Dominik Ziemke, Johan W. Joubert, Kai Nagel

Accessibility in a Post-Apartheid City: Comparison of Two Approaches for Accessibility Computations

WISSEN IM ZENTRUM
UNIVERSITÄTSBIBLIOTHEK

Technische
Universität
Berlin

Accessibility in a post-Apartheid city

Comparison of two approaches for accessibility computations

**Dominik Ziemke · Johan W. Joubert ·
Kai Nagel**

Abstract Many authors argue that issues related to interpretability, lack of data availability, and limited applicability in terms of policy analysis have hindered a more widespread use of accessibility indicators. Aiming to address these aspects, this paper presents two accessibility computation approaches applied to Nelson Mandela Bay in South Africa. The first approach, a *household-based accessibility indicator*, is designed to account for the high diversity both among the South African society and in terms of settlement patterns. Besides OpenStreetMap as its main data source, this indicator uses a census and a travel survey to create a synthetic population of the study area. Accessibilities are computed based on people's daily activity chains. The second approach, an *econometric accessibility indicator*, relies exclusively on OSM and computes 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. Neither a synthetic population nor travel information is used. It is found that the econometric indicator, although associated with much lower input data requirements, yields the same quality of insights regarding the identification of areas with low levels of accessibility. It also possesses advantages in terms of interpretability and policy sensitivity. In particular, its exclusive reliance on standardized and freely available input data and its easy

Dominik Ziemke · Kai Nagel
Transport Systems Planning and Transport Telematics, Institute for Land and Sea Transport Systems, TU Berlin, Sekr. SG12, Salzufer 17-19, 10587 Berlin, Germany
Centre of Transport Development, Industrial and Systems Engineering, University of Pretoria, Private Bag X20, Hatfield, 0028, South Africa
Tel.: +49-30-31421383
Fax: +49-30-31426269
E-mail: ziemke@vsp.tu-berlin.de

Johan W. Joubert
Centre of Transport Development, Industrial and Systems Engineering, University of Pretoria, Private Bag X20, Hatfield, 0028, South Africa

portability are a novelty that can support the more widespread application of accessibility measures.

Keywords Accessibility · Transport · Land Use · Open Data · Volunteered Geographic Information · OpenStreetMap · South Africa · Townships

1 Introduction

Accessibility describes the *ease with which activities may be reached from a given location using a particular transport system* (Morris et al, 1979; Litman, 2010; Knowles, 2009; Bocarejo and Oviedo, 2012; Chen et al, 2007; Venter and Cross, 2014; Ziemke, 2016). It is, therefore, also referred to as *the potential for interaction* (Hansen, 1959; Reggiani and Martín, 2011). As such, accessibility is what ultimately makes a location more attractive than others (El-Geneidy and Levinson, 2011). Accordingly, the improvement of accessibility is often stated as a main goal of proposed land use and transport policies (Handy and Niemeier, 1997; Vandenbulcke et al, 2009; El-Geneidy and Levinson, 2011; Geurs et al, 2012).

1.1 Accessibility indicators as planning tools

Typical infrastructure assessment instruments are mostly based on travel alone (like measuring and monetizing changes in travel times, highway levels of service, or delays) (Handy and Niemeier, 1997; El-Geneidy and Levinson, 2011). As El-Geneidy and Levinson (2011) point out, such measures have limited utility as they only consider the ability of residents to transport themselves under certain conditions, while they neglect the actual purpose of travel, i.e. the ease of reaching valued destinations. This is what the concept of accessibility, which focuses more strongly on the needs of individuals and households (Litman, 2010), is intended to capture. As such, quantitative computations of accessibilities can be used as a comprehensive and efficient planning instrument and are seen as a potential alternative or supplement to traditional planning tools (Koenig, 1980; Handy and Niemeier, 1997; Vandenbulcke et al, 2009; Venter and Cross, 2014).

Accessibility is determined both by the patterns of land use and by the characteristics of the transport system (Reggiani et al, 2011). Hence, measures of accessibility usually consist of two components, a *land-use (or activity) component* and a *transport component* (Koenig, 1980): The land-use component reflects the spatial distribution of opportunities and is characterized by both the amount and the location of different types of activities: The more opportunities and the greater the variety, the greater the accessibility. The transport component reflects the ease of travel between locations, determined by the quality of service provided by the transport system: The less time and money spent in travel, and thus the more places that can be reached within a certain travel budget, the greater the accessibility (Handy and Niemeier, 1997).

Handy and Niemeier (1997) distinguish three types of accessibility measures: (1) *Isochrone-based (or cummulative opportunities) measures* count the number of opportunities reachable within a given travel time or distance cut-off (isochrone). Because all reachable destinations within that isochrone are weighted equally, this measure emphasizes the number of potential destinations.

(2) *Gravity-based measures*, related to the well-kwown gravity model for trip distribution, weight opportunities by travel impedance, which is usually a function of travel time or travel cost. The negative exponential function is most commonly used. The closer an opportunity is, the more it contributes to accessibility. Next to many others, the approach by Hansen (1959), which is generally regarded as the start of quantitative accessibility computations (Reggiani and Martín, 2011), uses a gravity-based measure (El-Geneidy and Levinson, 2011). Note that the isochrone-based measure is a special case of the gravity-based measure, with the impedance function equal to one if within the isochrone, and zero otherwise.

(3) Finally, *utility-based measures*, based on random utility theory, assume that individuals assign a utility to each possible destination and, based on this, select the alternative which maximizes their utility. In Sec. 3.1, it is pointed out in more detail why the expected maximum utility, the logsum term, can be interpreted as a measure of accessibility. The utility function reflects both the attractiveness of the potential destination and the travel impedance that must be overcome to reach that destination. Handy and Niemeier (1997) highlight that this measure, while sometimes similar in form to gravity-based measures, has theoretical advantages.

In contrast to isochrone-based measures, which only count the number of opportunities that are located within a defined boundary, gravity-based and utility-based measures, in principle, sum *all* available opportunities. Therefore, the analyst does not have to specify a cutoff distance or travel time (Venter and Cross, 2014). Östh (2011, p.585) describes and illustrates issues, which can arise from using measures that apply a predefined catchment area.

Next to questions of interpretability und usability (Handy and Niemeier, 1997; Naudé et al, 1999; Venter and Cross, 2014), a major obstacles to a more widespread use of accessibility measures in policy analysis consist in lack of data availability (Pozzi et al, 2010; Venter and Cross, 2014). Freely available volunteered geographic information (VGI) like information from OSM (OpenStreetMap, 2016), which is increasingly becoming a worldwide standard for geospatial data, offer a solution. While many 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 on a broader basis. Two approaches for accessibility assessment for Nelson Mandela Bay 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 (Statistics South Africa, 2001). The corresponding travel demand is generated based on a travel survey (Nelson Mandela Bay Municipality, 2006). Based on local expert knowledge, a *household-based accessibility indicator* is designed, which takes into account various characteristics of land use and travel, such as (1) travel time to work and/or education and travel time to the nearest healthcare/shopping facility, (2) availability of different transport options, (3) walking time to transport options, and (4) presence of facilities within walking distance. Weights are used to combine the respective values of these characteristics into a composite, household-based accessibility score. The household-based indicator can be regarded a *person-based accessibility measure*. Some authors have argued that these types of measures are more sensitive to individual activity patterns and accessibility in space and time than its counterpart, *place-based accessibility measures* (Reggiani and Martín, 2011). Accordingly, such an approach appears particularly suitable in South Africa with its high diversity in society and residential patterns.

The second approach relies *exclusively* on OSM data, which are used to create the network and activity facilities. The approach applies an *econometric accessibility indicator*, which calculates accessibilities on a set of given measuring points (e.g. a regular grid) as the weighted sum over the utilities of *all* opportunities including the costs of reaching them. It can, therefore, be regarded a *place-based accessibility measure* (Reggiani and Martín, 2011). In contrast to the first approach, no synthetic population is used for the calculation. No assumptions about levels of acceptability have to be made. The approach is highly portable since no input data other than those from OSM are used. Furthermore, the approach can be easily applied for policy analysis.

Besides data utilization, the two accessibility computation approaches also differ in terms of modeling philosophy: The first approach is based on a representation of households and individuals of the study region and their travel behavior (person-based measure), enriched with local expert knowledge, to perform accessibility computations. In particular because of its reliance on a travel survey, the measure is largely based on *stated behavior*. The second approach, by contrast, operates on a regular spatial grid of measuring points (place-based measure) and applies an econometric accessibility measure without relying on a population representation. In contrast to the household-based measure, it focuses on the *potential* to participate in activities that are available to individuals. In this paper, the insights obtainable from both approaches are presented and the particular strengths and weaknesses of the two approaches are discussed.

As study area, Nelson Mandela Bay, a smaller of the eight metropolitan municipalities of South Africa, located in the Southern part of the Eastern Cape province, is chosen. It is located relatively distant from other conurbations, which facilitates the delineation of the scenario. As depicted in Fig. 1, it consists of the city of Port Elizabeth, the nearby towns of Uitenhage and Despatch, several townships, and surrounding rural areas.
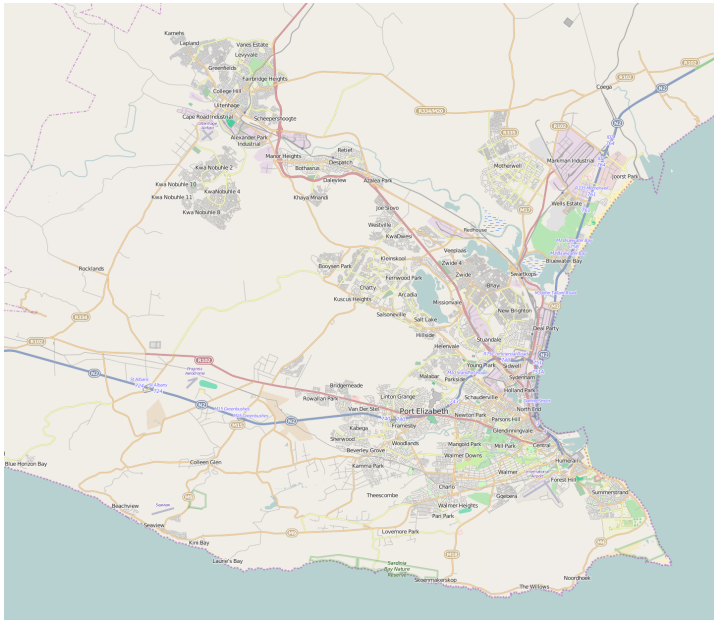
Fig. 1: Southern portion of Nelson Mandela Bay Metropolitan Municipality (Source: OSM)

1.2 Accessibilities in South Africa

There is a broad general consensus that accessibility is one of the major determinants of growth and development, both regarding economic and social dimensions. Naudé et al (1999) assert that a lack of accessibility to market centers is one of the major constraints of economic development in South Africa. Pozzi et al (2010) find that remote or inaccessible areas are likely to have high concentrations of chronic poverty. Venter and Cross (2014) point out that accessibility is central to promote sustainable livelihood and that deficits thereof may lead to social exclusion and the denial of basic human rights. In particular in the context of development economics, such relationships have been studied with a focus on African countries (e.g., Naudé and Krugell, 2004; Christiaensen et al, 2003), where different explanatory variables like descriptions of geographic location, market connectedness, remoteness, and distances to certain points of interest are applied. While many studies do not consider accessibilities explicitly or perform related computations, it is obvious that aforementioned variables are related to the concept of accessibility. Despite the amount of literature on regional development in Africa and the widely acknowledged relationship between accessibility and economic growth, there are few studies that directly focus on accessibilities and its quantitative analysis (Pozzi et al, 2010; Cheruiyot and Harrison, 2014).

This is noteworthy since the relationship of growth and accessibility is of particular interest in post-Apartheid South Africa. Apartheid laws like the 'Group Areas Act' of 1950 relocated the majority of black people to places where levels of accessibility were low, including ethnic homelands in the spatial peripheries of the country and townships in the spatial peripheries of cities (Cheruiyot and Harrison, 2014). As such, the natural growth and development of South Africa's cities and towns was artificially curtailed (Naudé and Krugell, 2004). The resulting spatial structures continue to exist today and are highly inert towards change. Once a low-income household resides in an inaccessible location, options to change that location are limited. As such, the general tendency that low-income households reside in areas that are less accessible, but more affordable in terms of housing costs, which can be observed in most countries, is reinforced in South Africa as townships are often located particularly remote from conurbations. Because of remote housing locations and low incomes, which exclude many transport options, a significant share of people residing in these locations travels very long distances by walking (Venter and Behrens, 2005). Consequently, South Africa has been characterized by significant inequality in spatial economic activity (Naudé and Krugell, 2004; Cheruiyot and Harrison, 2014). The Gini index, which measures the extent to which the distribution of income within an economy deviates from a perfectly equal distribution, has been higher in South Africa than in any other country of the world over the last decade (The World Bank, 2015).

The following studies with explicit computations of accessibilities in the South African context could be identified: Naudé et al (1999) measure the accessibility of rural centers in a remote region of the Wild Coast in South Africa, where residents from 188 out of 900 villages have to travel more than one hour to reach the nearest market center. Analysis is done with a GIS-based tool that applies an isochrone-based procedure (cf. Sec. 1.1). The authors identify inaccessible villages (with travel times to the nearest market center exceeding one hour) and analyze the effects of both a land-use and a transport policy. They conclude that their procedure provides a useful indication of the scale of possible accessibility improvements. Although the authors highlight the importance of such approaches to be easy to use and, especially in the context of developing countries, to be affordable for users, they mention that their tool is only partly open and not fully free of charge.

Tanser et al (2006) model the accessibility of primary health care (PHC) facilities in the rural district of Hlabisa in KwaZulu-Natal, South Africa, intending to identify vulnerable populations with limited access. They compute both walk times and public-transport-based travel times to the nearest PHC facility. Based on log-likelihood maximization, they fit their model against data obtained from interviews with 23,000 homesteads and obtain a model that predicts the PHC attended by a given homestead (cf. person-based measures in Sec. 1.1) with an accuracy of 91%. They find that a homestead, which can reach a PHC facility within 30 minutes, is ten times as likely to use this facility as a homestead that has to travel 90 to 120 minutes. While providing very insightful results for the specific analysis under consideration, the

model's reliance on a high number of interviews effects that the model is not transferable to other thematic and spatial contexts.

In a study conducted in Uganda, Pozzi et al (2010) explore the relationships between poverty and market accessibility in Uganda. Next to beeline distances to roads and markets, travel time to roads and markets, and population density, they use a 'travel-time-weighted population', which belongs to the class of gravity-based approaches (cf. Sec. 1.1), with the number of people in other locations as a proxy for opportunities. Only a weak correlation is found between poverty and distance and travel time to roads, while poverty is more strongly correlated with distance and travel time to markets and population density. Also, the authors find a strong correlation of the poverty indicators with their accessibility index, which reaffirms the necessity to take into account reachable opportunities apart from pure transport-system-related characteristics (cf. Sec. 1.1). The authors do, however, report difficulties related incompleteness of the required input data, in particular those required to determine the existence and location of markets, which are taken from a geospatial research database.

Cheruiyot and Harrison (2014) examine the relationship between economic activity, measured as gross value added (GVA), and accessibility, measured as travel times to the four major metropolitian areas in South Africa (Gauteng (Johannesburg and Pretoria), Cape Town, eThekwini (Durban), and Nelson Mandela Bay), for South African municipalities. The authors report that they cannot prove the proposed relationship between accessibility measure and GVA on a nationwide scale and speculate that their model may be unable to deal with a multi-polar spatial landscape as it is the case when considering South Africa as a whole. This seems plausible, since in their model a higher distance only translates into a worse accessibility score, while it does not take into account the fact that a remote destination is likely of less importance and should, therefore, only have a reduced impact on the score. As will be shown in Sec. 2 and Sec. 3, the two models of the present paper do account for this effect. On a regional level, with less complex spatial structures (e.g. the metropolitan areas of Cape Town and the Eastern Cape (Nelson Mandela Bay)), they find the hypothesized correlation between GVA and accessibility.

Venter and Cross (2014) describe 'access envelopes' as a new accessibility measure, which they regard suited to assessing the impact of transport and spatial development strategies. The computation result is referred to as *net wage after commute* (NWAC) and represents the potential income obtainable at a specific job location minus the time and money required to reach that location. The authors apply the method to a number of case studies, e.g. the extension of the bus rapid transit (BRT) system in the City of Tshwane (Pretoria) and conclude that their method is capable of presenting complex spatial relationships in an intuitive manner suited for land use and transport analyses. As input data, their method requires information on jobs, potential wage levels, public transport coverage, walking times, and public transport costs. The authors see this as a limitation to their approach, in particular as accurate spatial data on public transport routes and fares are seldom available. In their computation for Tshwane, potential incomes in potential destinations

Table 1: Review of Accessibility Studies in South Africa

Author(s)	Aim	Spatial scope	Indicator type	Impedance measure	Data and other limitations	Policy sensitivity	Transfer-ability
Naudé et al (1999)	Measure accessibilities to rural centers and analyze policies	Eastern Cape province in South Africa	Isochrone-based	Network-based travel times	Tool only partly open and not free of charge	Used to analyze a land-use and a transport policy	No
Tanser et al (2006)	Analyze accessibilities of primary healthcare facilities	Hlabisa district in KwaZulu-Natal, South Africa	Isochrone-based	Theoretical walking and PT travel time based on costgrow algorithm	Reliant on a high number of interviews; therefore not transferable	No	No
Pozzi et al (2010)	Explore relationship between poverty and accessibility	Uganda	'Travel-time-weighted population' (gravity-based)	Network-based travel time	Difficulties with data incompleteness	No	No
Cheruiyot and Harrison (2014)	Examine relationship between economic activity and accessibility	South Africa (national scope)	Gravity-based (simple)	Network-based travel times	Problems with multipolar spatial landscape	No	No
Venter and Cross (2014)	Analyze accessibility improvements because of different policies	City of Tshwane (Pretoria), South Africa	'Access envelopes'	Network-based and schedule-based PT travel times	Very high input data burdens	Used to analyze transport policies	No

were obtained from the Gauteng transport models updated job location data, which will likely limit the transferability of the approach to other regions.

In summary, out of the five discussed studies, four are focused on South Africa, one another African country (cf. Table 1). In terms of methodology (cf. Sec. 1.1), two studies (Cheruiyot and Harrison, 2014; Pozzi et al, 2010) apply gravity-based measures, two isochrone-based measures (Naudé et al, 1999; Tanser et al, 2006), and one a novel measure ('access envelope' Venter and Cross, 2014). At least three of these studies report rather high input data requirements and/or difficulties with inconsistency or incompleteness of required input data (Venter and Cross, 2014; Pozzi et al, 2010; Tanser et al, 2006). Cheruiyot and Harrison (2014) also report conceptual difficulties in that their measure may not be suitbale for a spatially complex study area. With its focus on transport (and potentially land use) policies at a metropolitan level, the study by Venter and Cross (2014) is most similar to the studies presented in this paper. In particular when compared to the econometric accessibility indicator (cf. Sec. 3), however, the input data requirements of the study by Venter and Cross are significantly higher. Also, similar to the other presented studies, it must be asserted that their approach, mainly due to context-specific input data, is not transferable to other spatial contexts. Such limitations are often regarded as hindering the more widespread application of otherwise favorable accessibility approaches. The present study is intended to address such limitations.

2 Household-based accessibility

As pointed out in Sec. 1.2, settlement patterns and people's needs and constraints are highly diverse in South Africa. In their review of existing accessibility approaches, Geurs and van Wee (2004) point out that there are, next to the land-use and transport components of accessibility (cf. Sec. 1.1), also a temporal and an individual component of accessibility. The latter addresses the fact that different socio-economic groups may have highly different needs and constraints. Especially in a highly diverse society like South Africa, it seems plausible to pay particular attention to these aspects. In line with that, Litman (2010) argues that 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. Consequently, the goal of the household-based accessibility measure is to take these factors explicitly into account.

2.1 Measure definition

According to Litman (2010), the factors that affect accessibility can be summarized 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 describes the interdependencies between land use and transport (cf. Sec. 1.1). Still, the more expanded classification given above is useful to develop the household-based accessibility measure.

2.2 Data

To be able to do account for individuals, a *synthetic population*, a disaggregate representation of the real population, is created. The concept of synthetic population generation was mainly developed in the context of microsimulation models, in particular activity-based transport models (e.g. MATSim (Horni et al, 2016)), which require a disaggregate representation of all relevant individuals and their attributes (Guo and Bhat, 2007). Synthetic population generation is used to create such a representation when individual records of all relevant members of the real population are (as in most cases) not available. The members of the synthetic population are intended to reflect the real population of the planning area in terms of relevant demographic and socioeconomic properties such that marginal distributions of these properties (mostly given by a household census) as well as the correlation structure among these properties (mostly given a population microsample) are met (Zhu and Ferreira, 2014). The conventional approach for synthesizing the population is based on the methodology first developed by Beckman et al (1996), which is based on the Iterative Proportional Fitting (IPF) procedure developed by Deming and Stephan (1940). An alternative, newer procedure, which can match household-level and person-level marginal distributions at the same time, is called Iterative Proportional Updating (IPU Ye et al, 2009).

For Nelson Mandela Bay, a synthetic population is created based on the following procedure: 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 has 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. For each individual in the synthetic population, an activity chain sampled from the travel survey is drawn and assigned. The sampling process ensures based on a Hamming-distance procedure (Hamming, 1950) that the

activity chain is drawn from an observed individual with similar demographics. Then, a random 10% sample from the full population is drawn, which is used to perform the household-based accessibility computation. 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 attached to the MATSim (Multi-Agent Transport Simulation) framework (Horni et al, 2016) and which uses MATSim infrastructure like a router to find shortest paths. The MATSim extension for accessibility computations (cf. Sec. 3.3) is *not* used for the computation of the household-based accessibility indicator. Running this computation within the MATSim framework is plausible because MATSim is an agent-based simulation framework where every individual is represented as an agent with an individual daily sequence of activities and intervening trips (mostly referred to as a *plan*). So, each agent can be followed during all stages of the simulation process, which is necessary to conduct the household-based accessibility computation.

Based on local expert knowledge, levels of acceptability, and a scoring scheme, which translates people's travel patterns into a person-specific accessibility indicator are defined. Inspired by the classification of Litman (2010), introduced in Sec. 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 carried out according to Table 2. The overall score is calculated as the average of the individual scores, where a score of 2 corresponds to high accessibility, 1 to medium accessibility, and 0 to low accessibility.

To consider the important fact that not all facility types are relevant for different individuals, 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.

Table 2: Score for mobility

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

The relevance of a facility is determined by observing the activities (cf. Table 3) that are contained within the daily plan of the 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. For employed adults, the travel time

Table 3: 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

from home to work is calculated by shortest-path routing on the uncongested network. Based on this, a score is assigned according to Table 2. 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 in 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 router is applied with an assumed walk speed of 3km/h. Otherwise, the router that has been used for work destinations 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 *attended* facilities are used, but the *nearest* facilities. This is done since it is assumed that, for basic services, the nearest facilities are used, which is in line with other studies. Tanser et al (2006), for instance, find that in low-income communities 91% of all households use the nearest healthcare facility.

Transport options This factor quantifies the number of transport modes 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 4 are added. The number of points that a person receives is summed. The total num-

Table 4: Points for available transportation options

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

ber of points is used to obtain a transport options score as shown in Table 5. If an agent has a car trip in their plan, the car mode is assumed to be available. The availability of the mode *short-distance walking* is considered available if

Table 5: Score for transportation options

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

the agent can reach their work or school location 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 minibus, it is checked whether the next stop is reachable by a 20-minutes walk. Just as for short-distance walking, the assessment refers to options that are available from a geographic perspective and not a person's actual choice (potential, cf. Sec. 1.1). The value of 20 minutes is chosen because a majority of people is able to reach a stop within that time (Venter, 2011). Since *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 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 using Table 6. If an agent has a car trip in their plan, the car mode is assumed to be available at an access time of zero. For walk trips, the whole trip duration is considered. For trips by train, bus, or minibus, the time needed to walk to

Table 6: Score for walking time to transport

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

the nearest stop of that mode is used. Travel times for both walking trips and access to stops by walking are computed with an assumed walk speed of 3km/h.

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 Sec. 2.2) include: Shop (food), shop (other), healthcare, police station, post office, education, petrol station, and bank including ATM. Work places are not considered for this assessment, since nearby workplaces will, in most cases, not be useful to the observed person because of qualification mismatch. Scoring is done based on Table 7.

Table 7: Score for access to facilities

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

The four scores for mobility, transport options, walk time to transport, and access to facilities are summed, which yields the person's accessibility score. The score of a household is calculated as the average of the scores of the individuals belonging to that household.

This household-based measure is designed to be a context-specific accessibility metric tailored to local conditions and the high diversity both among the South African society and in terms of settlement patterns. To account for this diversity, the household-based measure combines different types of the measures that were introduced in Sec. 1.1. While, for instance, the *access to facilities* component stands representative for an isochrone-based measure, other components are pure reflections of travel times. Notably, many components of the household-based measure are focused on stated behavior as reported in the travel survey. Other components like the *mobility* score for healthcare and shopping facilities as well as parts of the availability of *transport options* score measure the potential of using these facilities, irrespective of their actual use, adhering to the notion of accessibility as a measure of potential (cf. Sec. 1).

2.4 Results

Fig. 2 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 Fig. 3, are analyzed in more detail. Each of these photographs is taken at the same scale.
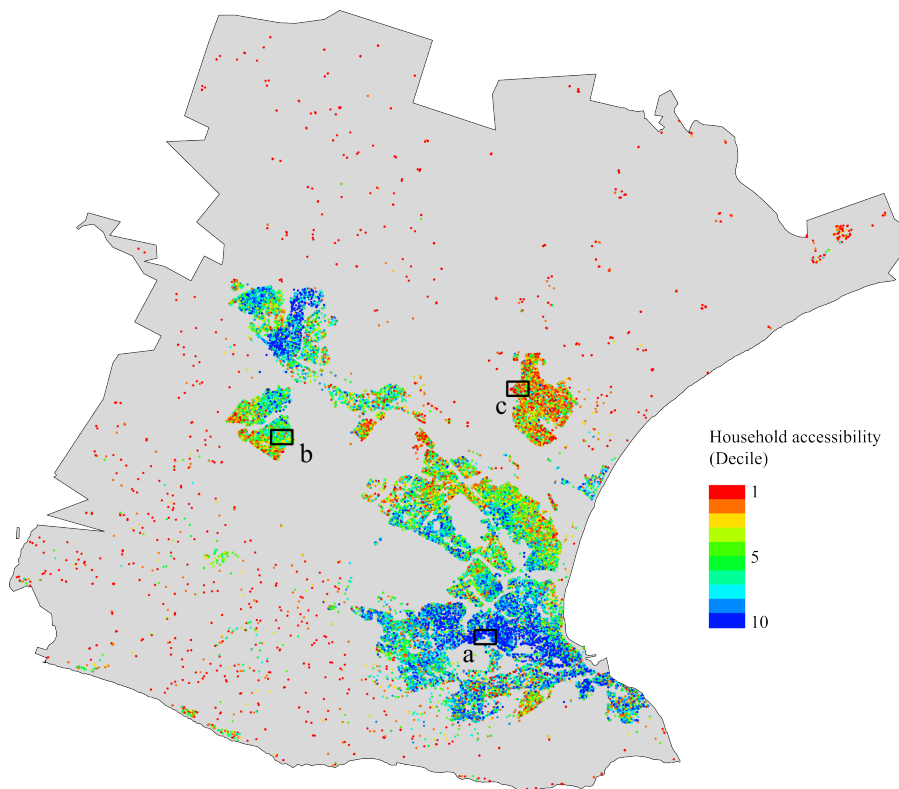
Fig. 2: Accessibilities according to the household-based accessibility indicator

Fig. 3a depicts an area situated in Port Elizabeth (marked as (a) in Fig. 2), 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

(a) Suburbs of Port Elizabeth

(b) Kwa Nobuhle township
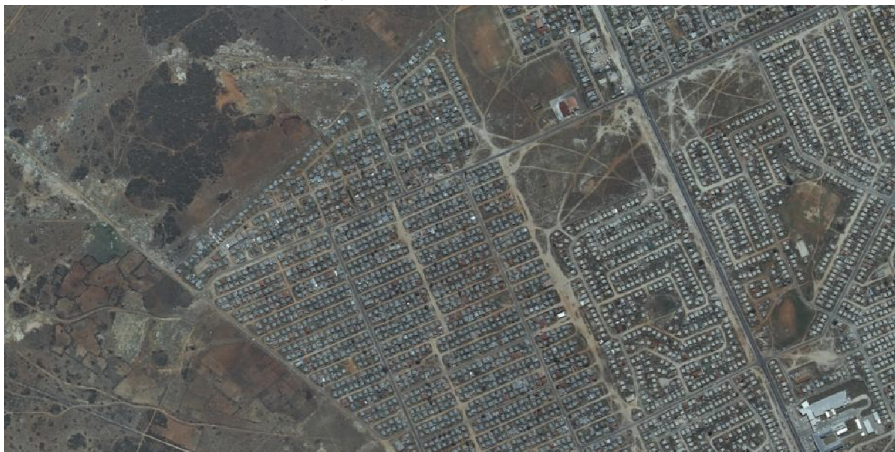
(c) Motherwell township

Fig. 3: Aerial photographs of areas with notably different accessibilities (Source: South African National Geospatial Institute (NGI))

(cf. Sec. 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.

Fig. 3b depicts Kwa Nobuhle (marked as (b) in Fig. 2), a township situated near Uitenhage, which is less accessible than the area in Fig. 3a. There are many houses in the area, but not many other facilities except primary and secondary schools (the larger buildings in Fig. 3b). The *access to facilities* score of this region is low because few 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.

Fig. 3c shows Motherwell (marked as (c) in Fig. 2), a township situated north of Port Elizabeth. This area shows low accessibility values because of a lack of facilities in that area. While there is an intermediary school in the bottom right of the map, the rest of the area is taken up by housing. Next to 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 centers of Port Elizabeth and Uitenhage are associated with lower accessibilities. Both in Fig. 3b and Fig. 3c, it can be seen that houses are smaller and the density of them is high, which indicates that people with lower income live here. This clearly demonstrates that low accessibility is generally associated with low-income households (Pozzi et al, 2010; Venter and Cross, 2014) and exemplifies the tradeoff between housing and transport costs (Lipman, 2006) as it has been discussed in Sec. 1.2.

3 Econometric accessibility

In Sec. 2, a household-based accessibility measure was introduced and applied. This can serve as the basis of household-based interventions as it explicitly considers individual households. However, Fig. 2 indicates that there are clear spatial patterns of large areas with low accessibilities. This observation may also be addressed by interventions into the infrastructure, either by improving the transport system or by adding required facilities (i.e. improving the land-use system).

3.1 Measure definition

To detect areas that are characterized by low levels of accessibility and to assess suitable interventions into the land-use and transport systems, an econometric accessibility measure (sometimes also referred to as utility-based measure, cf. Sec. 1.1) may be used, which computes the accessibility A_i at location i as

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

where j are opportunities somewhere in the study area and C_{ij} are the generalized costs of traveling from i to j. The mathematical form of Eq. 1, known as the *logsum* term, can be justified as follows:

1. Assume that the utility of each opportunity j, seen from the perspective of an agent a with a given home location i, is given by

$$U_{ij}^a = V_{ij}^{a,trav} + V_j^{a,perf} + \epsilon_{ij}^a \ ,$$

   where $V_{ij}^{a,trav}$ is the (usually negative) utility of traveling from i to j, and $V_j^{a,perf}$ is the (usually positive) utility of performing the activity at j. The error terms ϵ_{ij}^a take into account unobserved attributes, i.e. attributes which are not contained in the utilities either by choice of the analyst or because they are difficult or impossible to observe (Ben-Akiva and Lerman, 1985; Train, 2003). It should be mentioned that j are not the opportunity locations, but the opportunities themselves, as multiple opportunities at the same locations are counted separately (Nicolai and Nagel, 2014; Ziemke, 2016).

2. Further assume that the person-specific aspects of traveling and performing an activity are absorbed into the error terms:

$$U_i^a(j) = V_{ij}^{trav} + V_j^{perf} + \tilde{\epsilon}_{ij}^a \ .$$

   The remaining V_j^{perf} describes the effect that locations j can vary with regard to their general (i.e. non-person-specific) attractiveness. Assume that this property is absorbed into the error terms as well:

$$U_i^a(j) = V_{ij}^{trav} + V^{perf} + \tilde{\tilde{\epsilon}}_{ij}^a \ .$$

3. As typical in random utility modeling, assume that the error terms are identically and independently (iid) Gumbel-distributed random variables. This results in a choice model, where the probability of a person located at location i to choose an opportunity j is given by

$$p_i(j) = \frac{e^{\mu \cdot (V_{ij}^{trav} + V^{perf})}}{\sum_j e^{\mu \cdot (V_{ij}^{trav} + V^{perf})}} \ ,$$

   where the scale parameter μ is related to the width of the distribution of the error terms; a smaller μ leads to a wider distribution. Notably, the iid assumption is a rather strong assumption, since, for instance, the attractiveness of opportunities can be correlated between agents.

4. The expected maximum utility (EMU) of a person a in this choice situation is the logsum term[1]

$$E(U_i^a) = \frac{1}{\mu} \ln \sum_j e^{\mu \cdot (V_{ij}^{trav} + V^{perf})} \ .$$

---

[1]   The derivation of this formula can be found in textbooks on random utility modeling (e.g., Ben-Akiva and Lerman, 1985; Train, 2003).

5. The utility of performing the activity V^{perf} can be extracted:

$$E(U^a) = \frac{1}{\mu} \ln \left(e^{\mu\, V^{perf}} + \sum_j e^{\mu\, V_{ij}^{trav}} \right) = V^{perf} + \frac{1}{\mu} \ln \sum_j e^{\mu\, V_{ij}^{trav}} \ .$$

The second term is the econometric accessibility

$$A_i := \frac{1}{\mu} \sum_j e^{\mu\, V_{ij}^{trav}} \ .$$

Using the generalized cost of traveling C_{ij} as the utility of traveling V_{ij}^{trav}, this formula equals Eq. 1.

The econometric accessibility equals the expected maximum utility, which is derived from the individually best opportunity, *minus* the *average* utility V^{perf}, which is derived from doing the activity at *any* location. In consequence, A_i subsumes the effects of (1) the excess utility that a location j can have beyond V^{perf}, and (2) the usually negative utility to get to that location. In consequence, A_i can be positive (when the expected excess utility is larger than the expected travel effort), or negative.

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

Computing accessibilities using this formulation, the scale parameter μ and the utility of traveling V_{ij}^{trav} need to be determined. Typically, they are estimated from a survey by maximum likelihood estimation (Ben-Akiva and Lerman, 1985; Train, 2003). Given the present problem, the correct choice situation is destination choice. Notably, μ and V_{ij}^{trav} cannot be estimated separately as both values depend on each other (Train, 2003). In consequence, if an estimation for V_{ij}^{trav} from a destination choice survey is available, it should be inserted 'as is' into Eq. 1, and μ should be set to one.

Such an estimation was neither the focus nor within the scope of this study. Therefore, the MATSim default utility values (Nagel et al, 2016) were used in the absence of specific estimates. These values are chosen such that the effective values of time are approximately the same as the consensus values of the well-known Vickrey bottleneck scenario (Vickrey, 1969). Using these values based on such model-theoretic considerations has proven to be a valid approach in a number of scenarios when region-specific utility values were absent. Furthermore, Kickhöfer (2009) reports that applying these values to a real-world study leads to comparable results in terms of representativeness of real-world traffic patterns (as e.g. measured by comparing the model to real-world traffic counts) than using utilities that were estimated based on a local survey.

Since it is clear that the scale parameter is highly influential, a sensitivity analysis has been conducted (cf. Ap. A). The results confirm that the chosen

approach of using utilities based on model-theoretic reason is viable for the purpose of this study. In situations were the capability of the econometric measure is used to obtain results in monetary terms, a region-specific estimation of the model parameters is required.

In contrast to the household-based measure (cf. Sec. 2.1), the econometric measure is a pure measure of potential (cf. Sec. 1). No information about the actual selection of activity opportunities is required, but only information about the transport system (the network) and the land-use system (facilities). Since policies can be quite easily introduced by adapting these input data, the metric is easily applicable for policy analysis. The authors of this study are not aware of another approach where an econometric (utility-based) accessibility metric has been used in South Africa.

3.2 Data

In contrast to the household-based accessibility indicator introduced in Sec. 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. The exclusive use of OSM data is a highly beneficial property as the lack of available data is an often-encountered and well-known issue in many land-use and transport studies (Ziemke et al, 2015). For instance, Pozzi et al (2010) report difficulties in terms of data completeness and consistency, while Venter and Cross (2014) mention that various parts of their input data will likely not exist in other study areas.

3.3 Methodology

To compute the econometric accessibility indicator, the MATSim accessibility computation (Ziemke, 2016), which is one of several so-called *extensions*[2] of the modular MATSim transport simulation (Horni et al, 2016), is used. The MATSim accessibility computation can be run either on a zone-based or coordinate-based level. The latter is the more obvious choice because the MATSim transport simulation framework 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 often not feasible to conduct zone-based assessments of accessibility at a high level of spatial resolution. In contrast, the coordinate-based MATSim accessibility computation can be conducted on the basis of VGI like OSM, which contain data on network and activity facilities on a coordinate-based level.[3]

---

[2] See http://matsim.org/extensions

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

To calculate the accessibility A_i of a given origin location i to opportunities j, both the origin location i and the opportunity locations j are assigned to the road network. For every origin location 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 opportunities 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 based on Eq. 1.

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 of Sec. 2.

The MATSim accessibility extension is also able to conduct accessibility computations based on a congested network with time-dependent travel times. Time-dependent travel times can, for instance, be provided by a MATSim transport simulation run, which illustrates why an accessibility computation implemented as an extension to an agent-based transport simulation framework is reasonable. Moya-Gómez et al (2017) describe situations, in which the consideration of time-dependencies in the analysis of accessibilities is favorable. In the present study, the capability to compute congestion-dependent accessibilities is, however, not applied. (1) To run a MATSim simulation, a representation of the travel demand would be required. While such information was created for the household-based measure (cf. Sec. 2.2) and are, therefore, available, these information were not used for the econometric measure, adhering to the paradigm of only using freely available data to maintain the portability of the measure. (2) The Nelson Mandela Bay municipality is a metropolitan region with comparatively low levels of congestion so that only insignificant changes in accessibilities are expected when taking congestion into account. This assumption is confirmed by the municipality's Comprehensive Integrated Transport Plan (CITP), which identifies only ten out of 2,938 kilometers of highways in the metropolitan area to be severely affected by congestion (Nelson Mandela Bay Municipality, 2011, p. 89). Ap. B contains congestion-based

accessibilities for the morning peak and compares them to accessibilities that
were computed based on free-flow travel times. In line with the authors' expe-
rience and the municipality's CITP, it is confirmed that respecting congestion
does not significantly change accessibility patterns in Nelson Mandela Bay.

3.4 Results

The results of the econometric accessibility computation are shown in Fig. 4,
with a color scheme similar to Fig. 2. 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 Fig. 2. As
pointed out in Sec. 3.3, the econometric accessibility computation operates on
a spatial grid. As such, accessibility values are computed for *each* tile of this
spatial grid. For visual clarity, however, Fig. 4 only shows the tiles that cover
areas with a certain minimum population, making it more easily comparable
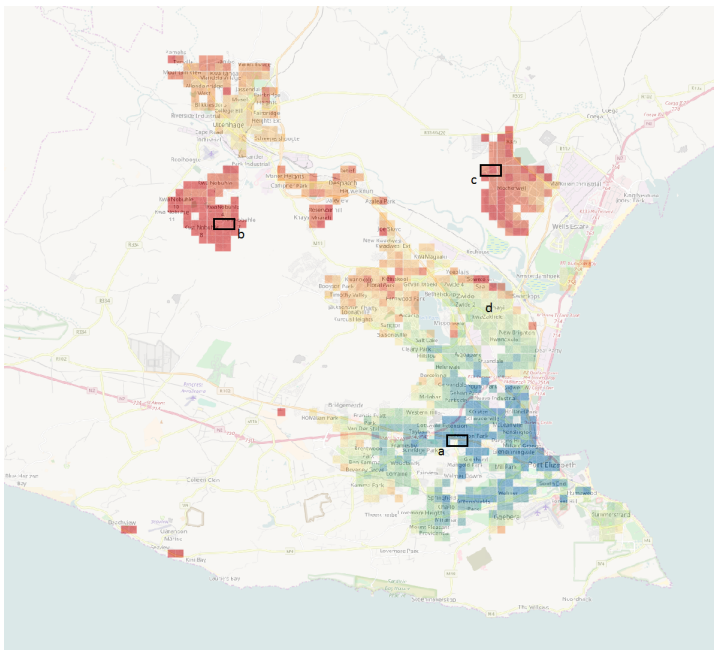to Fig. 2. Like in Fig. 2, it can be seen that the Port Elizabeth area (marked

Fig. 4: Accessibilities according to econometric accessibility indicator; decile
color ramp

as (a) in Fig. 4) has the highest accessibility. Just as for the household-based
accessibility measure, the two townships of Kwa Nobuhle (marked as (b) in Fig.
4) and Motherwell (marked as (c) in Fig. 4) have the lowest accessibility scores

of all larger populated areas, as visible in red shades of color. A bit different to the household-based accessibility, where Kwa Nobuhle showed somewhat better accessibilities than Motherwell, the two townships are associated with more similar accessibilities according to the econometric accessibility measure.

In summary, results show areas with low incomes to be those that are affected by poor accessibility. This is in line with previous findings from other studies (Pozzi et al, 2010; Venter and Cross, 2014). Ultimately, however, it is the geographic location that matters rather than the type of settlement. As such, the township of iBhayi (marked as (d) in Fig. 4; not shown in the aerial photographs in Fig. 3) has a much better accessibility than the two other township because of its less remote location.

4 Discussion and outlook

As shown in Sec. 2.4 and Sec. 3.4, the two accessibility measures, although quite different in terms of definition and properties, yield similar results. This is notable because the econometric accessibility computation requires fewer input data, less model assumptions, and no particular local expert knowledge. Both accessibility indicators clearly detect the significantly lower accessibilities of the two townships of Kwa Nobuhle and Motherwell. These results confirm the correlation of low incomes and low accessibility (Pozzi et al, 2010; Venter and Cross, 2014, also cf. Sec. 1.2).

As pointed out in Sec. 3.4, the results are slightly 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, they show more similar accessibilities according to the econometric measure. This is likely due to the fact that workplaces are contained in the household-based accessibility measure (cf. Sec. 2.3), while they are not (explicitly) considered by the econometric measure owing to the fact that extracting data for workplaces from OSM is more complicated than extracting data for education, leisure, shopping, or healtcare facilities. Interestingly, Zhan et al (2014) report similar difficulties of extracting work-related activity locations (e.g., industrial/manufacturing) in their procedure of inferring such information from social media ckeck-ins (e.g. Twitter). A significant part of the work facilities is still implicitly contained in the econometric accessibility measure because each leisure facility, hospital, or school also constitutes the location of a workplace. 'Pure' workplaces with no other function like some types of offices, however, are not taken into account. As Kwa Nobuhle is comparatively close to Uitenhage with its factories, this area reaches a better accessibility according to the household-based measure. First approaches to also infer the locations of work places from OSM and include them into the accessibility computation show promising results.

It was also found that the township of iBhayi (marked as (d) in Fig. 4) has considerably higher accessibilities than the other two townships. As such, one can conclude that having low accessibility is *not* equivalent with being located

in a township. Ultimately, it is the location that matters. A low-income area like a township can still have respectable prospects if it is more centrally located, which is in line with general wisdom about remoteness and centrality (Handy and Niemeier, 1997; Cheruiyot and Harrison, 2014; Christiaensen et al, 2003; Naudé and Krugell, 2004, also cf. Sec. 1).

The household-based measure (cf. Sec. 2.1) is intended to be a context-specific accessibility metric tailored to local conditions and the high diversity both among the South African society and in terms of settlement patterns. As it explicitly considers individual households, it is also capable of assessing impacts of household-based interventions, e.g. changed patterns of car availability. Notably, many components of the household-based measure are focused on stated behavior as reported in the travel survey, which limits its applicability to policy analysis.

In contrast to the household-based measure, the econometric measure is a pure measure of potential (cf. Sec. 1). No information about the actual selection of activity opportunities is required, but only information about the transport system (the network) and the land-use system (facilities). Individually specific attributes like car ownership are not picked up. Its computation operates on a regular spatial grid of measuring points. Like for the household-based measure, also the econometric measure is shown to be capable of detecting areas with high levels of accessibility deprivation and of identifying areas where interventions in the land-use and transport systems are advisable (cf. Sec. 3.4).

The fact that the econometric measure does not rely on a representation of the study area's population nor that it requires a travel survey, constitutes an advantage in terms of *policy evaluation*. While for the household-based measure a projection of the travel patterns into the future or under changed circumstances would be required, the econometric indicator can analyze the proposed policy by introducing according alterations into network and facility representations of the model. As such, changes in infrastructure will show up in accessibilities after rerunning the computation based on adapted input data.

Because of its exclusive reliance on freely-available, standardized input data (from OSM), the econometric measure is *easily portable to other regions*. An initial version of a system that captures the benefits of this portability has already been set up on a Geoserver[4]. While currently a few manual adjustments have to be made to perform accessibility computations for a new study region, the remaining steps to fully automatize this procedure and to make this procedure accessible to others like stakeholders or interested NGOs seem straightforward.

Finally, the econometric accessibility has an *interpretation in economic terms* as the expected maximum utility that can be derived at a given location from opportunities at other location. Other indicators, like potentially also the household-based indicator introduced in this study, are associated

---

[4] The current version of this Geoserver with results for other places like Cape Town, South Africa and Nairobi, Kenya can be accessed via the *VSP Geoportal* under `http://geo.vsp.tu-berlin.de`.

with the risk of being "appreciated as providing some interesting information; but their apparently subjective and empirical origin and formulation usually confines their practical use to a minor role" (Koenig, 1980, p. 145). Its clear interpretability also makes the econometric measure ready to be applied to project appraisal in monetary terms. For this, however, a region-specific estimation of the underlying utility function is required, which was not part of the present study.

Recalling the long-discussed obstacles to a more widespread use of accessibility measures in policy analysis, i.e. interpretability, usability, and lack of data availability (Handy and Niemeier, 1997; Naudé et al, 1999; Venter and Cross, 2014; Pozzi et al, 2010, also cf. Sec. 1.1), this study should be seen as a step of overcoming these obstacles.

5 Conclusion

In this study, two accessibility measures were introduced and applied. The two measures, a *household-based* measure and an *econometric* measure, are different in terms of data utilization and modeling philosophy. The household-based measure uses various types of data (OSM, a census, a travel survey, different sorts of facility information) and calculates accessibility by following people along their daily activity chain. In particular because of its reliance on the travel survey, the household-based measure is largely a measure of stated behavior. 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 indicator measures of potential of reaching opportunities. It does neither require information on individuals nor a travel survey. Its computation is performed on a regular spatial grid taking into account the variety of activity opportunities reachable from each of these grid points (measuring points) and the cost of overcoming the distance to these activity opportunities. Based on is mathematical formulation, the measure can be interpreted in economic terms as the expected maximum utility that a person at a given location can gain from opportunities at other locations. In contrast to the household-based measure, volunteered geographic information (VGI) are used as the exclusive data source. Because such VGI are available for many locations of the world based on the same standard, the econometric accessibility analysis procedure is easily portable to other study areas. Based on a Geoserver, a system to capture the benefits of this high level of portability has already been set up and shown its feasibility. Because of its sole reliance on OSM data, the approach has only a very low vulnerability to input data inconsistencies or incompleteness that many other accessibility computation procedures are struggling with. Notably, data coverage on OSM is steadily increasing, which will likely resolve issues regarding data coverage that may exist today.

While the household-based measure has strength in assessing impacts of household-based interventions because of its explicit representation of individual households, both accessibility indicators are suitable to detect areas that are accessibility-disadvantaged and to evaluate improvements in the land-use and transport systems to overcome accessibility deficits that may impair social and economic development.

As was shown, the econometric measure has significant advantages in terms of interpretability, portability, and usability and, at the same time, is associated with lower input data requirements. It therefore addresses some of the obstacles that have been seen as hindering the more widespread use of accessibility measures in policy evaluation. Especially, its exclusive reliance on standardized and freely available input data and its easy portability are a novelty that can support the more widespread application of accessibility measures.

References

Beckman R, Baggerly K, McKay M (1996) Creating synthetic baseline populations. Transportation Research Part A 30(6):415 – 429

Ben-Akiva M, Lerman SR (1985) Discrete choice analysis. The MIT Press, Cambridge, MA

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

Chen A, Yang C, Kongsomsaksakul S, Lee M (2007) Network-based accessibility measures for vulnerability analysis of degradable transportation networks. Networks and Spatial Economics DOI 10.1007/s11067-006-9012-5

Cheruiyot K, Harrison P (2014) Modeling the relationship between economic growth and time-distance accessibility in South Africa. Review of Urban & Regional Development Studies

Christiaensen L, Demery L, Paternostro S (2003) Reforms, remoteness and risk in Africa. Tech. Rep. 2003/70, World Institute for Development Economics Research

Deming WE, Stephan FF (1940) On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. The Annals of Mathematical Statistics

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

El-Geneidy A, Levinson D (2011) Place rank: Valuing spatial interactions. Network and Spatial Economics DOI 10.1007/s11067-011-9153-z

Geurs K, Krizek K, Reggiani A (2012) Accessibility Analysis and Transport Planning, Edward Elgar, chap Accessibility analysis and transport planning: an introduction, pp 1–12

Geurs KT, van Wee B (2004) Accessibility evaluation of land-use and transport strategies: review and research directions. Journal of Transport Geography 12(2):127–140

Guo J, Bhat C (2007) Population synthesis for microsimulating travel behavior. Transportation Research Record 2014:92–101, `http://www.caee.utexas.edu/prof/bhat/ABSTRACTS/SPG_Guo_Bhat.pdf`

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

Handy SL, Niemeier DA (1997) Measuring accessibility: an exploration of issues and alternatives. Environment and Planning A 29:1175–1194

Hansen W (1959) How accessibility shapes land use. Journal of the American Planning Association 25(2):73–76

Horni A, Nagel K, Axhausen KW (eds) (2016) The Multi-Agent Transport Simulation MATSim. Ubiquity, London, DOI 10.5334/baw, URL `http://matsim.org/the-book`

Kickhöfer B (2009) Die Methodik der ökonomischen Bewertung von Verkehrsmaßnahmen in Multiagentensimulationen. Diplomarbeit (Diploma Thesis), TU Berlin, Institute for Land and Sea Transport Systems, Berlin, Germany, also VSP WP 09-10, see `http://www.vsp.tu-berlin.de/publications`

Knowles R (2009) Transport geography. International Encyclopedia of Human Geography 1:441–451

Koenig J (1980) Indicators of urban accessibility: Theory and application. Transportation 9:145–172

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

Lipman B (2006) A heavy load: The combined housing and transportation burdens of working families. Tech. rep., Center for Housing Policy

Litman T (2010) Evaluating accessibility for transportation planning. Tech. rep., Victoria Transport Policy Institute

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

Moya-Gómez B, Salas-Olmedo MH, García-Palomares JC, Gutiérrez J (2017) Dynamic accessibility using big data: The role of the changing conditions of network congestion and destination attractiveness. Networks and Spatial Economics DOI 10.1007/s11067-017-9348-z

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

Nagel K, Kickhöfer B, Horni A, Charypar D (2016) A closer look at scoring. In: Horni et al (2016), chap 3, DOI 10.5334/baw, URL `http://matsim.org/the-book`

Naudé A, de Jong T, van Teefelen P (1999) Measuring accessibility with GIS-tools: A case study of the wild coast of South Africa. Transactions in GIS

Naudé W, Krugell W (2004) An inquiry into cities and their role in subnational economic growth in South Africa. Tech. Rep. 2004/8, World Institute for Development Economics Research

Nelson Mandela Bay Municipality (2006) Public transport plan 2006. Tech. rep.

Nelson Mandela Bay Municipality (2011) Comprehensive integrated transport plan 2011/12. Tech. rep.

Nicolai TW, Nagel K (2014) High resolution accessibility computations. In: Condeço A, Reggiani A, Gutiérrez J (eds) Accessibility and spatial interaction, Edward Elgar, pp 62–91

OpenStreetMap (2016) `http://www.openstreetmap.org/`, accessed 07 July 2016

Östh J (2011) Introducing a method for the computation of doubly constrained accessibility models in larger datasets. Networks and Spatial Economics DOI 10.1007/s11067-010-9129-4

Pozzi F, Robinson T, Nelson A (2010) Accessibility mapping and rural poverty in the horn of africa. Tech. Rep. 02-10, IGAD Livestock Policy Iinitiative

Reggiani A, Martín JC (2011) Guest editorial: New frontiers in accessibility modelling: An introduction. Networks and Spatial Economics DOI 10.1007/s11067-011-9155-x

Reggiani A, Bucci P, Russo G (2011) Accessibility and network structures in the german commuting. Networks and Spatial Economics DOI 10.1007/s11067-010-9149-0

Statistics South Africa (2001) Census. `http://www.statssa.gov.zahttp://www.statssa.gov.za/?page_id=3892`

Tanser F, Gijsbertsen B, Herbst K (2006) Modelling and understanding primary health care accessibility and utilization in rural south africa: An exploration using a geographical information system. Social Science & Medicine 63

The World Bank (2015) Gini index. `http://data.worldbank.org/indicator/SI.POV.GINI`, accessed 15 July 2015

Train K (2003) Discrete choice methods with simulation. Cambridge University Press

Vandenbulcke G, Steenberghen T, Thomas I (2009) Mapping accessibility in Belgium: a tool for land-use and transport planning? Journal of Transport Geography 17(1):39–53, DOI 10.1016/j.jtrangeo.2008.04.008, URL `http://www.sciencedirect.com/science/article/pii/S096669230800032X`

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

Venter C, Behrens R (2005) Transport expenditure: Is the 10% policy benchmark appropriate? In: Proceedings of the 24th Southern African Transport Conference, Pretoria, South Africa

Venter C, Cross C (2014) Access envelopes: A new accessibility mapping technique for transport and settlement planning. Town and Regional Planning

Vickrey W (1969) Congestion theory and transport investment. The American Economic Review 59(2):251–260

Ye X, Konduri K, Pendyala R, Sana B, Waddell P (2009) A methodology to match distributions of both household and person attributes in the generation of synthetic populations. Tech. Rep. 09-2096

Zhan X, Ukkusuri SV, Zhu F (2014) Inferring urban land use using large-scale social media check-in data. Networks and Spatial Economics DOI 10.1007/s11067-014-9264-4

Zhu Y, Ferreira JJ (2014) Synthetic population generation at disaggregated spatial scales for land use and transportation microsimulation. Transportation Research Record: Journal of the Transportation Research Board

Ziemke D (2016) Accessibility. In: Horni et al (2016), chap 18, DOI 10.5334/baw, URL http://matsim.org/the-book

Ziemke D, Nagel K, Bhat C (2015) Integrating CEMDAP and MATSim to increase the transferability of transport demand models. Transportation Research Record 2493:117–125, DOI 10.3141/2493-13

A Sensitivity test for the scale parameter

As pointed out in Sec. 3.1, the MATSim default scoring function (Nagel et al, 2016) has supplied the scale parameter μ and the utility of traveling V_{ij}^{trav}, which are required to compute the econometric accessibility scores according to Eq. 1. These values are based on model-theoretic considerations based on the well-known Vickrey bottleneck scenario (Vickrey, 1969). It was pointed out that using MATSim default values has proven to be a valid approach in a number of scenarios when region-specific utility values are absent. Also recall that μ and V_{ij}^{trav} are dependent on each other and can, therefore, not be estimated separately (Train, 2003).

As the scale parameter is highly influential, a sensitivity analysis has been conducted. Fig. 5 shows accessibilities to education facilities computed based on the MATSim default scoring function that was used to create the results shown in Sec. 3.4, for which the scale parameter ws set to one ($\mu = 1.0$).

In Fig. 6, the scale parameter was increased to $\mu = 2.0$. While the overall accessibility pattern does not change, one can detect some changes in Kwa Nobuhle (the township in the northwest, just south of Uitenhage). Some measure points now fall into the higher-scored deciles (i.e. are drawn in green colors). By increasing the scale parameter, the presence of facilities on a very local scale is given a higher weight. A measure point that is in the direct vicinity of very few facilities (education facilities in this case), will rank in the upper deciles of the evaluation. This is confirmed by the observation that the heterogeneity of accessibility values (i.e. different colors) in a given areas is somewhat higher than in Fig. 5.

In line with this, the opposite effect can be observed when the scale parameter is decreased to $\mu = 0.5$ (cf. Fig. 7). Here, Kwa Nobuhle is quite homogeneously red, i.e. people residing in any part of Kwa Nobuhle experience a low accessibility. The reduction of the scale parameter has effected that a higher number of facilities is required to reach a good score, but that longer trips are accepted to reach these facilities. Accordingly, the suburbs to the Southwest of the center of Port Elizabeth receive a better evaluation with a scale parameter of $\mu = 0.5$. The availability of a larger number of facilities in central Port Elizabeth that are reachable within an acceptable travel time brings them into the scope of people who reside in these suburbs and, thus, results in a quite good accessibility score.

The results confirm that the taken approach of using utilities out of model-theoretical reason is viable for a study with a comparative focus. When the capability of the econometric measure is applied to monetize results, a region-specific estimation of the model parameters is required.
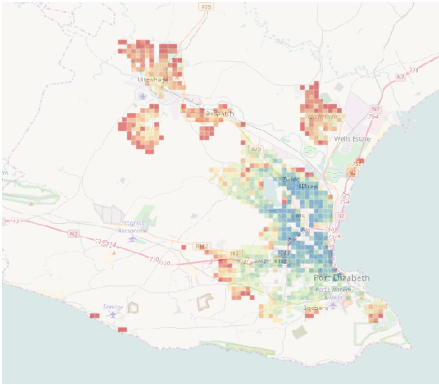
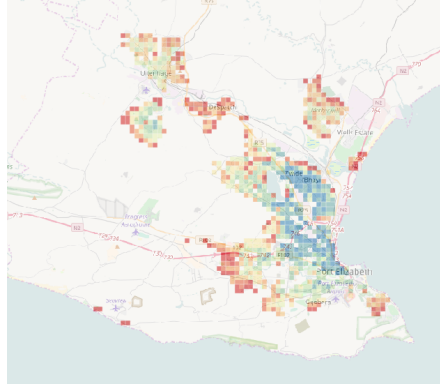Fig. 5: Scale parameter $\mu = 1.0$; decile color ramp.

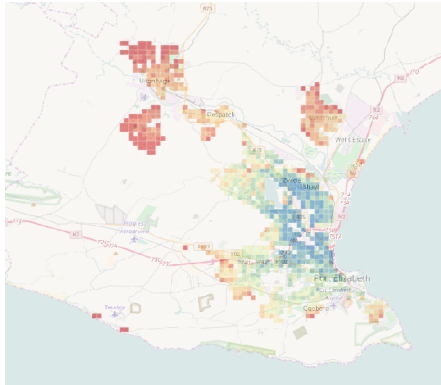Fig. 6: Scale parameter $\mu = 2.0$; decile color ramp.

Fig. 7: Scale parameter $\mu = 0.5$; decile color ramp.

B Considering time-dependent accessibilities

Some authors (e.g., Moya-Gómez et al, 2017) highlight that accessibilities may change during the course of a day due to time-dependent variance in the transport and land-use systems. As pointed out in Sec. 2 and Sec. 3, all computations carried out for this study are implemented within the MATSim transport simulation framework. As described in more detail in Sec. 3.3, this offers the opportunity to run accessibility computations that respect time-dependent (i.e. potentially congested) traffic conditions without significant alterations of the procedure. The reason why the results in Sec. 3.4 have been created under *un*congested traffic conditions (i.e. using free-flow network speeds) is the goal to exclusively use freely available and standardizes input data (OSM).

For the computation of congestion-based accessibilities, the synthetic population that was described in section 2.2, which relies on a (not publicly available) travel survey is required. The results of the congestion-based accessibility computation for the morning peak (8 o'clock) of a regular day are shown in Fig. 9. It can be observed that there are some detectable, but no significant changes in accessibilities when compared to *un*congested conditions as depicted in Fig. 8.

Under congested conditions, the accessibility values in the area southwest of Port Elizabeth (Mount Pleasant, Broadwood) range in worse quantiles than under free-flow traffic conditions. This seems plausible because from these locations only smaller roads (M9 – Buffelsfontein Road/Heugh Road and M7 – Main Road/RiverRoad), which are more prone to be affected by congestion, lead to the center of Port Elizabeth, while from the other directions bigger thoroughfares (R102 and N2 from the West; R75 from the Northwest, and R102 and N2 from the North) lead to the center of Port Elizabeth.

In line with the municipality's Comprehensive Integrated Transport Plan (CITP Nelson Mandela Bay Municipality, 2011), it can, therefore, be concluded that congestion effects do not seem to play an overly important role in Nelson Mandela Bay and that accessibility computations under free-flow conditions yield viable results.
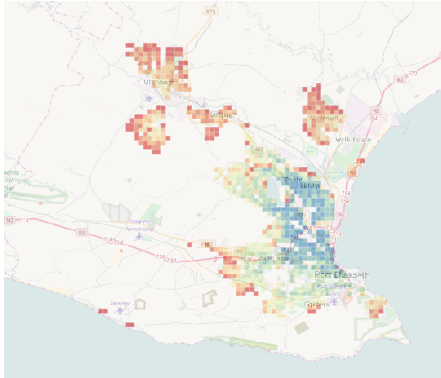

Fig. 8: Accessibility to education facilities under *un*congested traffic conditions; decile color ramp
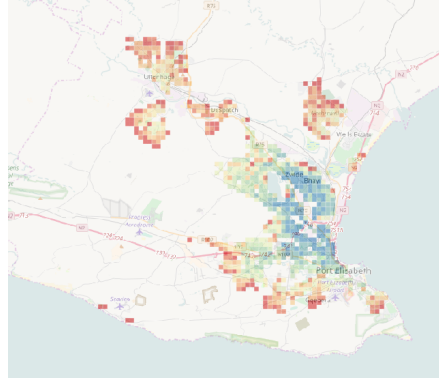
Fig. 9: Accessibility to education facilities under congested traffic conditions at 8 o'clock; decile color ramp