Accessibility in a post-apartheid city: Comparison of two approaches for the computation of accessibility indicators

J.W. Joubert[a], D. Ziemke[b], K. Nagel[a,b]

[a] *Centre of Transport Development, Industrial and Systems Engineering, University of Pretoria, Private Bag X20, Hatfield, 0028, South Africa*

[b] *Transport Systems Planning and Transport Telematics, Institute for Land and Sea Transport Systems, TU Berlin Sekr. SG12, Salzufer 17-19, 10587 Berlin, Germany*

Abstract

In many areas, spatial data to perform accessibility computations are hard to obtain or not available at all. Freely available volunteered geographic information (VGI) like from OpenStreetMap (OSM) may be a solution to this problem. While accessibility studies have been using data from OSM to perform network-based computations, this paper presents two approaches for accessibility assessment which exceed the use of OSM data for network creation. In the first approach, the transport network and parts of activity facilities are taken from OSM. Additionally, a synthetic population is created based on a census, enriched with travel demand from a travel survey. Based on local expert knowledge, a household-specific accessibility indicator is designed, which takes into account travel time, transport options, and facilities in direct vicinity. This approach respects the high diversity in residential locations and travel characteristics in South Africa. The second approach relies exclusively on OSM data to create network and activity facilities. It applies an econometric accessibility indicator, which calculates the accessibility of a given location as the weighted sum over the utilities of all opportunities reachable from that location including the costs of overcoming the distance. Since neither a synthetic population nor travel information is used, the approach is highly portable. It is found that the second approach, though being much more lightweight in terms of data requirements, yields the same quality of insights concerning accessibilities of different areas of the region. Both approaches detect areas where levels of accessibility deprivation are high and interventions in the transport-land-use system are advisable.

Keywords: Accessibility, Transport, Land Use, Open Data, Volunteered Geographic Inforamtion, OpenStreetMap, South Africa, Africa, Townships

1. Introduction

1.1. Accessibility indicators as planning tools

The improvement of accessibility is often stated as a central goal of proposed transport or infrastructure schemes (Geurs et al., 2012). Accessibility describes the *ease with which activities may be reached from a given location using a particular transportation system* (Morris et al., 1979; Litman, 2010; Knowles, 2009; Bocarejo and Oviedo, 2012; Ziemke, in preparation). Quantitative computations of accessibilities can be used as a comprehensive and efficient planning instrument. In contrast to other instruments whose focus is mostly based on travel alone (like measuring changes in travel times), the concept of accessibility focuses more strongly on the needs of the households (Litman, 2010). Accessibility computations are, therefore, seen as a potential alternative or supplement for/to traditional transport planning tools to assess land use and transport policies.

In many areas, however, spatial data to perform accessibility calculations are hard to obtain or not available at all. Freely available volunteered geographic information (VGI) like from OpenStreetMap (OSM, 2012), which is increasingly becoming a world-wide standard for geo-spatial data, offer a solution to this

Email addresses: johan.joubert@up.ac.za (J.W. Joubert), ziemke@vsp.tu-berlin.de (D. Ziemke), nagel@vsp.tu-berlin.de (K. Nagel)

problem. While transport and accessibility studies have been using data from OSM to create representations of the transport network and to perform network-based computations, this study uses OSM data more broadly. Two approaches for accessibility assessment for Nelson Mandela Bay Municipality in South Africa are presented. The approaches possess different levels of utilization of OSM data, which both exceed the use of OSM data for network creation.

In the first approach, the transport network as well as locations and types of activity facilities are taken from OSM. Additionally, a synthetic population is created based on a census. The corresponding travel demand is generated based on a travel survey. Local expert knowledge is applied to design a household-specific accessibility indicator that takes into account various characteristics of travel and land use, such as travel time to work and/or education, travel time to the nearest healthcare/shopping facility, availability of and walking time to different transport options, and availability of various facilities within walking distance. Weights are used to combine the respective values of aforementioned characteristics into a composite, household-based accessibility score. This approach appears particularly suitable in the South African context where housing locations and travel characteristics are highly diverse among the population.

The second approach relies exclusively on OSM data, which is – as before – used to create the network and activity facilities in the model. The approach applies an econometric accessibility indicator, which calculates the accessibility of a given measuring point as the weighted sum over the utilities of all opportunities including the costs of reaching them. In contrast to the first approach, no synthetic population is used for the calculation. The approach is highly portable since no input data other than those from OSM are used.

Besides data utilization, the two accessibility computation approaches also differ in terms of modeling philosophy: The first approach is based on a represenation of households and individuals of the study region (synthetic population) and their travel behavior, enriched with local expert knowledge, to perform accessibility compuations. The second approach, by contrast, operates on a regular spatial grid of measuring points and applies an econometric accessibility measure without relying on a population representation. In this paper, the insights obtainable from both approaches are presented and the particular strengths of the two approaches are discussed.

1.2. Accessibilities in South Africa

In general, low-income households are more likely to reside in areas that are less accessible and have less affordable transport, but are more affordable in terms of housing costs. Therefore, transport inaccessibility is generally a symptom of poverty rather than its cause (Gannon and Liu, 1997).

This general phenomenon is exacerbated in South Africa where differences in many aspects of social and economic life are bigger than in other countries. This is exemplified by the Gini index, which measures the extent to which the distribution of income within an economy deviates from a perfectly equal distribution. Its value has been higher in South Africa than in any other country of the world over the last decade (The World Bank, 2015). Huge differences also exist in the layout of residential and commercial areas in South Africa, which is a result of the Apartheid era when non-white people were relocated to settlements outside the cities. These spatial structures are still evident today and are quite inert towards change because of aforementioned transport-housing costs tradeoff. Since townships are oftentimes located particularly remote from the conurbation they pertain to, the effects of this general phenomenon are particularly pronounced in South Africa. People who reside in remote settlements have mostly a low income and cannot afford private vehicles for their daily commute. A significant share of people residing in these locations has to walk overly long distances (Venter and Behrens, 2005).

For the two approaches of accessibility assessments presented in this paper, the Nelson Mandela Bay Metropolitan Municipality, a smaller of the eight Metropolitan Municipalities in South Africa, located in the Southern part of the Eastern Cape province (see Figure 1), is used. It is located relatively distant from other conurbations, which facilitates the setup of the scenario. As depicted in Figure 2, it consists of the city of Port Elizabeth, the nearby towns of Uitenhage and Despatch, several townships, and surrounding rural areas.

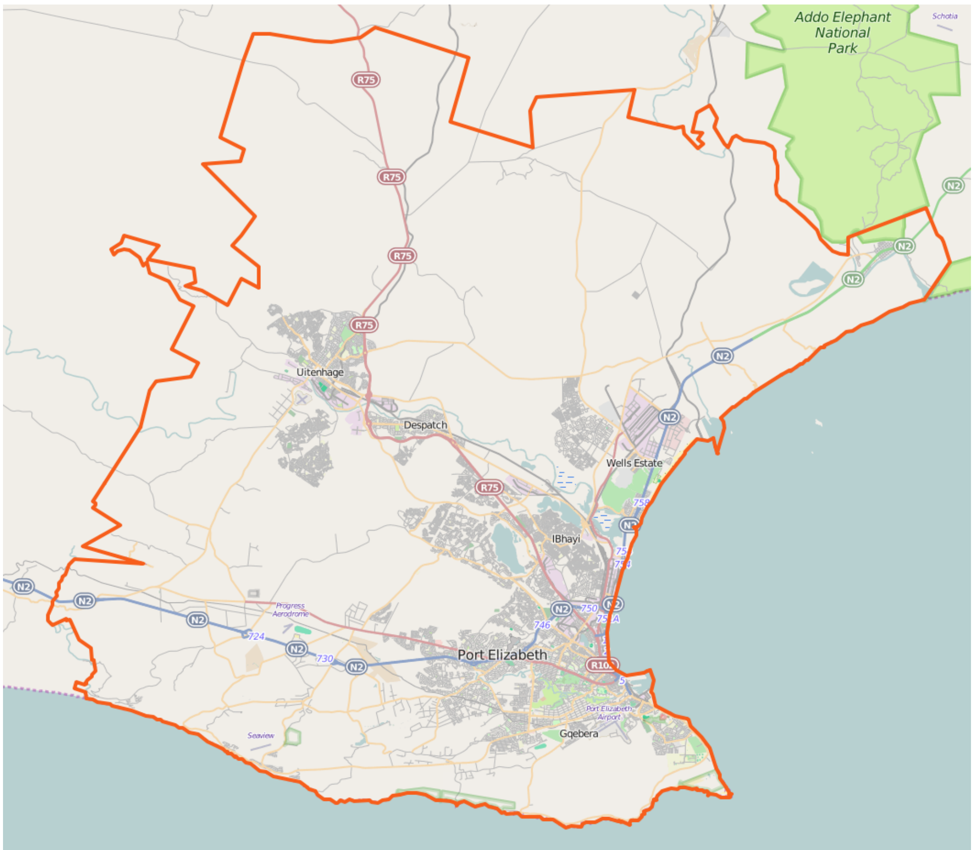

Figure 1: Nelson Mandela Bay Metropolitan Municipality with its boundaries (Source: OSM).

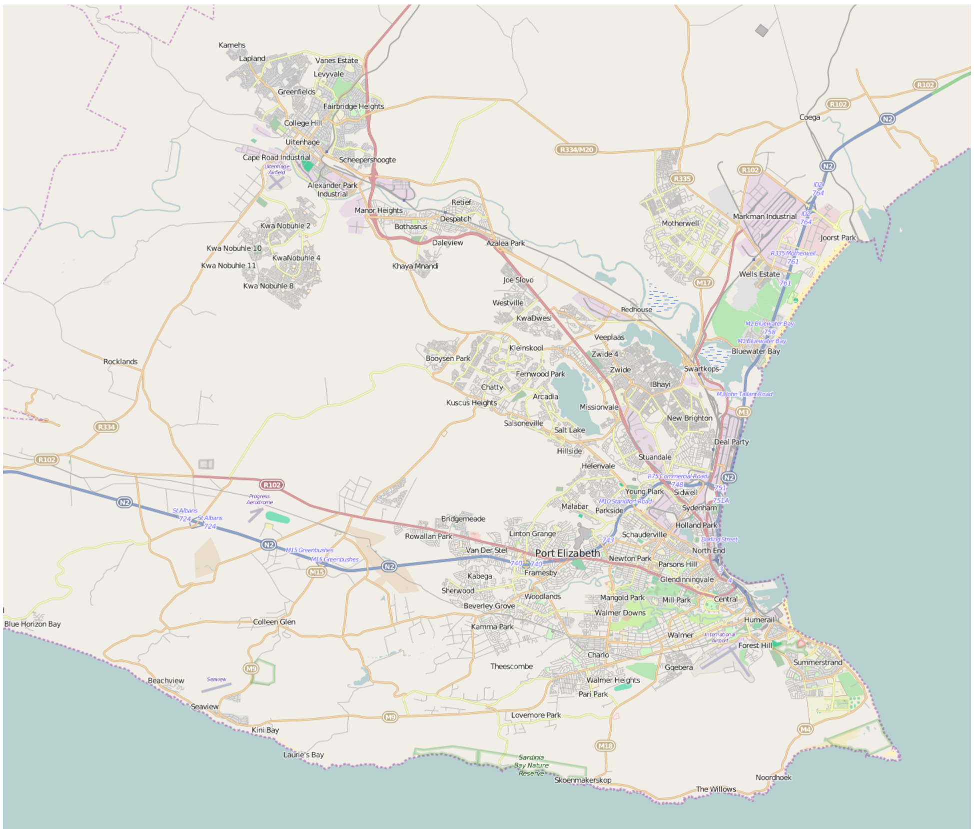

Figure 2: Southern portion of Nelson Mandela Bay Metropolitan Municipality (Source: OSM).

2. Household-based accessibility

2.1. Measure definition

As pointed out in Section 1.2, settlement patterns in South Africa and people's transport needs are highly diverse. According to Litman (2010), the evaluation of accessibility requires a detailed understanding of people's needs and abilities, travel mode constraints, and the quality of service at a destination. The goal of the household-based accessibility measure is to take these influence factors explicitly into account.

Litman (2010) summarizes the factors affecting accessibility as follows: Transport demand, mobility, transport options, security, information, integration, roadway design and management, and land use. It is obvious that this classification could be done differently, e.g. one could consider the integration aspect, which describes the ease to switch from one mode to another, as a part of the mobility aspect. The same could be done with the security aspect, which may arguably be understood as an additional factor affecting the actual quality of mobility. Indeed, all aforementioned characteristics can be collapsed and combined, which ultimately leads to the notion of accessibility as a measure that decribes the interdependencies between transport and land use in an explicit and interpretable fashion (see Section 1). Still, the more expanded classification given above is useful to develop the household-based accessibility measure. The detailed methodology of computing the measure is given in Section 2.3.

2.2. Data

The goal of the household-based accessibility indicator is to consider the diversity of the South African population to account for different needs. To do so, a synthetic population is created and stratified based on its members' demographics.

First, iterative proportional fitting (IPF), similar to the implementation by Müller and Axhausen (2011) is applied with the 2001 South African Census (Statistics South Africa, 2001) as source data. Based on the census, household size, population group, age, gender, household relationship, employment status, level of schooling, and household income are known for each member of the synthetic population.

Second, the Nelson Mandela Bay Metropolitan Travel Survey of 2004 (Nelson Mandela Bay Municipality, 2006), for which a 1% random sample of people had been queried in form of a 24-hour trip diary, is used. Each activity is described using a predetermined activity type, start and end time, as well as the mode connecting the different activities in the chain.

For each individual in the synthetic population, an activity chain sampled from the travel survey chains is drawn and assigned. The sampling process ensures based on a Hamming-distance procedure (Hamming, 1950) that the chain is from an observed individual with similar demographics. Then, a random 10% sample of the population is drawn, on which the household-based accessibility computations is based. The sample of 10% is considered large enough to be statistically representative of the entire population, while offering computing time savings.

2.3. Methodology

The calculation of the household-based accessibility indicators is done in a Java class, which is within the MATSim (Multi-Agent Transport Simulation) framework (Horni et al., in preparation) and which uses MATSim infrastructure, e.g. a MATSim router to find shortest paths. The official MATSim accessibility computation (see section 3.3) is, however, not used for computation of the household-based accessibility indicator.

It is resonable to run this computation within the MATSim framework because MATSim is an agent-based simulation framework where every individual is represented as an agent with an individual daily sequence of activities and the intervening trips (mostly referred to as a *plan*). So, each agent can be followed during all stages of the simulation process, which is essential to conduct the household-based accessibility computation.

Based on local expert knowledge, levels of acceptability and a scoring that translates the characteristics of the transport-land-use system into an person-specific accessibility indicator are defined. Inspired by the classification of Litman (2010), introduced in Section 2.1, the following four factors have been chosen to calculate the accessibility of each individual.

Mobility The mobility factor measures the travel time from the household location of a given member of the synthetic population

- to the location of the education facility used by the person,
- to the location of the workplace of the person,
- to the *nearest* healthcare facility, and
- to the *nearest* shopping facility.

For each of these four possible destinations, a mobility scoring is done according to Table 1. The overall score is calculated as the average of the individual scores, where a score of 2 corresponds to a high accessibility, 1 to a medium accessibility, and 0 to a low accessibility.

Table 1: Score for mobility.

Destination	**Score**		
	2	**1**	**0**
	Travel time (min)		
Work	≤ 30	31-90	> 90
Education	≤ 30	31-60	> 60
Healthcare	≤ 30	31-60	> 60
Shopping	≤ 15	16-30	> 30

To consider the important fact that not all facility types are relevant for a given person, the members of the synthetic population are classified into five groups:

- Children who attend school
- Employed adults who take children to school
- Employed adults who do not take children to school
- Unemployed adults who take children to school
- Unemployed adults who do not take children to school

Only relevant facilities are scored, where relevance is determined by observing the activities (outlined in Table 2) that are contained within the plan of a given agent. For example, a person with the symbols "e_3" and "w" in their activity chain is assigned to the *employed adults who take children to school* group.

Table 2: Activity symbols used in activity-travel plans.

Symbol	**Activity**
e_1	Going to school
e_2	Tertiary education
e_3	Dropping children at school
h	Home
l	Leisure
o	Other
s	Shopping
w	Work

For employed adults, the travel time from home to work is calculated by an A*-Euclidian algorithm, which finds the free-speed shortest path on the network. Dependent on the travel time based on this routing, a score is assigned according to Table 1. In case a person is marked as employed, but does not have a work activity in their plan, it is assumed that the person is a home-based worker with zero travel time.

For agents with an education activity in their plan, the travel time to the attended education facility is calculated. In case the activity contained within the agent's plan is *going to school* (either at primary or secondary school level, i.e. symbol "e_1"), it is assumed that the agent (the child) walks to school and the A*-Euclidian router is run with an assumed walk speed of 3km/h. Otherwise, the router that has been used for work destinations (see above) is applied, assuming that the agent travels approximately with car speed.

For healthcare and shopping, travel times based on walking speed (3km/h) are assumed. In contrast to work and education facilities, not the actually used facilities of a given synthetic traveler are taken into account, but the *nearest* facilities. This is done since it is assumed that, for a basic accessibility to services, the nearest healthcare and shopping facilities are sufficient. In contrast, work and education facilites cannot (or only in a limited and long-term fashion) be substituted.

Transport Options This factor quantifies the number of transport modes that are available to a household based on the information in the agent's plan. If a given mode of transport is available, quantitative points according to Table 3 are added.

Table 3: Points for available transportation options.

Transport Mode	Points
Car	5
Short-distance walking	4
Taxi	3
Bus	2
Train	1
Long-distance walking	0

The number of points that a person receives is summed. The total number of points is used to give a transport options score as shown in Table 4.

Table 4: Score for transportation options.

Score		
2	**1**	**0**
Points		
10-15	3-9	0-2

If an agent has a car trip in their plan, the car mode is assumed to be available. The availability of the mode of *short-distance walking* is considered available if the agent can reach their work or school facility by a walk of 20 minutes or less (if the agent is an employed adult or a child who attends school). Otherwise, the first shopping activity is considered. Here, it is not relevant whether the agent actually went to the considered activity by walking.

For the modes of train, bus, and taxi (i.e. minibus), it is checked whether the next station is reachable by a walk of 20 minutes or less, based on the A*-Euclidian algorithm with an assumed walk speed of 3km/h. Just as for short-distance walking, the assessment refers to options that are available from a geographic perspective and not a person's actual use of them. The value of 20 minutes is taken from the Household Travel Survey 2003 (Department of Transport, 2005), where a majority of people are able to reach a station within 20 minutes (Venter, 2011). As *long-distance walking* ($> 20min$) is always available, it does not contribute to the scoring.

Walking time to Transport This factor measures the walking time to reach the modes of transport that are contained in the daily plan of a given agent. For each activity in the plan, the walking time from the activity to the chosen mode is checked. Then, the average over these walking times is calculated and scored accoring to Table 5.

If an agent has a car trip in their plan, the car mode is assumed to be accessible at an access time of zero. For walking trips, the whole trip duration is considered. For trips by train, bus, or taxi, the time needed to walk to the nearest stop of the respective mode is used. Travel times for both walking trips and access to stops by walking are calculated by using the A*-Euclidian router with an assumed walk speed of 3km/h.

Table 5: Score for walking time to transport.

Score		
2	**1**	**0**
Walking time (min)		
≤ 15	15-30	> 30

Access to Facilities This factor assesses the availability of facilities near a person's home – the more facilities nearby, the more accessible the location of the person's home. If a facility is reachable within a 20 minutes' walk, it is considered accessible. The facilities (taken from OSM, see Section 2.2) include:

- Shop (food).
- Shop (other).
- Healthcare.
- Police station.
- Post office.
- Education.
- Petrol station.
- Bank including ATM.

Work places are not considered for this assessment, since nearby workplaces may not be useful to the observed person because of qualification mismatch. The scoring based on the number of facilities is given in Table 6.

Table 6: Score for access to facilities.

	Score		
	2	**1**	**0**
Number of Facilities	> 10	$6 - 10$	$0 - 5$

The four scores for mobility, transport options, walking time to transport, and access to facilities are each multiplied by 10 and summed, which yields the accessibility score of a given person. The score of a given household is calculated as the average of the scores of the individuals belonging to that household.

2.4. Results

Figure 3 depicts the spatial distribution of different levels of accessibility. It shows areas with distinctly different accessibility scores (red color = low accessibility; dark blue color = high accessibility). It can be seen that areas tend to be clustered in terms of their accessibility rating. The Port Elizabeth area, situated in the bottom right of the map, is blue, while green to red is dominant in many surrounding areas. Three areas, shown in the aerial photographs in Figure 4, are analyzed in more detail. Each of these photographs is taken at the same scale.

Figure 4a depicts an area situated in Port Elizabeth (a), which possesses a very high accessibility. Various types of facilities are located in this suburban area. The houses are big and many have swimming pools. In the northeastern part, there are shopping centers, restaurants, a primary school, a post office, a clinic, a petrol station, a bank, a library, a church, and sport facilities, which leads to a high score in the *access to facilities* category (see Section 2.3). The distance to the city center of Port Elizabeth is comparatively low so that many other facilities, in particular workplaces, are easily accessibile, which increases the *mobility* score.

Figure 4b depicts Kwa Nobuhle (b), a township situated near Uitenhage, which is less accessible than the area in Figure 4a. There are many houses in the area, but not many other facilities except primary and

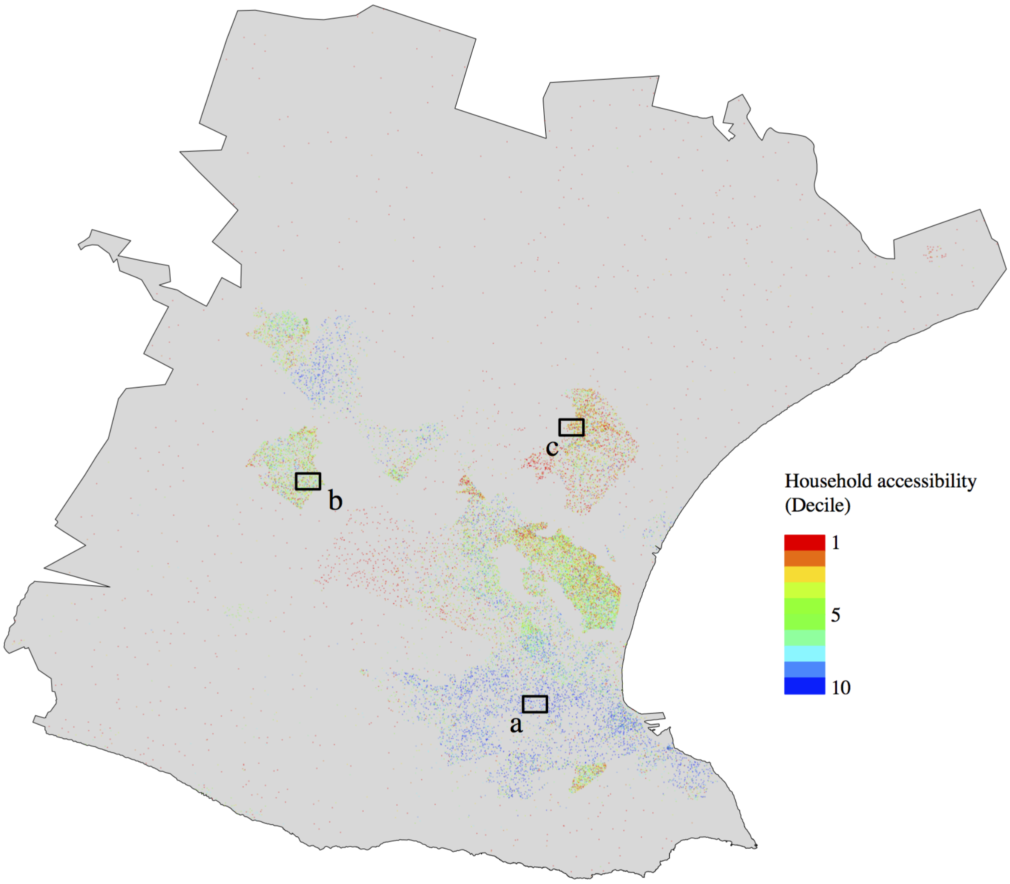

Figure 3: Accessibilities according to the household-based accessibility indicator.

secondary schools (the larger buildings). The *access to facilities* score of this region is low because not many facilities are accessible by walking. Since Uitenhage, where many workplaces are located, is near, the overall accessibility of this area ranges at a medium level.

Figure 4c shows Motherwell (c), a township situated north of Port Elizabeth. This area has a low accessibility because of a lack of facilities in that area. While there is an intermediary school in the southeastern corner, the rest of the area is taken up by housing. Next to the insufficient supply with facilities in the direct vicinity, Motherwell is also relatively far away from areas where such facilities are located.

In conclusion, the areas that are farther away from the city center and facilities are associated with lower accessibilities. Both in Figures 4b and 4c, it can be seen that houses are smaller and the density of them is high, which is often a property of areas with a high amount of informal dwellings and indicates that people with lower income live there. This clearly demonstrates that low accessibility is generally associated with low-income households and exemplifies the tradeoff between housing and transport costs (Lipman, 2006) as it has been discussed in Section 1.2.

(a) High accessibility

(b) Medium accessibility

(c) Low accessibility

Figure 4: Aerial photographs of areas with notably different accessibilities (Source: South African National Geospatial Institute (NGI).

3. Econometric accessibility

3.1. Measure definition

In Section 2, a household-based accessibility metric is introduced and applied. This can serve as the basis of household-based interventions. However, Figure 3 indicates that there are strong spatial structures of areas with deprived accessibilities. This may also be addressed by interventions into the infrastructure, either by improving transport or by adding facilities. In order to address such questions, an econometric accessibility measure may be used. Following Hansen (1959), accessibility A_i at location i to opportunities j is defined as

$$A_i = \ln \sum_j e^{-C_{ij}} \ , \tag{1}$$

where $-C_{ij}$ is the generalized cost of traveling from i to j. This is a high resolution accessibility measure as multiple opportunities j at the same location are all counted separately (Ziemke, in preparation; Nicolai and Nagel, 2014).

It is well known that this measure, up to an additive constant V, can be interpreted as the expected maximum utility at location i that can be extracted from opportunities j. For this, the true utility of using location j, U_j, is first decomposed into $U_j = V - C_{ij} + \epsilon_{ij}$, where V is a uniform intrinsic utility of all opportunities j, and ϵ_{ij} are person- and opportunity-specific corrections to V and C_{ij}. If the ϵ_{ij} are independently and identically Gumbel-distributed, a logit model and the logsum term (de Jong et al., 2007) in Equation 1 can be derived.

The behavioral interpretation is that the person, at i, weighs the utility of all opportunities j including the generalized cost $-C_{ij}$ of getting there, and picks the one that provides the best tradeoff between intrinsic utility and costs. The person does *not* always pick the opportunity with the smallest generalized cost since opportunities are expected to have intrinsic utility which is random from the perspective of the analyst.

In consequence, Equation 1 is a way to obtain the equivalent of the *mobility* score in Section 2.3 without having to make assumptions about the actually selected opportunity. With the econometric measure, similar information as in Section 2 can be derived without having to resort to the synthetic population with its activity chains.

3.2. Data

In contrast to the household-based accessibility indicator introduced in Section 2, the econometric accessibility measure is neither based on a synthetic population nor on a travel diary. Only a transport network and facility information, both OSM-based, are used, which offers two advantages. Firstly, no data except those from OpenStreetMap are used. This is a highly beneficial property as the lacking availability of data is an often-encountered and well-known issue in many (transport) modeling procedures (Ziemke et al., 2015). Secondly, the econometric accessibility indicator has an economic interpretation and can, therefore, be applied for project appraisal.

3.3. Methodology

To calculate the values of the econometric accessibility indicator, the standard MATSim accessibility computation (Ziemke, in preparation) is used. It can either be used on a zone-based or coordinate-based level. The latter is the more common case since MATSim itself is coordinate-based. Details concerning the interpretation of coordinate- and zone-based accessibility measures and a discussion of issues associated with zone-based calculations are given by Nicolai and Nagel (2014).

Since spatial data are mostly only available at some level of administrative zones, e.g. municipalities, it is mostly not feasable to conduct zone-based approaches of accessibility assessment at a high level of spatial resolution. The coordinate-based MATSim accessibility calculation can be conducted on the basis of VGI like OSM, which contains data on activity facilities on a coordinate-based level[1].

[1] For an example of how to apply the coordinate-based MATSim accessibility computation see `http://matsim.org/javadoc` → accessibility → `RunAccessibilityExample`.

To calculate the accessibility A_i of a given origin location i to opportunity locations j, both the origin location i and the opportunity locations j are assigned to the road network. For every i, a so-called *least cost path tree* computation (Lefebvre and Balmer, 2007) is carried out, which determines the best route and, therby, the least cost $-C_{ij}$ between origin i and opportunity j based on Dijkstra's algorithm (Dijkstra, 1959). Once the least cost path tree has explored all nodes, the resulting costs $-C_{ij}$ for all opportunities j are queried and the accessibility is calculated as stated in Equation 1. The computation can also be done based on a congested network with time-dependent travel times, which can e.g. come from a MATSim transport simulation. This property (not used in the current application) illustrates why an integration of accessibility computation with transport simulation is reasonable.

The costs of traveling on the network are supplemented by the costs of accessing the network from the origin i (network access) and the costs of accessing the destination j from the network (network egress). For origin locations i, the shortest distance to the network is either given by the Euclidean distance to the nearest node or the orthogonal distance to the nearest link on the network. For destination locations j, the Euclidean distance to the nearest node is used to determine the shortest distance to the network. The assumption that opportunity locations are attached to the nearest network *node* rather than the nearest network *element* is minor in terms of changes in accessibility results, but major in terms of computational optimization. It significantly reduces the exploration of the network by the least cost path tree, which is the computationally most expensive task of the accessibility computation.

The econometric accessibility measure is calculated for shopping, education, leisure, and other facilities (including healthcare facilities) separately. For each tile of a spatial grid, the accessibility values for these four types of facilities are averaged, which leads to a composite econometric accessibility measure that can be compared to the composite household-based accessibility from Section 2.

3.4. Results

The results of the econometric accessibility computation are shown in Figure 5, with a color scheme similar to Figure 3. Since the household-based measure is based on a synthetic population and, thus, the locations where people actually reside, accessibility values are only shown for these locations in Figure 3. The calculation of the econometric accessibility works differently: Accessibility values are computed for *each* tile of the spatial grid (see Section 3.3), but only those tiles that cover areas with a certain minimum population are depicted in Figure 5.

Again, it can be seen that the Port Elizabeth area (a) has the highest accessibility. Just as for the household-based accessibility measure, the two townships of Kwa Nobuhle and Motherwell have the lowest accessibility scores of all larger populated areas, as visible in red and yellow shades of color in Figure 5. Contrary to the household-based accessibility in Figure 3, where Kwa Nobuhle (b) had better accessibilities than Motherwell (c), the two townships show more similar accessibilities according to the econometric accessibility measure.

4. Discussion

As shown in Sections 2.4 and 3.4, the two accessibility measures, though being quite different in the way they are defined and computed, yield similar results. This is notable because the econometric accessibility computation has much lower input data requirements, fewer assumptions have to be made, and no local expert knowledge is needed to define levels of acceptability and scoring.

Both accessibility calculation procedures clearly detect the significantly lower accessibilities of the two townships of Kwa Nobuhle and Motherwell. The results based on both measures are thus in line with general wisdom that remote areas lack accessibility, whilst central and suburban areas have higher accessibilities. It is interesting to note that having low accessibility is *not* equivalent with being a township area: The township of iBhayi, marked by “d” in Figure 5, has considerably higher accessibilities than the other two townships.

Thus, both accessibility indicators detect areas where levels of accessibility deprivation are high and interventions in the transport-land-use system may be advisable. These could consist in improving transport infrastructure and/or adding facilities in deprived neighborhoods.

In contrast to the econometric accessibility indicator, the household-based accessibility indicator is also capable of assessing impacts of household-based interventions like increasing car availability of households.

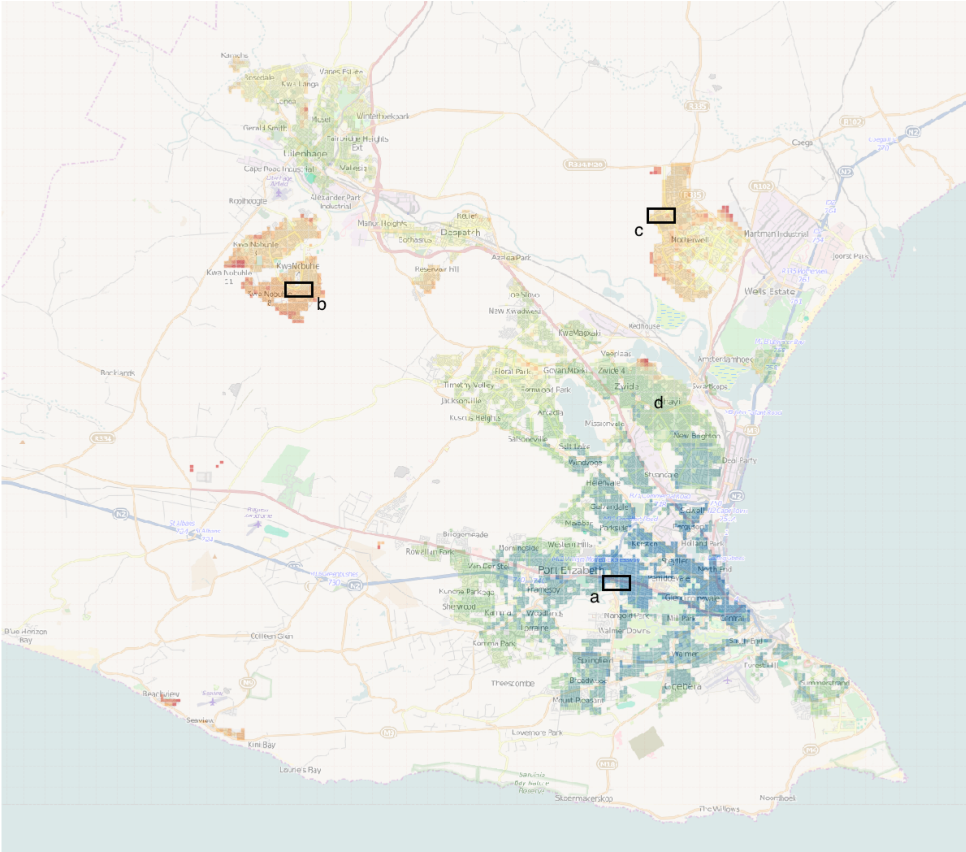

Figure 5: Accessibilities according to econometric accessibility indicator.

On the other hand, the econometric indicator does not require the analyst to consider which opportunitiy locations have actually been visited by an agent. Neither does the analyst have to make moral judgements. For instance, the travel survey, which the household-based indicator is based on, contains a significant number of working children although the children reported an age at which they are not allowed to work. These cases where not included into the synthetic population, which constitutes a moral judgement that one could argue.

As pointed out in Section 3.4, the results are somewhat different when comparing the accessibilities of the two townships of Kwa Nobuhle and Motherwell against each other. While Kwa Nobuhle has better accessibilities according to the household-based measure, Motherwell is somewhat better accessible based on the econometric measure. This is likely due to the fact that workplaces are contained in the household-based accessibility measure (see Section 2.3), while they are not (explicitly) considered by the econometric measure owing to the fact that extracting data concerning leisure or shopping facilities is simpler than extracting workplaces from OSM. A part of the work facilities is still implictly contained in the econometric accessibility measure because each leisure facility, hospital, or school constitutes a workplace, too. Workplaces like offices, however, are not taken into account by the computation. Kwa Nobuhle is comparatively close to Uitenhage with its factories, which helps reaching a better accessibility according to the household-based measure. First approaches to also extract work facilities from OSM and include them into the accessibility computation show promising results.

5. Conclusion

In this study, two accessibility measures were introduced and applied. The two measures, one labeled as *household-based* and the other as *econometric* are different in terms of data utilization and modeling philosophy. The houshold-based measure uses various sorts of data (OSM, a census, a travel survey, different sorts of facility information) and calculates accessibility by following people along their daily activity chain. Travel times, availability of transport modes, walking times to transport, and local access to facilities are each evaluated and then combined into a composite accessibility measure using local expert knowledge.

In contrast, the econometric accessibility measure leaves the consideration of individuals aside. Instead, accessibility computations are performed on a regular spatial grid taking into account the variety of activity opportunities reachable from each of these grid points (measuring points) along with the cost of overcoming the distance to each of these activity opportunities. Owing to the mathematical formulation of the measure, it is interpretable in economic terms as the maximum expected utility that a person at a given location extracts from facilities at other locations. In contrast to the household-based measure, volunteered geographic information (VGI) are used as the exclusive data source to calculate the measure. Because such VGI are available for many locations of the world based on the same standard (OSM), the econometric accessibility assessment procedure is highly portable, which is an important feature in infrastructure modeling (Ziemke et al., 2015).

While the household-based measure is capable of assessing impacts of household-based interventions, both accessibility indicators are suitable to evaluate improvents both in the transport and the land-use system and assess its impacts, also in terms of localizing communities affected by projects. Consequently, the paper is a contribution to accessibility analyses based on easily and freely obtainable OSM data and a highly portable and econometrically sound accessibility evaluation framework.

Acknowledgements

JWJ is grateful to the South African National Treasury for sponsoring the research leading to the household-based accessibility measure, and also to Ms Jeanette de Hoog for her inputs during her final year Industrial Engineering degree project in the initial stages of the research. KN thanks the University of Pretoria for hospitality during a research semester. All authors thank ERAfrica for funding.

Bibliography

Bocarejo, J. and D. Oviedo. Transport accessibility and social inequities: a tool for identification of mobility needs and evaluation of transport investments. *Journal of Transport Geography*, 24:142–154, 2012.

de Jong, G., A. Daly, M. Pieters, and T. van der Hoorn. The logsum as an evaluation measure: Review of the literature and new results. *Transportation Research Part A: Policy and Practice*, 41(9):874–889, 2007. ISSN 0965-8564. doi: 10.1016/j.tra.2006.10.002.

Department of Transport. Key Results of the National Household Travel Survey, August 2005.

Dijkstra, E. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1:269–271, 1959.

Gannon, C. and Z. Liu. Poverty and Transport. In *TWU Papers TWU-30*, Washington, DC., 1997. World Bank.

Geurs, K., K. Krizek, and A. Reggiani. *Accessibility Analysis and Transport Planning*, chapter Accessibility analysis and transport planning: an introduction, pages 1–12. Edward Elgar, 2012.

Hamming, R. Error detecting and error correcting codes. *Bell System Technical Journal*, 29(2):147–160, 1950.

Hansen, W. How Accessibility Shapes Land Use. *Journal of the American Planning Association*, 25(2): 73–76, 1959.

Horni, A., K. W. Axhausen, and K. Nagel, editors. *The Multi-Agent Transport Simulation MATSim*. in preparation. URL http://ci.matsim.org:8080/view/All/job/MATSim-Book/ws/main.pdf.

Knowles, R. Transport Geography. *International Encyclopedia of Human Geography*, 1:441–451, 2009.

Lefebvre, N. and M. Balmer. Fast shortest path computation in time-dependent traffic networks. In *Proceedings of the Swiss Transport Research Conference (STRC)*, Monte Verita, Switzerland, 2007. URL http://www.strc.ch.

Lipman, B. A Heavy Load: The Combined Housing and Transportation Burdens of Working Families. Technical report, Center for Housing Policy, October 2006.

Litman, T. Evaluating Accessibility for Transportation Planning. Technical report, Victoria Transport Policy Institute, March 2010.

Morris, J., P. Dumble, and M. Wigan. Accessibility indicators for transport planning. *Transportation Research Part A*, 13A:91–109, 1979.

Müller, K. and K. Axhausen. Hierarchical IPF: Generating a synthetic population for Switzerland. In *45th Congress of the European Regional Science Association (ERSA)*, 2011.

Nelson Mandela Bay Municipality. Public Transport Plan, 2006.

Nicolai, T. W. and K. Nagel. High resolution accessibility computations. In Condeço, A., A. Reggiani, and J. Gutiérrez, editors, *Accessibility and spatial interaction*, pages 62–91. Edward Elgar, 2014.

OpenStreetMap. http://www.openstreetmap.org/, 2012. Accessed 07 April 2012.

Statistics South Africa. Census. http://www.statssa.gov.zahttp://www.statssa.gov.za/?page_id=3892, 2001. URL Accessed10July2015.

The World Bank. Gini Index. http://data.worldbank.org/indicator/SI.POV.GINI, 2015. Accessed 15 July 2015.

Venter, C. Transport expenditure and affordability: The cost of being mobile. *Development Southern Africa*, 28:1:121–140, March 2011.

Venter, C. and R. Behrens. Transport Expenditure: Is the 10% policy benchmark appropriate? In *Proceedings of the 24th Southern African Transport Conference*, Pretoria, South Africa, 2005.

Ziemke, D. Accessibility. In Horni et al. (in preparation), chapter 18. URL http://ci.matsim.org:8080/view/All/job/MATSim-Book/ws/main.pdf.

Ziemke, D., K. Nagel, and C. Bhat. Integrating CEMDAP and MATSim to increase the transferability of transport demand models. Annual Meeting Preprint 15-5516, Transportation Research Board, Washington D.C., 2015. Also VSP WP 14-15, see http://www.vsp.tu-berlin.de/publications.