroid and each police station (destination) was connected to its nearest road segment (Fig. 4a), and the connecting links were added to the road network (Fig. 4b).

Next, a triangular irregular network (TIN) was created. A TIN is a vector based representation of a surface, comprising a number of points connected by lines (arcs) to create triangular shapes (Lloyd, 2010). In our case, the TIN comprises straight lines between small area centroids in non-overlapping triangles (Fig. 4c). TIN lines were removed if they crossed (or intersected) major dams, rivers or steep mountainous terrain (Fig. 4d). The latter was identified by converting a 30-m DEM into a raster slope. Areas with a slope steepness of >16 degrees (strong slope) were considered to be mountainous and not suitable for travelling (https://geographyfieldwork.com/SlopeSteepnessIndex.htm). The remaining TIN lines were added to the travel network and serve as a proxy for unmapped dirt roads and footpaths. We refer to this network as the augmented travel network (Fig. 4e). Fig. 5 displays the road network and the augmented travel network for the O.R. Tambo District Municipality. A comparison (by visual inspection) confirms that the augmented travel network has much wider coverage than the road network.

4.2. *Results for the three distance measurement techniques*

4.2.1. *Geographic accessibility by the population*

In Fig. 6, small area centroids are coloured according to their distance from the closest police station. The three maps – (A), (B) and (C) – represent the three different distance measuring techniques. An estimated coverage area for each police station was calculated based on the respective distance measurement technique: ‘normal’ straight-line Thiessen polygons, a coverage area based on routes in the road network, and a coverage area based on routes in the augmented travel network. From visual inspection, it is clear that most small area centroids are within a straight-line distance of 16 km from the closest police station (blue dots). For the other two techniques, a significantly larger number of small area centroids are further away and even more centroids are further than 24 km from a police station (yellow and red dots).

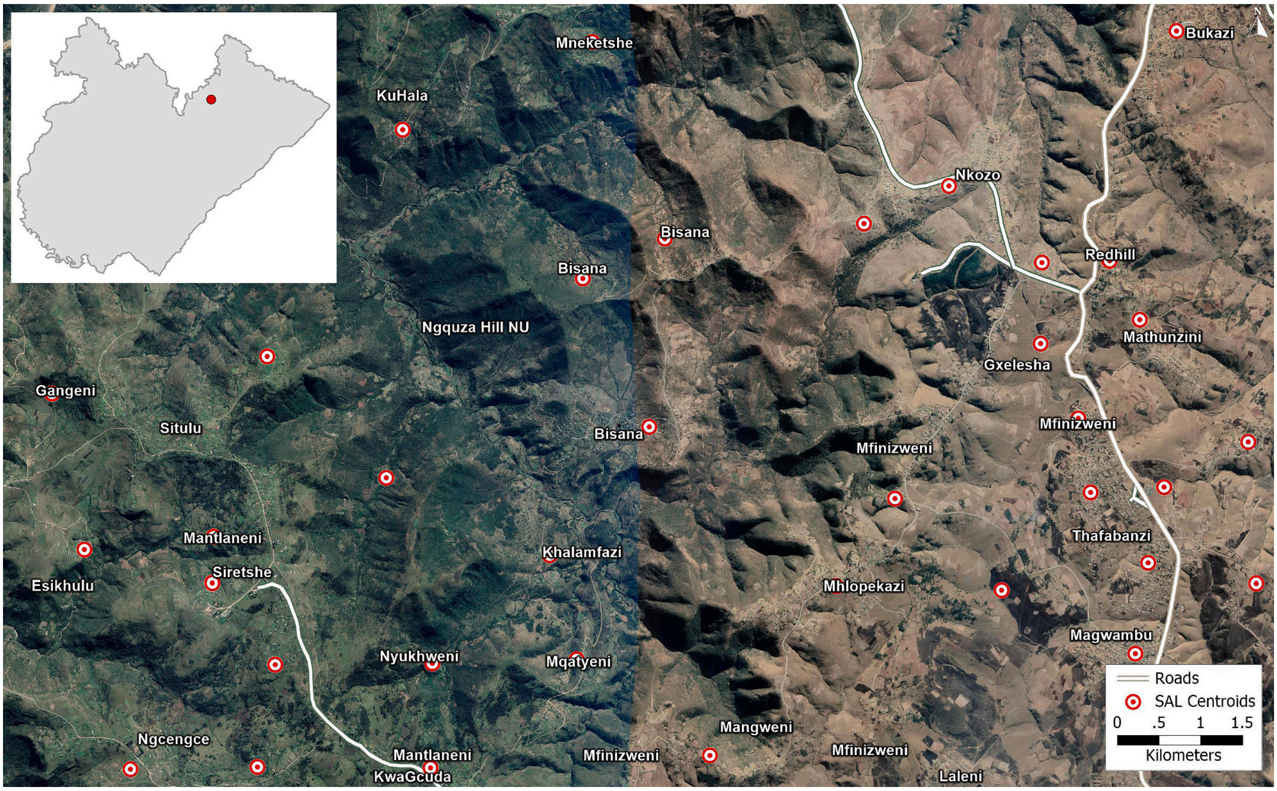


Fig. 3. Satellite image revealing unmapped footpaths and dirt roads in a non-urban area of the O. R. Tambo District Municipality (Data sources: see 3.3).

Areas 1, 2, 3 and 4 identified by the ovals on the maps show that in comparison to routes based on the road network (B), routes based on the augmented travel network (C) improved the accessibility results in remote non-urban areas where there are very few government maintained roads. This shows that geographic accessibility to services in non-urban areas is underestimated for the road network techniques.

Area 5 shows that a much longer distance to the closest police station was calculated for measurement techniques B and C (>32 km), compared to A, the straight-line distance technique (8 to 24 km). A major river (see Fig. 2) runs between the settlements (small area centroids) and the two nearby stations. Since B and C consider natural barriers, such as major rivers, in their routes, the travel distances for them are notably longer.

4.2.2. *Geographic accessibility of the police station*

Fig. 7 shows, for each of the three distance measuring techniques, how many people live within different threshold distances from the nearest police station. At a threshold distance of approximately 16 km, the three techniques are very similar, but for smaller or larger threshold distances they differ significantly. The graph reveals that straight-line distances (orange line in the graph) overestimate the population within 0–8 and 8–16 km from a police station and underestimate the population for larger threshold distances. When comparing the population covered by the road network (blue line) with that covered by the augmented travel network (black line), one can see that for smaller threshold distances the augmented travel network ‘covers’ a larger population because the TIN connects the population ‘beyond’ the road network. For larger threshold distances, the augmented travel network ‘corrects’ the overestimation of population coverage due to from routes based on the road network only: in the latter the distance between origin and road network is ignored, while the former incorporates this distance through the TIN augmentation.

When comparing the number of people within a distance band against the average for the three techniques, it can be observed that the population size calculated for the augmented travel network is closest to the average. See Table 1. This suggests that shortest routes based on the augmented travel network provide a more generalised (or averaged out) result compared to the others, and that the other techniques tend to over-, or underestimate the size of the population.

4.2.3. *Spatial pattern of the variance between the three distances*

For each small area centroid (origin), we calculated the variance between the three distances to the nearest police station (destination). A high value indicates high variance between the three distances, and vice versa. In Fig. 8, the small area centroids are categorised according to the variance calculated for the three distances. Low variance values (blue and purple dots) can be observed for the majority of small area centroids in proximity of the urban areas where there are police stations, indicating that there is no significant difference in the three distances. The high variance values (yellow and red dots) are further away from the urban areas. Closer inspection of the map shows that these dots are sometimes separated by a river from the urban areas.

In Fig. 9A, only small area centroids with variance values above the average of 26 (see Table 2) are displayed. This represents 591 (22.7%) of the 2604 small area centroids in the study area, or 232,135 people (19.45%) from a total population of 1,193,688 in the rural parts of O.R. Tambo District Municipality. A similar pattern, with a less distinct spatial distribution, was observed when small area centroids with variance values above the median of 4.6 were displayed, see Fig. 9B. A total of 1296 (49.8%) small area centroids are included on the map, which represents 535,823 people, i.e., 44.89% of the population in the district. These results confirm that there is significant local variation in the three distance measurement techniques. The spatial distribution of variances further confirms that the distance measurement technique is sensitive to the urban / rural classification of an area: the choice of distance measurement technique in non-urban areas has a more significant impact than in urban areas.

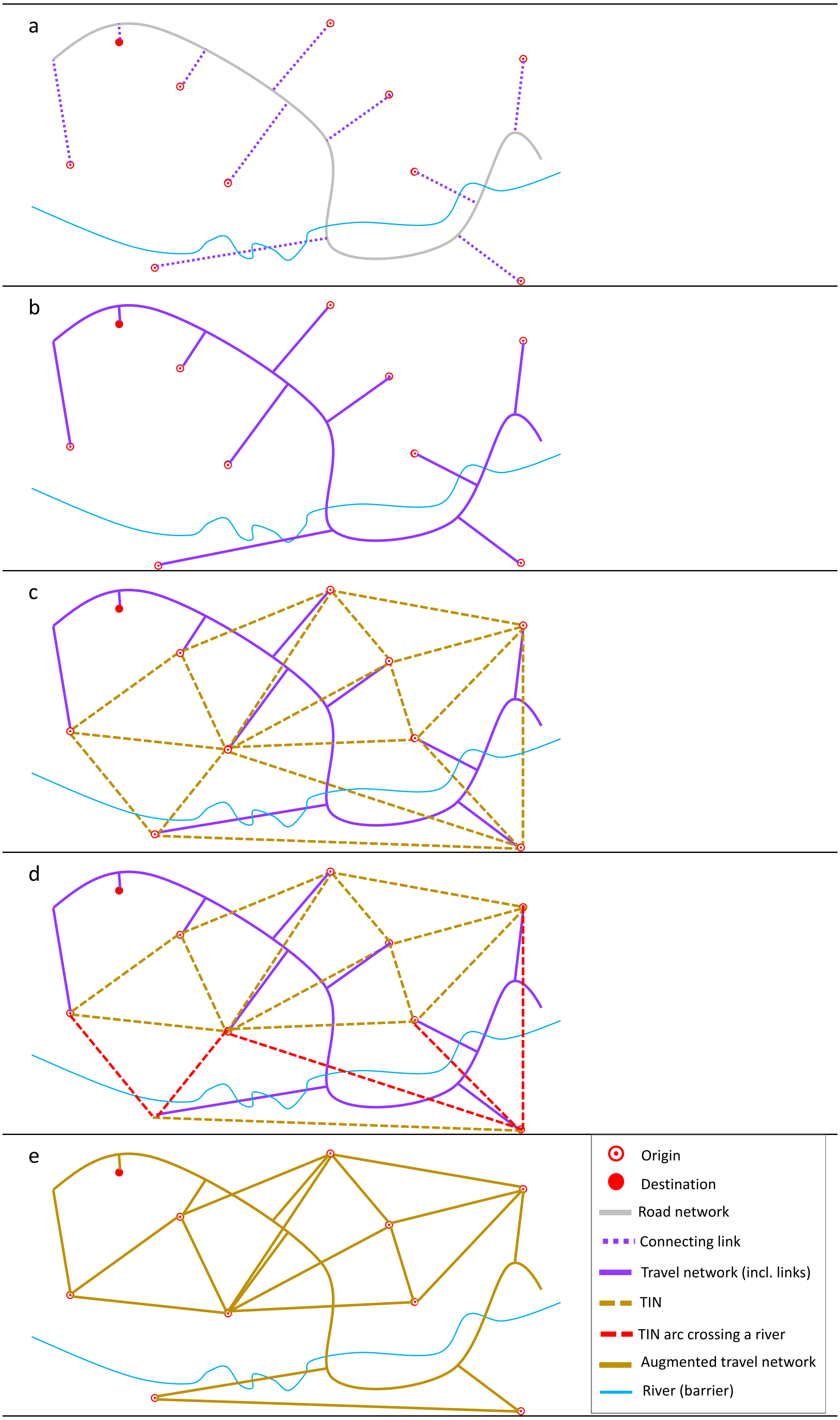


Fig. 4. Preparation of the augmented travel network.

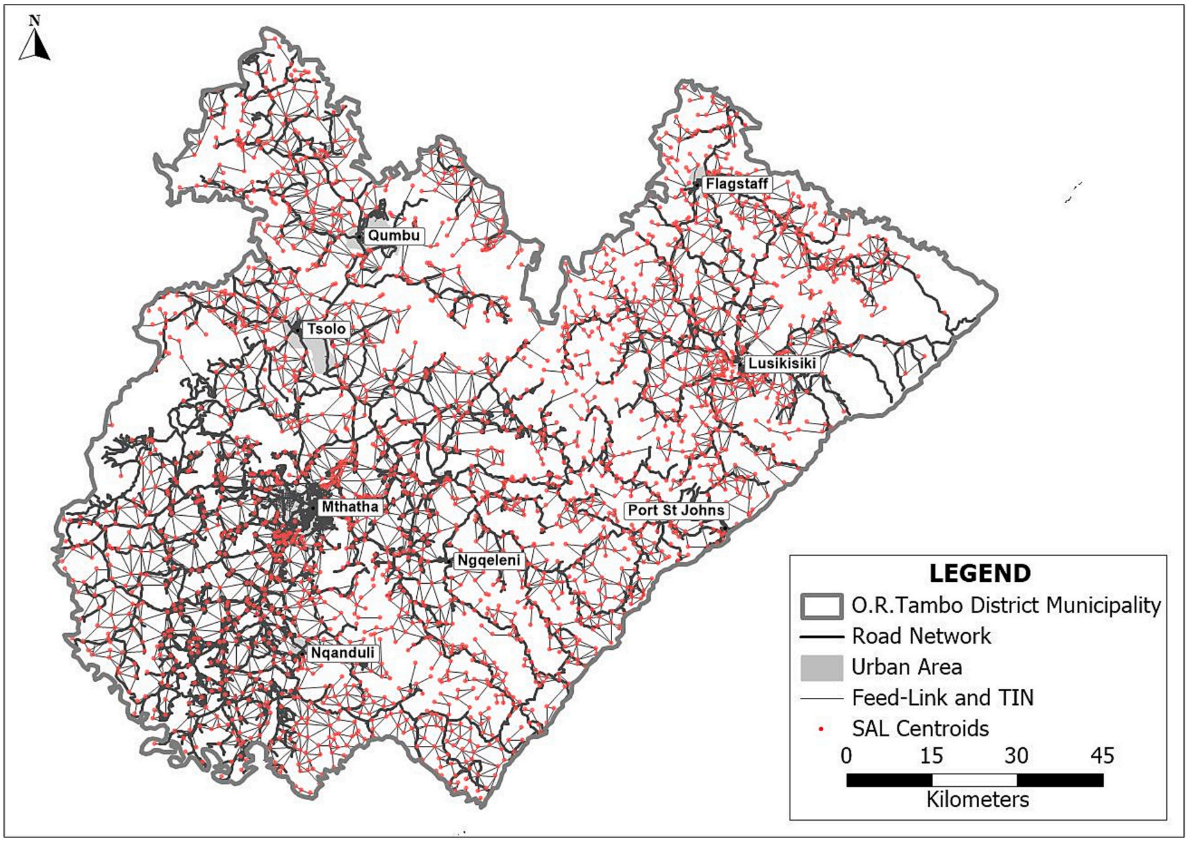


Fig. 5. The road network (thick black lines) and the augmented travel network (thin grey lines), which connects the centroids of small areas (orange dots) to the road network and is augmented with a TIN (Data sources: see 3.3).

4.2.4. Total squared error (TSE) mean squared error (MSE)

The TSE and MSE measure the overall error rate of the augmented travel network (C) against each of the other two techniques (A and B). These values show how the accessibility measurements of A and B differ from C. The TSE was calculated as follows:

$$TSE = \sum \left(a_{i-}a_{Aug,i}\right)^2$$

Where $a_{Aug,i}$ is the distance measured per small area centroid for the augmented travel network technique (C) and a_i represents the distance measure for the other two techniques; A and B. MSE is the TSE divided by the number of distances.

Both the TSE and MSE between C and A are higher than corresponding values between C and B, indicating greater distance variation between the accessibility measurements of C and A (See Table 3). The lower TSE and MSE between C and B are the result of both techniques using a road network, considering natural barriers such as mountains and large rivers which influences the distance measurements in rural areas; this is not the case for straight-line distances.

4.2.5. Spatial pattern of the error rate per locality

To identify the spatial patterns of these error rates per locality, the following formula was used:

$$error\ rate\ per\ locality = \frac{a_i - a_{Aug,i}}{a_{Aug,i}}$$

A local indicator of spatial association (LISA) statistic was used to detect hot or cold spots of high and low error rates as well as outliers. A local Moran's I was calculated to determine whether adjacent small area centroids have similar error rates and whether they are part of a spatial cluster. The LISA statistic groups the small area centroids into four categories: High-High (dark red), Low-Low (dark blue) and the outlier categories of Low-High (light blue) and High-Low (light red), see Fig. 10. For the purpose of this study, High-High represents a high error rate of small area centroids that are also in close proximity to each other. Low-Low represents clusters of low error rates.

When comparing the augmented network technique (C) against the other two techniques, the highest number of hot-spot clusters with High-High and Low-Low error rates was between the augmented travel network (C) and the straight-line distance (A), see Fig. 10a. The clusters are generally situated in areas where the augmented travel network improved the level of accessibility because small area centroids are connected to the road network and augmented with the TIN. Similar to the variances, these error rate clusters are more prominent in non-urban parts of the study area, suggesting significant differences in the distance calculations and therefore accessibility results. A positive local Moran's I of 0,518 also emphasizes a high level of spatial clustering. 740 small area centroids are not significant (grey dots), i.e., there is no significant difference in the distance calculations, but the map shows that many of these centroids are situated closely to a police station.

Fewer spatial clusters are observed when the error rate per locality between the augmented travel network (C) and the road network (B) is determined, see Fig. 10b. The clusters (High-High and Low-Low) in this map are more concentrated in certain areas, compared to the map in Fig. 10a. An even larger number of small area centroids (1464) is not significant.

4.2.6. Lorenz curve and Gini coefficient for measuring inequality

The Lorenz curve provides a visual representation of the cumulative distribution of population vs access (distance) for the three techniques, see Fig. 11. The grey line indicates maximum equality.

The Gini coefficient calculates a value between 0 and 1, where 0 indicates perfect equality and 1 indicates complete inequality: 0 indicates that all people have the same level of access (distance) to the nearest

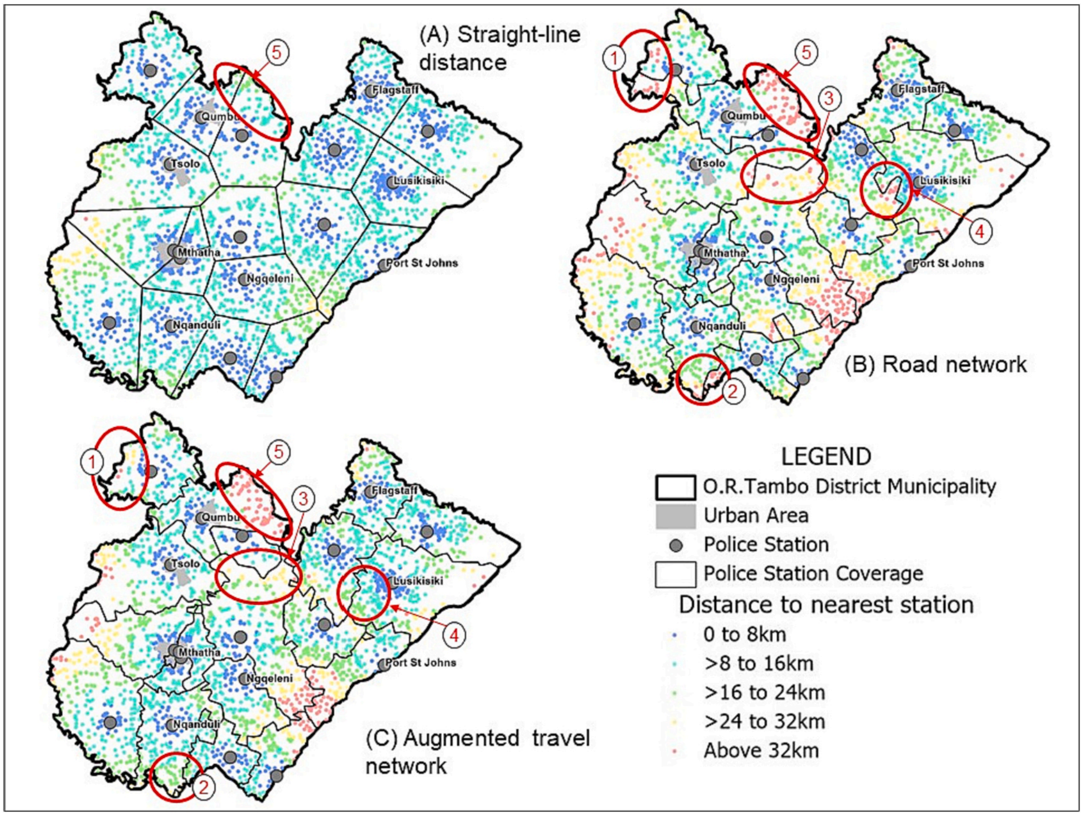
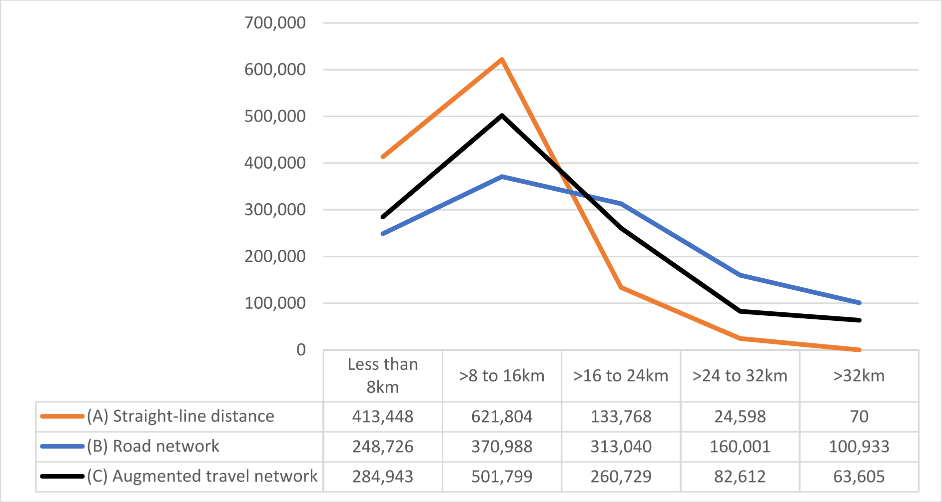


Fig. 6. Geographic accessibility by the population: shortest distances between small area centroids and their nearest police station based on the three distance measurement techniques.



	Less than 8km	>8 to 16km	>16 to 24km	>24 to 32km	>32km
(A) Straight-line distance	413,448	621,804	133,768	24,598	70
(B) Road network	248,726	370,988	313,040	160,001	100,933
(C) Augmented travel network	284,943	501,799	260,729	82,612	63,605

Fig. 7. Geographic accessibility of the service provider: size of the population within different threshold distances based on the three distance measurement techniques.

police station.

Results from the three techniques indicate a high level of inequality in each case. The straight-line distance technique (A) measured the lowest level of inequality (0.43), see Table 4. The road network technique (B) measured the highest level of inequality, with a Gini coefficient of 0.49, followed by the travel network (C) with a value of 0.48, indicating a balance between the A and B. In other words, if the straight-line distance technique is used, governments underestimate inequality and overestimate service delivery.

5. Discussion

In this research, we implemented the government recommended approach for measuring geographic accessibility in rural areas, and analysed and compared the results to other techniques. Geographic accessibility results were compared for three different types of distances. As expected, the TSE and MSE confirmed that the distance differences between the augmented travel network (C) and straight-line distances (A) are larger than those between the augmented travel network (C) and the road network (B).

Table 1
Size of the population within different distance bands, based on the three techniques and compared to the average population size for the band. The smallest difference to the population average is shaded in grey.

Population average based on the three different techniques	315 706	498 197	235 846	89 070	54 869
Technique	**Less than 8km**	**>8 to 16km**	**>16 to 24km**	**>24 to 32km**	**>32km**
(A) Straight-line distance	97 742	123 607	-102 078	-64 472	-54 799
(B) Road network	-66 980	-127 209	77 194	70 931	46 064
(C) Augmented travel network	-30 763	3 602	24 883	-6 458	8 736

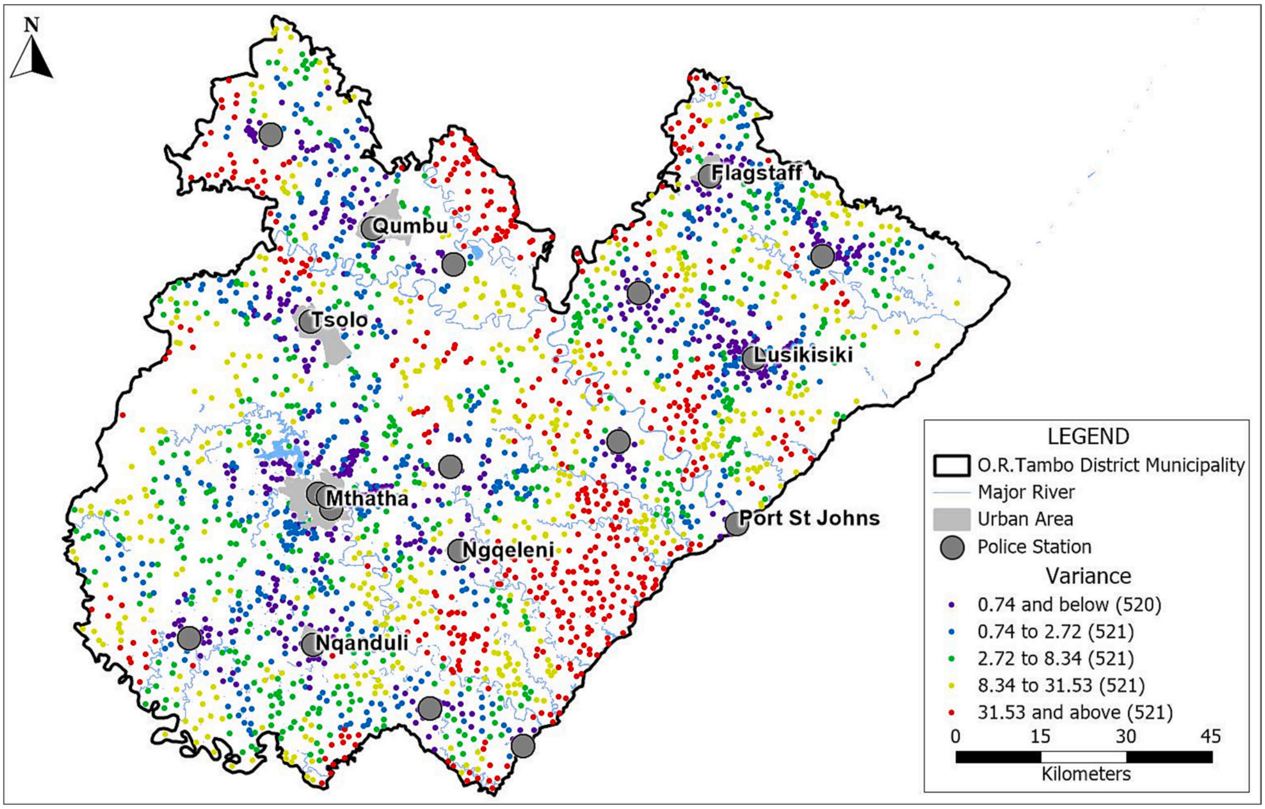


Fig. 8. Variance between the three distances calculated from each small area centroid (origin) to the nearest police station (destination).

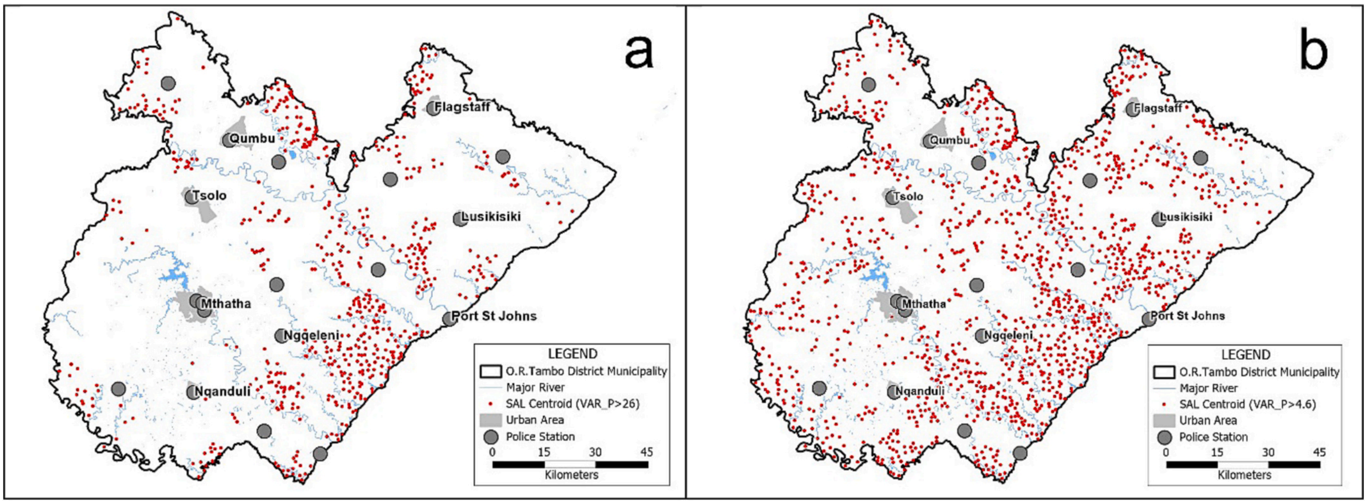


Fig. 9. Small area centroids with high variance in the three distances: above the average of 26 (a) and above the median of 4.6 (b).

Table 2
Descriptive statistics of the variance values for the three distances.

Variance	Value
Average	26.0
Median	4.6
Min	0.0
Max	865.7
90% Percentile	58.5
80% Percentile	31.5
20% Percentile	0.7

Table 3
TSE and MSE of the augmented travel network (C) against both (A), straight-line distance and (B) the road network.

Comparison	TSE	MSE
(C) Augmented travel network and (A) Straight-line distance	168,035	64.5
(C) Augmented travel network and (B) Road network	100,773	38.7

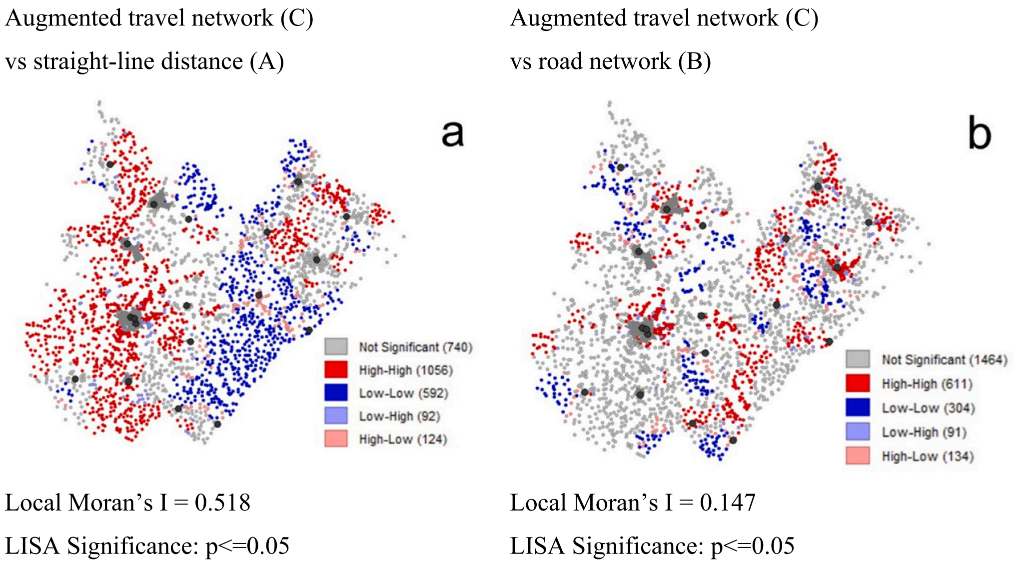
When there are physical barriers in an area, which is the case in the study area, straight-line distances do not correlate well with route-based distances (Boscoe et al., 2012); straight-line distance calculations tend to overestimate catchment areas. The results in 4.2.1 confirm that straight-line distances are not appropriate for measuring geographic accessibility by population (shortest distance to a police station) in the study area: many people have to travel >16 km to a police station but the straight-line distances do not reflect this. Similar to what others have found (e.g., Mavoa et al., 2012; Polo et al., 2013), straight-line distances overestimate the population within smaller distances, e.g., 0–8 km from a police station in this study, and underestimate the population for larger threshold distances, such as 24–32 km and also distances further than 32 km in the study area. Based on the Lorenz curves for the Gini coefficient of the three distance measurement techniques, straight-line distances underestimate inequality and overestimate service delivery. Therefore, if decision makers in government consider accessibility measured with straight-line distances, this could lead to a slow-down in the much needed roll-out of services in previously disadvantaged non-urban areas.

A comparison of the population averages shows that route-based distances in an augmented travel network provide a generalised result, while the other two techniques either over-, or underestimate the population coverage. For example, the road network technique underestimates the accessibility of the population in non-urban areas in comparison to the augmented travel network. This can be explained by the fact that when measuring geographic accessibility based on a road network, the distances between villages and their nearest road segment are not considered. Furthermore, natural barriers that may occur on these routes are not considered, nor the fact villages may be connected by footpaths. These results also confirm that the distance measuring technique impacts geographic accessibility results (Higgs et al., 2012; La Rosa, 2013). Based on the above, one can conclude that the government recommended method of calculating distances in a road network augmented with a TIN is suitable in the study area.

The three techniques are specifically sensitive to a threshold distance of 8 km because the size of the population within this threshold distance differed significantly between the three techniques. For a threshold distance of 24 km, the differences were not as significant. This confirms that the threshold distance should not be specified universally, but rather in consideration of local specificities (Boscoe et al. 20,212). For a threshold of 8 km or less, the government recommended method of augmenting the travel network with a TIN improved the accuracy of accessibility results, confirming the suitability of the method for the threshold distance specified in the guidelines.

Similar to Boscoe et al. (2012), significant localized differences were detected in the variance between the distances measured for each small area centroid in the study area. A high variance was observed in non-urban areas, indicating substantial differences in the distance calculations for small area centroids in these areas. Closer to the urban areas, the variance was significantly lower, i.e. smaller differences in the distance calculations. Similarly, LISA maps showed more clusters of error rates per locality in non-urban areas. These results can be explained by the fact that there are more roads in and around urban areas and also because these roads are typically maintained by a municipality and therefore included in the data. This confirms that the government recommended method improves the accuracy of accessibility results in data poor rural areas. The method could provide a much needed scalable approach in data poor areas, also in other countries (Noor et al., 2006).

There are some limitations to our study. Firstly, our method of removing TIN edges, may not have considered all natural barriers. To improve the accuracy, one could consider other methods, such as the convex path presented by Hong and Murray (2013) or one could employ machine learning techniques to detect barriers on imagery or


Augmented travel network (C) vs straight-line distance (A)

Augmented travel network (C) vs road network (B)

Local Moran's I = 0.518
LISA Significance: p<=0.05

Local Moran's I = 0.147
LISA Significance: p<=0.05

Fig. 10. LISA results and Moran's I for the augmented travel network technique (C) against the straight-line distance (A) and road network (B) techniques respectively.

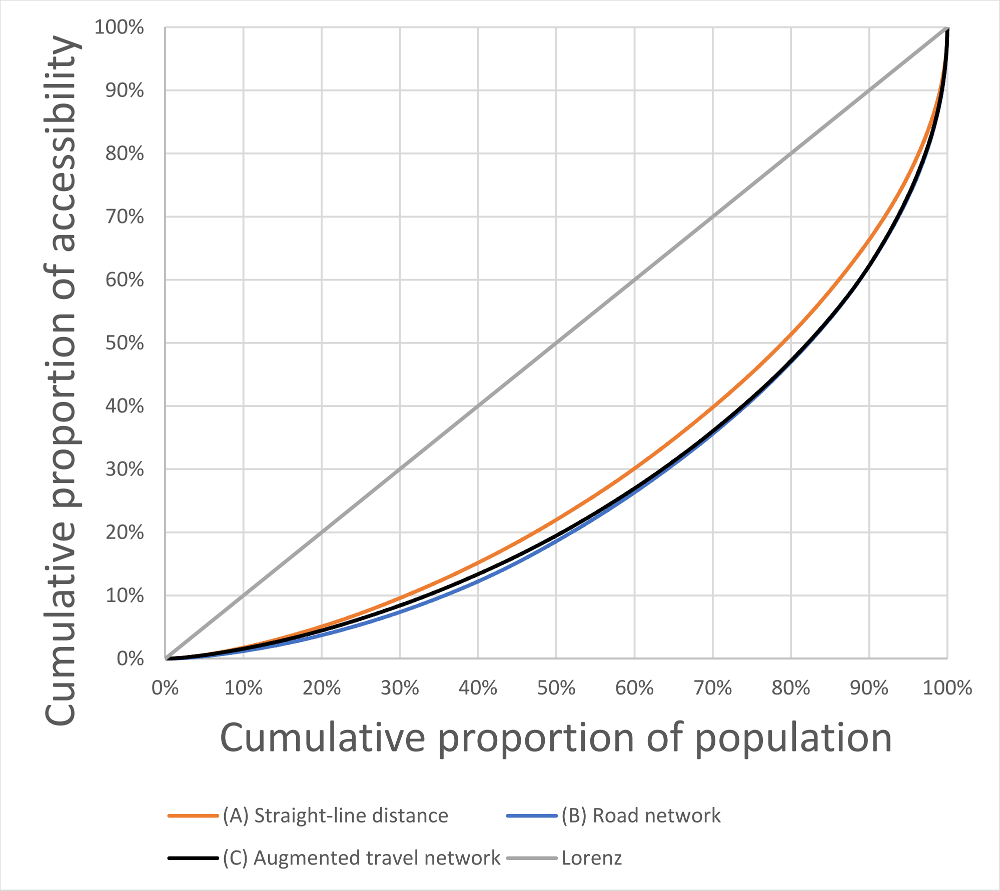


Fig. 11. Lorenz curve showing the cumulative level of accessibility over population based on the three distance measurement techniques.

Table 4
Gini coefficient for the three distance measurement techniques.

Distance measurement technique	Gini coefficient
(A) Straight-line distance	0.43
(B) Road network	0.49
(C) Augmented travel network	0.48

incorporating additional topographical data sources. If the situation warrants it, one could even consider manual processing but care should be taken not to spend so much time that the method is not cost effective anymore. Secondly, because rural areas tend to be sparsely populated, they tend to have larger enumeration areas and consequently the small areas are also larger. The accuracy with which a centroid represents the population, depends on how the population is distributed across an area (the checkerboard problem). For example, a single village situated in a ‘corner’ of the enumeration area is not accurately represented by a centroid in the ‘middle’ of the enumeration area. In this study, centroids were determined based on the geographic centre of a polygon. One could experiment with different algorithms to determine whether they influence the geographic accessibility results. Alternatively, populated areas could be detected from satellite areas.

Our study reveals the impact of geospatial data scarcity on geographic accessibility results and confirms the complexity of measuring equitable access in countries with vastly different urban and rural areas, not only in developing countries with rural areas similar to those in this study area (Gabrysch et al., 2011), but also in sparsely populated developed countries such as Canada (Law et al., 2013). Machine learning techniques that employ raster-based image recognition algorithms to identify and map footpaths could also contribute to improve data paucity (Fabris-Rotelli et al., 2022).

Our study worked with distance-based thresholds for accessibility, as this is how they are currently specified in South African guidelines (Green and Argue, 2012). Working with distance only (not considering travel speeds on different kinds of roads) is a limitation, especially in rural areas, where people often walk to public transport pickup points and then travel the last part of the route in a vehicle – to travel 20 km to a police station in a rural area will take much longer than 20 km in an urban area. Further research should look into how one could automatically assign a travel mode to different segments of the TIN so that accessibility can be measured in terms of travel time. Additionally, a comparison of distance-based accessibility results vs travel time-based accessibility results would provide evidence and justification for changing policies accordingly.

6. Conclusion

In this research, we implemented the government recommended approach for measuring geographic accessibility in rural areas, and analysed and compared the results to other techniques. While this method is recommended, a description of an implementation of this method has not been published, and the suitability of its accessibility results has not been assessed.

As a first contribution, the paper provides a detailed description of how the recommended method was implemented in the study area. Next, we presented the results of a comparison of geographic accessibility computed based on three different distance measuring techniques: straight-line distances, shortest routes in a road network and shortest routes in an augmented travel network (small area centroids are connected to a road network and augmented with a TIN). In the government recommended approach, the TIN serves as a proxy for dirt roads and footpaths that are seldomly mapped, change frequently or vary depending on the season. The augmented travel network proved to be useful for threshold values of >8 km because it provides a more generalised result, reducing under- or over-estimations observed in other techniques.

Results show that one needs to consider the threshold distance when choosing a suitable distance measuring technique for geographic accessibility. Furthermore, the suitability of a distance measurement technique depends on the proximity to or distance from urban areas. Closer to urban areas, the techniques deliver similar results, but remote areas are more sensitive to the choice of distance measuring technique. The results of our study could inform future guidelines for measuring accessibility in South Africa.

Further work could consider travel time, instead of distance only, by specifying travel speeds for different parts of the augmented travel network, however, the challenge would be to determine which of the TIN arcs are accessible by vehicles (dirt roads) and which not (footpaths) so that one can differentiate between in-vehicle and out-of-vehicle travel. The results could inform the revision of policies which currently specify accessibility based on distance only. The implementation of the government recommended method could be further refined by more accurately identifying barriers, for example by employing machine learning techniques to imagery (such as least cost path analysis where various raster-based calculations could be used to identify potential footpaths) or incorporating other topographical data sources. In this study, small area centroids were calculated based on the geographic centre of a polygon. It would be interesting to see whether a different way of determining centroids, such as a weighted centroid calculation which considers density patterns inside a polygon, would influence the results.

The results of this study are important for policy makers who are challenged with uplifting and integrating rural areas, and have to measure geographic accessibility in the absence of data. They also need to understand how different methods for measuring geographic accessibility may affect their planning of supply and demand. The study was done for the O.R. Tambo District Municipality in the Eastern Cape of South Africa, however, results are also useful for other rural areas of Africa where villages are scattered across the landscape and are not in close proximity to roads included in government maintained datasets.

Data availability

Data will be made available on request.

Acknowledgements

The authors appreciate the guidance and support of Mr Trevor Holdsworth from the Department of Public Service and Administration (DPSA) in an earlier project related to this research.

References

Al-Taiar, A., et al., 2010. Physical accessibility and utilization of health services in Yemen. Int. J. Health Geogr. 9 (1), 1–8.

Anselin, L., 1995. Local indicators of spatial association—LISA. Geogr. Anal. 27 (2), 93–115.

Apparicio, P., et al., 2008. Comparing alternative approaches to measuring the geographical accessibility of urban health services: distance types and aggregation-error issues. Int. J. Health Geogr. 7 (7) https://doi.org/10.1186/1476-072X-7-7.

Apparicio, P., et al., 2017. The approaches to measuring the potential spatial access to urban health services revisited: distance types and aggregation-error issues. Int. J. Health Geogr. 16, 32. https://doi.org/10.1186/s12942-017-0105-9.

Azmoodeh, M., et al., 2021. Proposing an integrated accessibility-based measure to evaluate spatial equity among different social classes. Environ. Plan. B Urban Anal. City Sci. 48 (9), 2790–2807.

Baloyi, E., et al., 2017. Evaluating public ambulance service levels by applying a GIS based accessibility analysis approach. South Afr. J. Geomat. 6 (2), 172–183. https://doi.org/10.4314/sajg.v6i2.3.

Blanford, J.I., et al., 2012. It's a long, long walk: accessibility to hospitals, maternity and integrated health centers in Niger. Int. J. Health Geogr. 11 (1), 24.

Boscoe, F.P., Henry, K.A., Zdeb, M.S., 2012. A Nationwide comparison of driving distance versus straight-line distance to hospitals. Prof. Geogr. 64.

Church, R.L., Murray, A.T., 2009. Business Site Selection, Location Analysis and GIS. John Wiley & Sons, Hoboken, NJ.

Cromley, G.A., 2019. Measuring differential access to facilities between population groups using spatial Lorenz curves and related indices. Trans. GIS 23 (6), 1332–1351.

Department of Public Service and Administration (DPSA), 2012. Guideline: Improving geographic access to government service points [Online]. Available. http://www.dpsa.gov.za/dpsa2g/documents/sdot/2012/DPSA_Guideline%20on%20Access%20to%20Service%20Points.pdf [16 August 2016].

Department of Transport, 2007. Rural Transport Strategy for South Africa [Online]. Available. https://www.gov.za/sites/default/files/gcis_document/201409/ruraltransportstrategy-forsa2007new0.pdf [6 August 2023].

Drezner, Z., Hamacher, H.W., 2002. Facility Location Application and Theory. Springer, Berlin.

du Plessis, E., 2018. An overview of land use management and the role of traditional leaders in terms of SPLUMA. KAS/NWU conference: Traditional leadership and authority versus local government, Potchefstroom, 13 & 14 September 2018. http://law.nwu.ac.za/sites/law.nwu.ac.za/files/files/Elmien%20du%20Plessis%20SPLUMA%20and%20TL.pdf [16 October 2022].

Fabris-Rotelli, I., Wannenburg, A., Maribe, G., Thiede, R., Vogel, M., Coetzee, M., Sethaelo, K., Selahle, E., Debba, P., Rautenbach, V., 2022. An informal road detection neural network for societal impact in developing countries. ISPRS Ann. Photogramm. Remote Sens. Spatial Inform. Sci. V-4-2022, 267–274. https://doi.org/10.5194/isprs-annals-V-4-2022-267-2022.

Feikin, D.R., et al., 2009. The impact of distance of residence from a peripheral health facility on pediatric health utilisation in rural western Kenya. Tropical Med. Int. Health 14 (1), 54–61.

Gabrysch, S., et al., 2011. The influence of distance and level of care on delivery place in rural Zambia: a study of linked national data in a geographic information system. PLoS Med. 8, e1000394.

Green, C., Argue, T., 2012. CSIR Guidelines for the Provision of Social Facilities in South African Settlements, 1st ed. CSIR Built Environment, Pretoria.

Green, C., et al., 2014. Planning for emergency services using GIS-based geographic accessibility analysis. Town Reg. Plan. 64, 53–64.

Guzman, L.A., et al., 2017. Assessing equity in transport accessibility to work and study: the Bogotá region. J. Transp. Geogr. 58, 236–246.

Higgs, G., Fry, R., Langford, M., 2012. Investigating the implications of using alternative GIS-based techniques to measure accessibility to green space. Environ. Plan. B Plan. Des. 39, 326–343.

Hong, I., Murray, A.T., 2013. Efficient measurement of continuous space shortest distance around barriers. Int. J. Geogr. Inf. Sci. 27 (12), 2302–2318. https://doi.org/10.1080/13658816.2013.788182.

Hundt, G.L., et al., 2012. The provision of accessible, acceptable health care in rural remote areas and the right to health: Bedouin in the North East region of Jordan. Soc. Sci. Med. 74 (1), 36–43.

Ilie, C.M., Brovelli, M.A., Coetzee, S., 2019. Monitoring SDG 9 with global open data and open software – a case study from rural Tanzania. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. XLII-2/W13, 1551–1558. https://doi.org/10.5194/isprs-archives-XLII-2-W13-1551-2019.

Kahabuka, C., et al., 2011. Why caretakers bypass primary health care facilities for child care-a case from rural Tanzania. BMC Health Serv. Res. 11 (1), 1–10.

Kwan, M.P., 2009. From place-based to people-based exposure measures. Soc. Sci. Med. 69 (9), 1311–1313. https://doi.org/10.1016/j.socscimed.2009.07.013.

La Rosa, D., 2013. Accessibility to greenspaces: GIS based indicators for sustainable planning in a dense urban context. Ecol. Indic. 42, 122–134.

Law, M.R., et al., 2013. The geographic accessibility of pharmacies in Nova Scotia: CPJRPC. Can. Pharm. J. 146 (1), 39–46.

Lloyd, C., 2010. Spatial Data Analysis: An Introduction for GIS Users. Oxford University Press, Oxford, NY.

Lorenz, M.O., 1905. Methods of measuring the concentration of wealth. Publ. Am. Stat. Assoc. 9 (70), 209–219.

Lucas, K., 2011. Making the connections between transport disadvantage and the social exclusion of low income populations in the Tshwane Region of South Africa. J. Transp. Geogr. 19 (6), 1320–1334.

Luo, W., Wang, F., 2003. Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. Environ. Plan. B Plan. Des. 30 (6), 865–884. https://doi.org/10.1068/b29120.

Masiya, T., et al., 2019. Assessing service delivery: public perception of municipal service delivery in South Africa. Theor. Emp. Res. Urban Manag. 14 (2), 20–40.

Mavoa, S., et al., 2012. GIS based destination accessibility via public transit and walking in Auckland, New Zealand. J. Transp. Geogr. 20, 15–22.

Mcgrail, M.R., Humphreys, J.S., 2009. Measuring spatial accessibility to primary care in rural areas: improving the effectiveness of the two-step floating catchment area method. Appl. Geogr. 29, 533–541.

Miller, H.J., 2007. Place-based versus people-based geographic information science. Geogr. Compass 1 (3), 503–535. https://doi.org/10.1111/j.1749-8198.2007.00025.x.

Mokgalaka, H., et al., 2013. GIS-based accessibility analysis-a mixed method approach to determine public primary health care demand in South Africa. In: Global Geospatial Conference, a AfricaGIS 2013 and GSDI14 Joint Conference at: United Nations Economic Commission of Africa Centre, Addis Ababa, Ethiopia.

Müller, I., et al., 1998. The effect of distance from home on attendance at a small rural health Centre in Papua New Guinea. Int. J. Epidemiol. 27 (5), 878–884.

Ndirangu, J., et al., 2009. Levels of childhood vaccination coverage and the impact of maternal HIV status on child vaccination status in rural KwaZulu-Natal, South Africa. Tropical Med. Int. Health 14 (11), 1383–1393.

Neutens, T., 2015. Accessibility, equity and health care: review and research directions for transport geographers. J. Transp. Geogr. 43, 14–27.

Neutens, T., et al., 2010. Equity of urban service delivery: a comparison of different accessibility measures. Environ. Plan. A 42 (7). https://doi.org/10.1068/a4230.

Noor, A., et al., 2003. Defining equity in physical access to clinical services using geographical information systems as part of malaria planning and monitoring in Kenya. Tropical Med. Int. Health 8 (10), 917–926.

Noor, A.M., et al., 2006. Modelling distances travelled to government health services in Kenya. Tropical Med. Int. Health 11 (2), 188–196.

Poku-Boansi, M., et al., 2010. Combating maternal mortality in the Gushegu district of Ghana: the role of rural transportation. J. Sustain. Dev. Afr. 12 (5), 274–283.

Polo, G., Acosta, C.M., Dias, R.A., 2013. Spatial accessibility to vaccination sites in a campaign against rabies in São Paulo city, Brazil. Prev. Vet. Med. 111, 10–16.

Pozzi, F., Robinson, T.P., 2008. Accessibility mapping in the Horn of Africa: Applications for livestock policy. In: IGAD LPI Working Paper 11–08.

Rodrigue, J.-P., Comtois, C., Slack, B., 2009. The Geography of Transport Systems. Routledge, London; New York.

Rofé, Y., et al., 2015. Accessibility and Social Equity in Tel-Aviv Metropolitan Area-Examination of the Current Conditions and Development Scenarios.

Salonen, M., et al., 2012. Critical distances: comparing measures of spatial accessibility in the riverine landscapes of Peruvian Amazonia. Appl. Geogr. 32, 501–513.

SAPS (South African Police Service), 2023. Website [Online]. Available. www.saps.gov.za [6 August 2023].

Schoeps, A., et al., 2011. The effect of distance to health-care facilities on childhood mortality in rural Burkina Faso. Am. J. Epidemiol. 173 (5), 492–498.

Statistics South Africa, 2012. Census 2011 Metadata. Statistics South Africa, Pretoria.

Statistics South Africa, 2019. Non-financial census of municipalities for the year ended 30 June 2019. http://www.statssa.gov.za/publications/P9115/P91152019.pdf [16 October 2022].

Statistics South Africa, 2021. SDG Goal Tracker. https://south-africa.goaltracker.org/platform/south-africa/explorer?indicator=1.2.2 [15 September 2021].

Stock, R., 1983. Distance and the utilization of health facilities in rural Nigeria. Soc. Sci. Med. 17 (9), 563–570.

Tanser, F., 2006. Methodology for optimising location of new primary health care facilities in rural communities: a case study in KwaZulu-Natal, South Africa. J. Epidemiol. Community Health 60 (10), 846–850.

Tanser, F., et al., 2001. New approaches to spatially analyse primary health care usage patterns in rural South Africa. Tropical Med. Int. Health 6 (10), 826–838.

Tanser, F., et al., 2006. Modelling and understanding primary health care accessibility and utilization in rural South Africa: an exploration using a geographical information system. Soc. Sci. Med. 63 (3), 691–705.

Transport & ICT Global Practice, 2016. Measuring Rural Access: Using New Technologies. Working paper. The World Bank Group, Washington DC. http://www.worldbank.org/transport (web archive link, 15 September 2021). http://www.worldbank.org/ict [15 September 2021].

United Nations Statistical Commission, 2016. Global Indicator Framework. https://unstats.un.org/sdgs/indicators/indicators-list/ [15 September 2021].

Van Rensburg, H.J., 2014. South Africa's protracted struggle for equal distribution and equitable access – still not there. Hum. Resour. Health 2014 (12), 26. https://doi.org/10.1186/1478-4491-12-26.