Partner acronyms

ETH	Eigenössische Technische Hochschule Zürich, Zurich, Switzerland
TUB	Technische Universität Berlin, Berlin, Germany
UoN	University of Nairobi, Nairobi, Kenya
UP	University of Pretoria, Pretoria, South Africa

Summary

Although the complex interactions between land use and transport are often acknowledged and appreciated, the state of practice is to fall back on purely transport and mobility-related metrics when evaluating infrastructure investment decisions. Especially when there is large economic inequality among citizens, answering the questions: “who gets the infrastructure benefits?” and “who pays for the infrastructure?” become loaded and controversial.

In this project we propose a more comprehensive metric, called “accessibility” that takes both land use and mobility into account. The metric is calculated at a disaggregate level allowing not only for results to be aggregated to a variety of useful levels, but also that the spread of the metric’s distribution is measurable for the areas of the study area.

The project’s point of departure is the emerging body of knowledge regarding agent-based transport simulation and the associated synthetic populations. In the project we aim to rely on open and freely available data to ensure the use of the metric is duplicable everywhere, but especially beyond well-funded studies, i.e. citizen’s or NGO efforts.

Early in the project the partners established two accessibility metrics. The first is a household-based metric where accessibility is calculated for each household member, taking the individual’s daily travel plan into account, and then aggregating the measure to the household level. The second is an econometric approach, which does not rely on a synthetic population or a computationally burdensome, agent-based transport model, and calculates a high-resolution spatial accessibility measure using freely available public data only. Both initial metrics were applied to the Nelson Mandela Bay Municipality in the Eastern Cape, South Africa. A GeoServer has also been established to host and visualize the metric results.

During the 2015/6 reporting period the work was extended to also apply the metric to Nairobi in Kenya, and the City of Cape Town in South Africa.

As a decision support tool this project was significant in allowing decision makers to better evaluate how to spend the national budget in an inclusive manner, accounting for the economic and social heterogeneity of the population, and the transport mode variety. The resulting metrics will be made available through high resolution, open-source, adaptive maps that will allow decision and policy makers to intuitively and easily access the data to support their planning endeavours.

Project objectives

The project concerned six main streams of work. Each of the following sections will address one of the streams, starting with a description of the work stream, followed by a report on the progress made.

Investigating microscopic accessibility metrics

The main objective of this work package was to come up with accessibility indicators that may be of interest to citizens or policy makers.

The concept of accessibility is usually credited to Hansen (1959). He defines accessibility as the "potential of opportunities for interaction". Accessibility is, at least in principle, always defined for a point location (e.g. Kwan, 1998), either as a measure of how well a certain location, e.g. for working or for shopping, can be accessed (access), or as a measure of how well all services/opportunities of a certain kind, e.g. for working or for shopping, can be accessed *from* a certain location (potential accessibility). In this project, it will be used primarily in the latter sense, i.e. as a measure how well households in arbitrary locations can access various services.

Work progress

The first attempt at the microscopic metric was done by Ms Jeanette de Hoog, a final year student at UP. The household-level metric currently requires the output from a Multi-Agent Transport Simulation (MATSim) model, and ist results are illustrated in Figure 1 for the Nelson Mandela Bay Municipality in the Eastern Cape, South Africa.

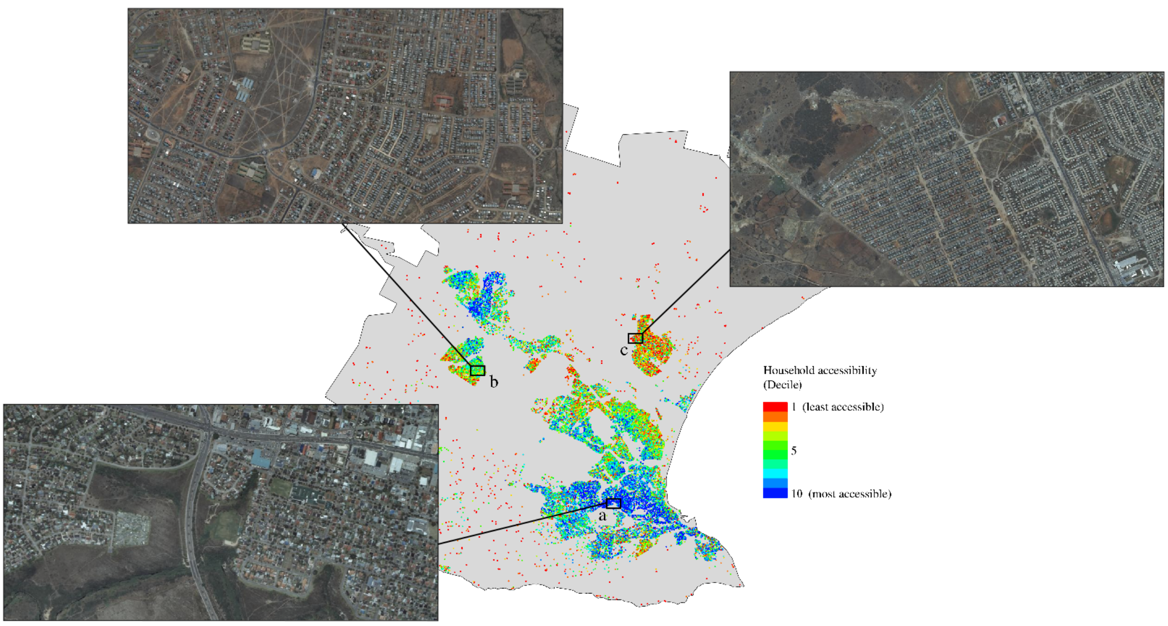

Figure 1 - Household accessibility for Nelson Mandela Bay metropole, South Africa.

The metric took the following aspects into account:

- ***Mobility***. This mobility factor measures the travel time from the household location of a given individual to a variety of activity locations such as education, workplace, healthcare and shopping.
- ***Transport options***. This factor quantifies the number of transport modes that are available to a household member, irrespective of whether the individual actually uses the specific mode.
- ***Walking distance to transport***. Here the metric looks at how long an individual walks to those modes that were used in the person's daily activity plan.
- ***Access to facilities***. This factor quantifies the variety of amenities within walking distance to an individual.

The household-based measure is strongly dependent on local expert knowledge, with associated thresholds inferred, and is also based on the output of a Multi-Agent Transport Simulation (MATSim) model, a data intensive exercise that is also computationally burdensome.

As an alternative, TUB developed an econometric accessibility metric that – in contrast to the household-based measure – does not rely on a synthetic population. This approach aims to capture the spatial structures observed in Figure 1. Accessibility at location i is denoted by A_i, and its accessibility to opportunities j is defined as

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

where $-C_{ij}$ is the generalized cost of travelling from i to j. The measure has the form of the so-called *logsum term*. It is also a high resolution accessibility measure and multiple opportunities j at the same location are all separately counted.

The calculation of the econometric accessibility work different than when using a synthetic popualtion as it does not use such data. Instead, accessibility values are computed for each tile of a given spatial grid and, thus, also for locations where nobody lives. As an example of the results, Figure 2 shows the econometric accessibility to work opportunities.

The details of the two metrics were presented at the European Regional Science Association (ERSA) conference[1], and a revised manuscript has since been accepted for publication in an ISI peer-reviewed journal:

Ziemke, D., Joubert, J.W., Nagel, K. *Accessibility in a post-apartheid city: Comparison of two approaches for accessibility computations*. Forthcoming in *Networks & Spatial Economics.* Doi: 10.1007/s11067-017-9360-3.

[1] http://EconPapers.repec.org/RePEc:wiw:wiwrsa:ersa15p1614

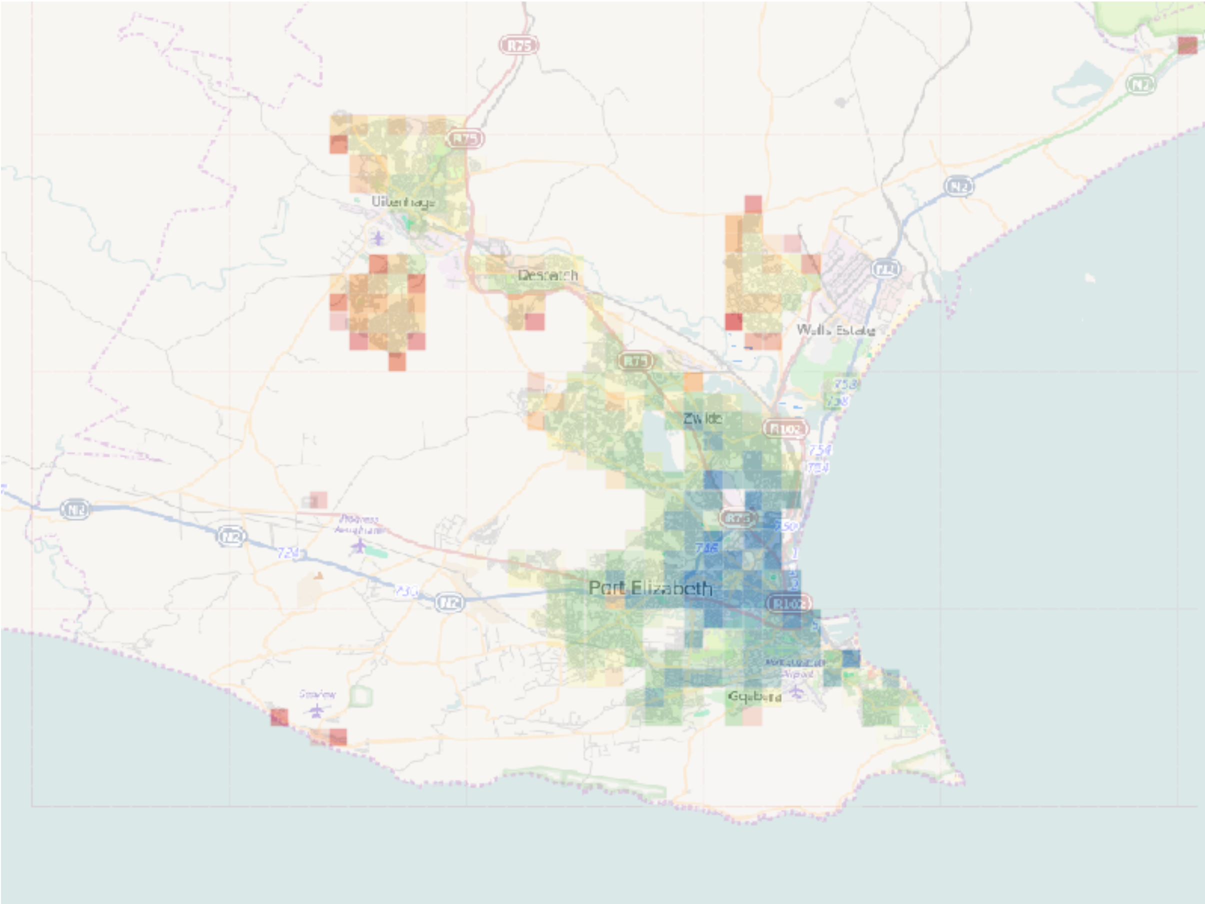

Figure 2 - Econometric accessibility to work locations by private car. Only tiles with a certain minimum population are shown and are superimposed on the *OpenStreetMap* layer.

Application of the accessibility metric

A recent advance is that much more input data is now publicly available, often crowd-sourced. Our primary source of open data was *OpenStreetMap* (http://osm.org), which seems to have established itself as a uniform and worldwide accessible standard for crowd-sourced and other geo-data. In fact, the use of *OpenStreetMap* is favoured in the South African environment, the result of strong advocacy for geospatial data to be more readily available for decision-making and decision support. To that extent, the South African National Geospatial Information (NGI) has started to move all its Geospatial Information System (GIS) data over to *OpenStreetMap*.

This offers the unprecedented opportunity to bind new computational methods together with standardized input data. This make this project's approach very portable, that is, it could be easily ported to a different spatial location, to the analysis of other types of services, or even a different spatial scale. Which is exactly what we've done in this project.

At present, the transportation network information provided by *OpenStreetMap* is already sufficient in most places to obtain plausible results; this holds, in particular, to urbanized locations also in developing countries. The data stock is also growing at an unprecedented pace, so even if input data is not fully available yet it pays off to invest into methods and infrastructure to harvest that data in the (near) future. Finally, the types of data covered by

OpenStreetMap are ever increasing. Concerning pedestrian access in urban areas, it may already be the best global data source around; for example, it contains a full walk network of Kibera in Nairobi which is more than what Google Maps can feature (accessed 6-Apr-2013). The services, to which access should be provided, are often already mapped in *OpenStreetMap*; if not, neighbourhood organizations may map and add these or other service locations to the existing data infrastructure.

Computations are done within the context of the Multi-agent Transportation Simulation (MATSim) (www.matsim.org) framework. Within MATSim, converters to process *OpenStreetMap* data or algorithms for high-resolution accessibility computations are already available. It comes with an established server infrastructure that allows shared programming and automatic regression testing. There already exists a MATSim application for the Gauteng area in South Africa [Fourie and Joubert, 2009], using both publicly and non-publicly available data, generating time-dependent traffic congestion patterns for the area. Within this work package, we computed accessibility maps according to the criteria developed in the first work stream. These will include maps that are local, regional (Figure 2 for Nelson Mandela Bay Metro in South Africa) or national in scope; they will refer to a variety of different services; they will compute accessibility by a variety of transport modes; and they will use different accessibility indicators. The maps will firstly be computed both for Kenya and South Africa based on publicly available data, primarily *OpenStreetMap*. In the case of Kenya, the *OpenStreetMap* data was augmented with supplementary data like the *Kenya Open Data Initiative* (KODI). This enabled an additional; distinction of facilities, for example primary versus secondary schools, and private versus public schools.

Figure 3 shows, as an example, the accessibility to drinking water taps in the township (slum) of Kibera in Nairobi, Kenya.

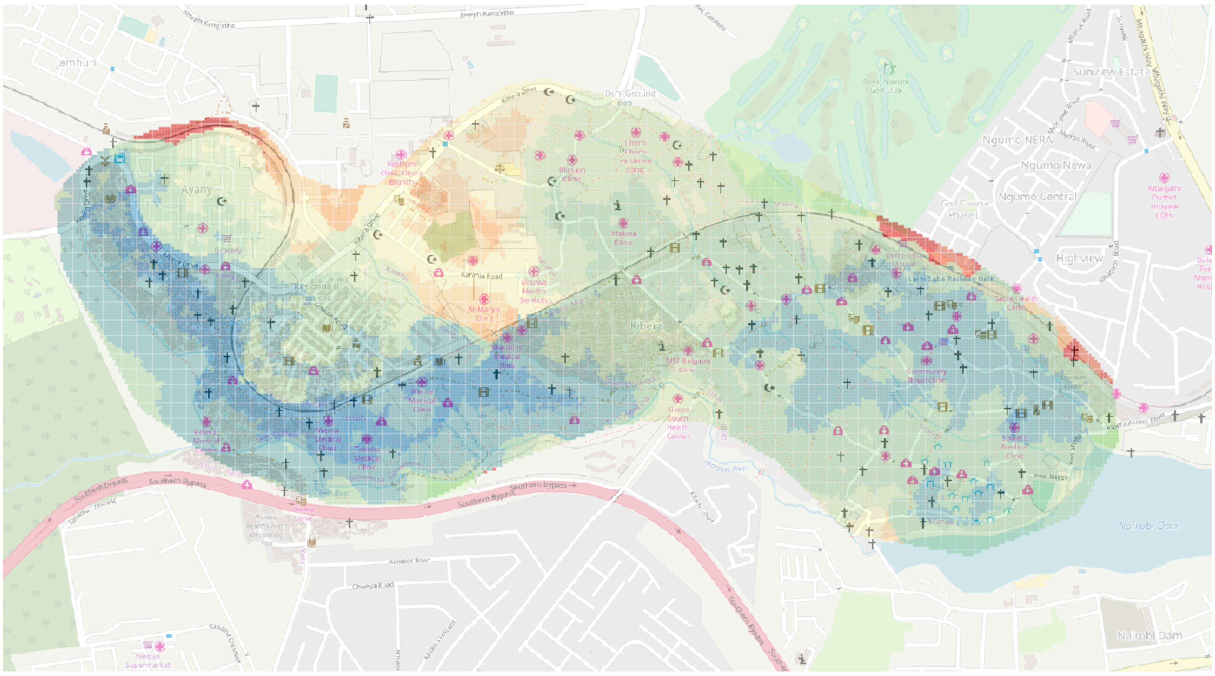

Figure 3 – Accessibility to drinking water taps in the Kibera township in Nairobi, Kenya. Results are superimposed on the *OpenStreetMap* data layer.

High accessibility is indicated by the blue shades while low accessibility is indicated by the red shades.

The accessibility maps are available on the VSP hosted GeoServer at http://geo.vsp.tu-berlin.de/geoserver/web/ with a brief overview at the geo portal, http://geo.vsp.tu-berlin.de.

Furthermore, the results were presented at the NECTAR Cluster 1 & 6 Workshop on Transport Infrastructures for better Accessibility, Equity and Territorial Cohesion: “Accessibility computations based on open data for different spatial scopes, transport modes, and activity types – Analyses for Nairobi, Kenya”. 2016. Warszawa, Poland.

The code for the computations are available at https://github.com/matsim-org/matsim/tree/master/contribs/accessibility in the public domain.

Accessibility of Southern African townships

A special topic was the consideration of the access and accessibility of South African townships to services, as they are often far away from city centres due to the legacy of Apartheid forced relocation, and as a result from (certain types of) jobs or (certain types of) services. The primary mode of transport between the townships and the city centres are shared minibus taxis, with a mode share of 50% or more for these relations. In Kenya, for example, this (para)transit mode is referred to as *Matatus* while they’re known in South Africa as *minibus taxis*. Conventional traffic models typically ignore this mode, or are not very good at it, reasons being sufficient spatial data (e.g. details about the minibus route system) and modelling challenges.

Röder (2013) and the subsequent Neumann et al (2015) developed an algorithm that is able to synthetically generate a minibus taxi simulation for a city or region that resembles the real system. Albeit improvements to this approach will surely follow, it will serve as an indispensable ingredient for a detailed investigation of the accessibility of SA townships to services. Subsequent computational improvements to the minibus contribution in MATSim is ongoing to ensure computational efficiencies. The TUB was granted a project extension to further the stream of research. They applied aforementioned evolutionary market simulation approach again to be able to synthetically create a sufficiently realistic ‘schedule’ of those informal forms of public transport so that corresponding accessibility computations (based on that schedule) can also be performed for other regions where no such schedule from a previous project is available yet. Based on this approach, accessibilities for different activity types were computed for the mode of Matatus in Nairobi, Kenya.

See the first section for the calculation of household and econometric accessibility for the accessibility results fort he Nelson Mandela Bay Municiaplity area. Work is ongoing in the thesis of PhD candidate, Mr Gerhard Hitge, to extend the accessibility work to the City of Cape Town.

Generation of a descriptive synthetic population

Census data is often captured at the resolution of an enumeration area. When census data is reported, aggregating the individual records to sub place level ensures anonymity. Sub places, in turn, can be aggregated to main places, which in turn can be aggregated to a variety of country-specific demarcations, and ultimately national level.

The subplace tables provided in census data give us, in South Africa for example, the number of males and females in the area, or the number of Blacks, Coloureds, Asian/Indian and Whites. The tables also tell us how many households there are with one, two, three, etc. individuals in the household, and how many people speak Afrikaans, English, Sepedi, etc. as first language. The tables, however, report independently on each characteristic. It therefor does not tell us how many Setswana-speaking black males between 15 and 18 years of age there are in an area.

Censuses do, however, provide the detailed records for an anonymous sample of unit records, in South Africa 10% of the respondents. The sample is representative of the geographic distribution, as well as the different demographic characteristics. For each record in the 10% sample we know all of the census attributes: age, gender, race, language, household size, income, etc.

The process to generate an entire 100% synthetic population was published in Joubert (2018) and follows a two-stage approach. The first, called *fitting*, uses a Bayesian Network approach to understand the causal relationships in household structures. The second stage, called *generation*, is then using the Bayesian Networks to generate a pool of households and individuals from where weighted sampling is done following a generalized raking procedure.

In the spirit of making data publicly available, and since the generation process is random, multiple instances of the synthetic populations for the nine (9) city/metropolitan/provincial areas shown in Figure 4 were generated and released into the public domain. The description of the data generation process is covered in

Joubert, J.W. (2018). Synthetic populations of South African urban areas. *Data in Brief*, **19**, 1012-1020,

while the data instances itself is released on Mendeley at
http://dx.doi.org/10.17632/dh4gcm7ckb.1

During the 2015/6 reporting year, a 100% sample for the City of Cape Town was also generated. The population is augmented with travel demand from the 2013 household travel survey conducted by the City, and specifically the travel diary component that provides a detailed description of a 24-hour period of travel for a subset of people being surveyed. The following activities are included in the diary:

- home;
- work;
- primary and secondary education;
- tertiary education;
- shopping;
- leisure;

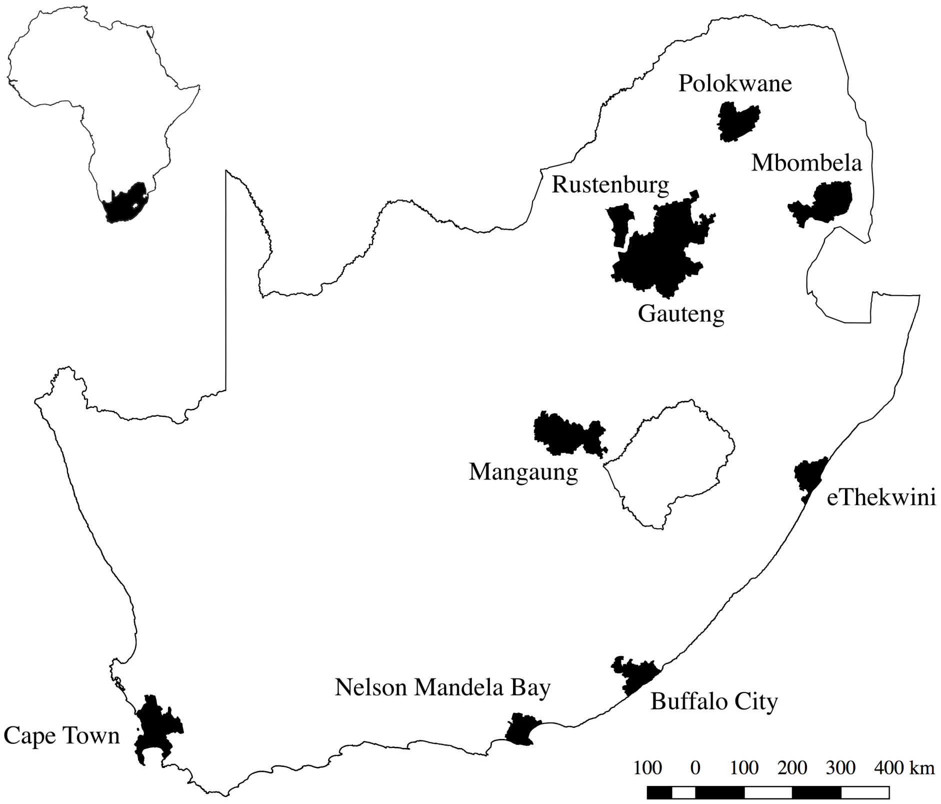

Figure 4 - Metropolitan areas for which full synthetic populations were generated (Source: Joubert, 2018).

- visiting friends; and
- other.

The modes connecting the activities include:

- private car (as driver);
- private car (as passenger);
- bus (the Golden Arrow Bus Service);
- minibus (paratransit);
- MyCiTi (bus-rapid transit system in Cape Town);
- walk;
- rail (the MetroRail); and
- other.

Matching activity chains to the synthetic population is complete. What remains to be done is relocating both primary and secondary activity locations of the sampled activity chains. This will be covered in a follow up project on transport land use interaction.

Constructing a daily schedule aware of accessibility

Accessibility by itself is a measure of how well certain services can be accessed from a given location with one trip (movement there and back again). This is a full assessment, when the population is indeed generally returning home between different activities, or when the policymakers are willing to ignore the remaining activities, say when focusing on health care or work.

If trip chaining is common in a study area, i.e. the linkage of multiple activities within chains of trips, the common accessibility measures are likely to be biased, by omitting the further possibilities available during the day. The incidence of trip chaining is well known to have links with home location, commuting and other socio-demographic variables. As trip chaining has positive impacts by reducing the total mileage travelled and the related emissions, it is generally encouraged by transport policy, even in the absence of a comprehensive measure of its benefit.

The starting point to the construction of the expanded measure of accessibility will be the class of activity-based travel demand models (Ben-Akiva, Bowman and Gopinah, 1996, and the many later models in their tradition). These models estimate the probability with which the population in the study area will choose a particular daily schedule, i.e. trip chain. Their final log-sum term is therefore an overall measure of benefits/utility available to a traveler of a given socio-demographic profile beginning the day at a particular location.

In this work package the aim was to extend the basic accessibility measure to also account for the trip planning behavior of travelers. The extended measure should be parsimonious and enable a comparison between European (Swiss) and African contexts.

At first a nested logit model was used as it considers destination and mode choice jointly. The modes were nested (one nest per mode) and the error terms were assumed to follow a Gumbel distribution. For each individual there exists one error term per mode, and one per destination/mode pair. The assumption is made that the unobserved attributes (not captured by the model or data) is spread equally between the mode and destination.

A method was developed to estimate a joint destination/mode choice model based on an activity-travel survey, encoded in the form of a MATSim population file. It uses socio-demographic attributes of the person, personal constraints (such as car availability) and travel times as computed from the MATSim input files, that represent the infrastructure in detail. The choice set of destinations is constrained by the activity chain observed in the survey.

Accessibility values can be computed for each member of a synthetic population, taking into account socio-demographics, car availability, transport system and constraints from the activity chain. In this way, several aspects that do not come out of other measures can be apparent:

- The effect of car availability, and in particular its spatial distribution, can be made apparent
- Persons with longer commutes have a more constrained destination choice set, having an influence on accessibility

The approach was used to data from Switzerland and Cape Town, South Africa. It allowed to reveal interesting effects that do not appear when using other measures: while car availability tends to level accessibility out between urban and rural areas in Switzerland (because almost anybody can own a car to compensate longer public transport travel times), car availability increases differences in accessibility between Cape Town city and the isolated township of Atlantis.

In both cases the revealed preference (mode and destination choice) was available. That said, the revealed demand for the bus rapid transit (MyCiTi) mode, for which we had the detailed public transit routes and schedules, were very low compared to the paratransit, bus and metrorail for which routes and schedules were *not* available.

The initial model was fit using leisure as an activity type since the logit assumptions made more sense. The base model samples, for each leisure activity, 200 additional leisure facilities as the choice set. The choice set area is an elipse anchored at the preceding and subsequent activity of the current leisure activity. Travel times to the choice set is calculated using the MATSim infrastructure. The base accessiblity is shown in Figure 5.

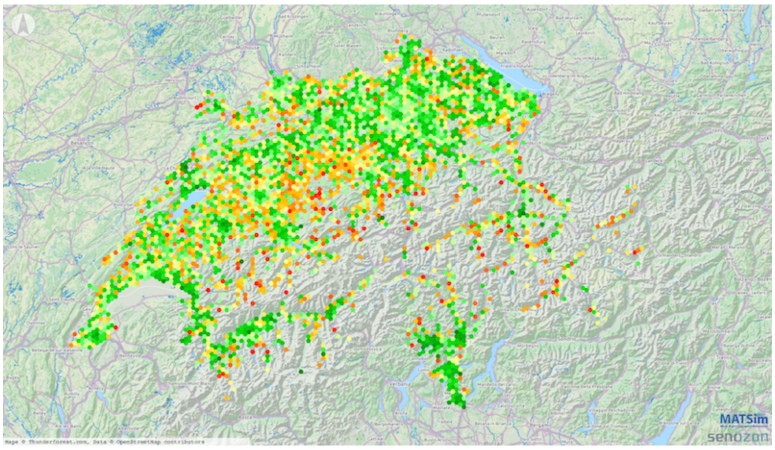

Figure 5 - Base accessibility to leisure. Green denotes higher accessibility, and red lower.

When all modes of transport is considered, the result is shown in Figure 6.

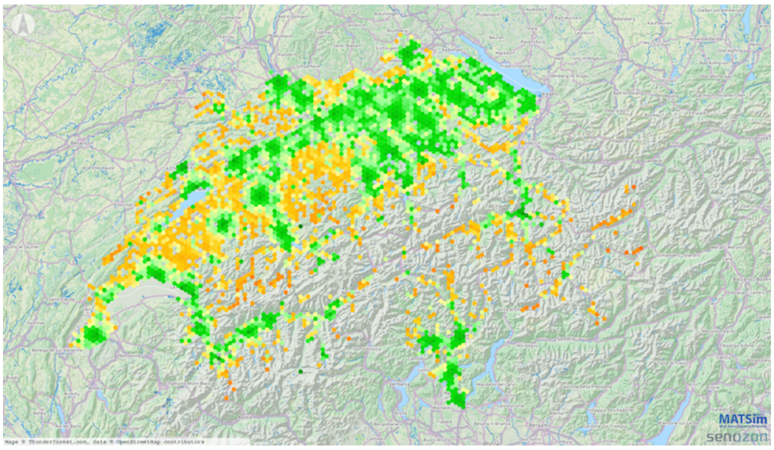

Figure 6 - Accessibility using all modes of transport. Green hotspots of high accessibility is notable around cities and denser urban areas where public transport is more regularly and densely available.

It then follows that the advantage of having access to a car is more pronounced in rural areas where public transport is less frequently available. This is highlighted in Figure 7.

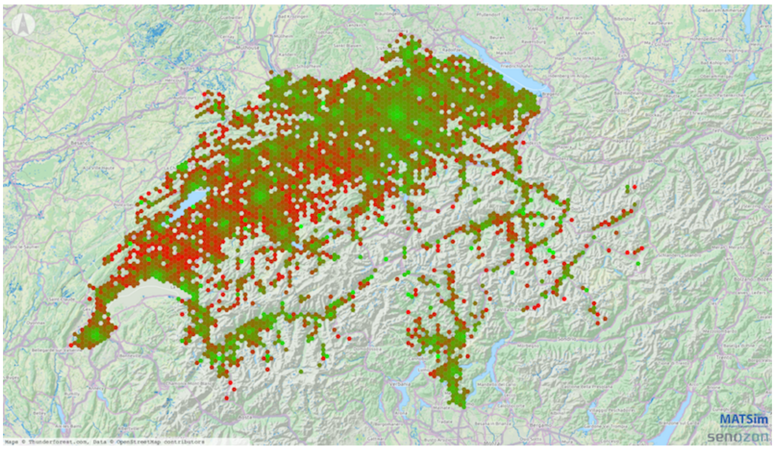

Figure 7 - The advantage of having a car to access leisure activities is higher (red) in the more rural areas.

For the data in Cape Town, more accurate activity locations were not available, so the initial assumption was made to assume more central (close to the centroid of the City's boundaries) has more leisure facilities. The results are shown in Figure 8.

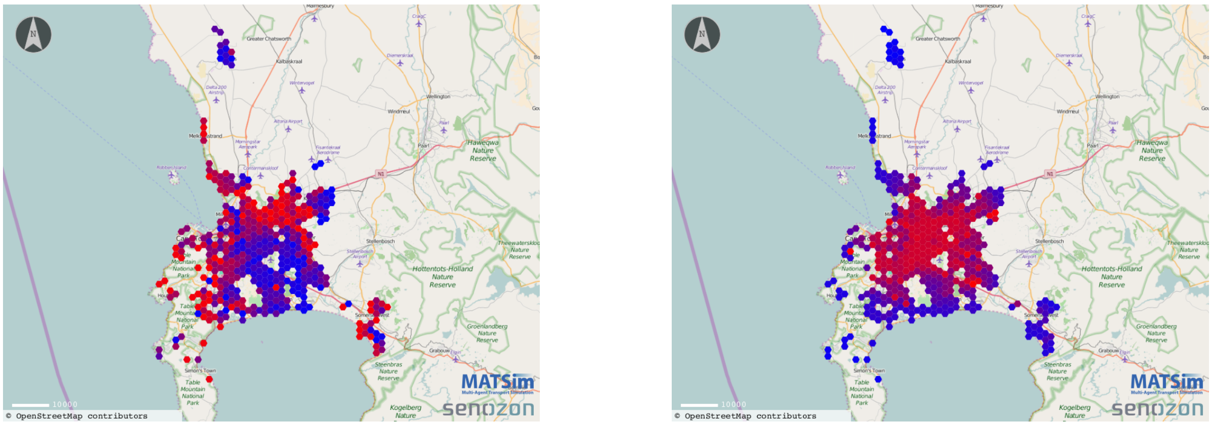

Figure 8 - Accessibility to leisure in Cape Town (left) where blue denotes higher accessibility. Having access to a car is again more advantages (blue) for those on the periphery of the city.

These results requires some more refinement in terms of leisure activity locations.

An additional action item that has already received some attention is considering competition when calculating accessibility. This is because accessibility essentially looks at two competing objectives: access to many (dense) opportunities, and the associated speed of accessing them, which decreases as higher density of people and vehicles reduces travel speeds.

The model used in this work package was described in Dubernet and Axhausen (2016). As such, the objectives listed in the proposal were achieved.

Dubernet, T., Axhausen, K.W. 2016. Using a joint destination-mode choice model for developing accessibility measures, in: STRC (Ed.), 16th Swiss Transport Research Conference. Ascona. Available online from
http://www.strc.ch/2016/DubernetT_Axhausen.pdf

An extension of the approach to take into account capacity constraints in the form of "shadow prices" as per Palma et al. (2007) or Vitins et al. (2015) was developed, though not published.

Visualisation and dissemination of results

The value of any quantitative metric is in its ability to provide insightful decision-support. To be usable, the metric needs to be intuitively understood, and visualised in a spatio-temporal manner that decision-makers can relate to. For example, even 20 years into South Africa's democracy, the racially segregated urban design is still apparent and should be reflected in the results.

The results of the accessibility measure is communicated through an openly accessible GeoServer by the TUB team. An overview is provided at http://geo.vsp.tu-berlin.de while the computational results are hosted at http://geo.vsp.tu-berlin.de/geoserver/web/. An example of the hosted results were shown in Figure 3, while Figure 9 shows an example screenshot of the interactive accessibility maps.

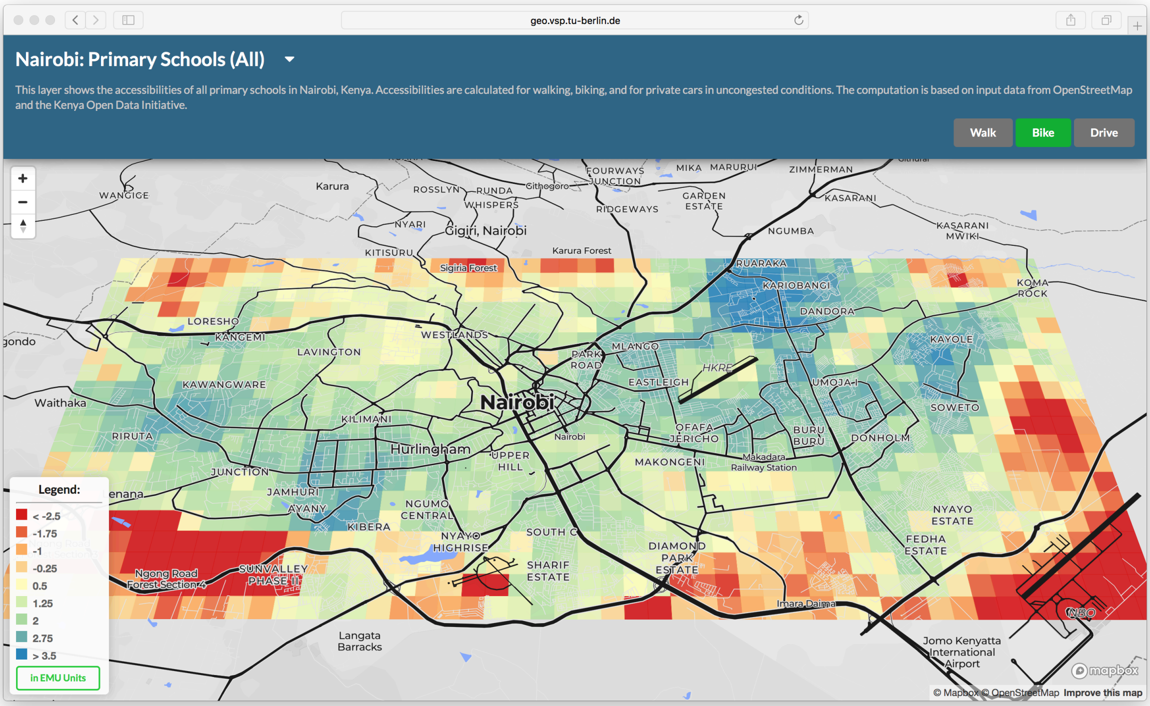

Figure 9 – Screenshot example of the interactive GeoServer results. Here the accessibility to all (public and private) primary schools, using bike (cycling) as the mode of transport, is shown.

References

Ben-Akiva, M., Bowman, J.L., Gopinath, D. (1996). Travel demand model system for the information era. *Transportation*, **23**(3), 241-266.

Hansen, W.G. (1959). How accessibility shapes land use. *Journal of the American Institute of Planners*, **25**(2), 73-76.

Kwan, M.-P. (1998). Space-time and integral measures of individual accessibility: A comparative analysis using a point-based framework. *Geographical Analysis*, **30**(3), 191-216.

Palma, A. de, Picard, N., Waddell, P.A. (2007). Discrete choice models with capacity constraints: An empirical analysis of the housing market of the greater Paris region. *Journal of Urban Economics*, **62**, 204-230.

Neumann, A., Röder, D., Joubert, J.W. (2015). Toward a simulation of minibuses in South Africa. *The Journal of Transport and Land Use*, **8**(1), 137-154.

Röder, D. (2013). Simulation of South African minibus taxis. Masters dissertation, FG Verkehrssystemplanung und Verkehrstelematik, Institut für Land- und Seeverkehr, Technische Universität Berlin, Germany.

Vitins, B.J., Erath, A.L., Axhausen, K.W. (2015). Integration of a Capacity Constrained Workplace Choice Model: Recent Developments and Applications for an Agent-Based Simulation in Singapore (Arbeitsberichte Verkehrs- und Raumplanung No. 1106). IVT, ETH Zurich, Zurich.

Deliverables and Milestones

See Table 1 for the list of deliverables, and Table 2 for the list of milestones. No explicit deliverables and milestones were identified in the proposal. Instead, the deliverables were planned around the four PhD candidates in terms of their thesis progression and article submissions. At the time of writing all four PhD candidates were still enrolled towards the completion of their theses.

Project management

All partners could only confirm funding from their respective agencies by September 2015. A consortium agreement was drafted, following the DESCA Horison 2020 Model Consortium Agreement. The agreement was signed by all four partners by 9 September 2015.

The following consortium meetings have already been concluded:

- 3-4 June 2015, Pretoria, South Africa.
- 1-2 December 2015, Zurich, Switzerland.
- 13-14 June 2016, Nairobi, Kenya.
- 8-9 December 2016, Berlin, Germany.

In the end, the Kenyan partners only received a single full year's funding from their agency. Both ETH and UP received their full funding, while TUB received an extended amount for additional deliverables agreed with their funding agency.

Resources

The main IT resources currently used are the GeoServer hosted on the TUB server, and the computational servers by the respective partners. Four PhD students are currently participating:

- Mr Thiebaut Dubernet, ETH;
- Mr Sammy Matara, UoN;
- Mr Dominik Ziemke, TUB; and
- Mr Gerhard Hitge, UP.

Finances

The summary financial report is shown in Table 3 and is updated as and when consortium members made their financial reporting available on the shared Dropbox folder.

The project was extended by the UP's funding agency, Department of Science and Technology (South Africa), on a no cost-basis from 2017/03 until 2018/03. One of the reasons was the late recruitment of a PhD candidate, Mr Gerhard Hitge.

Consortium member TUB was granted an extension to the project to the by the German Federal Ministry of Education and Research to include computation of accessibilities for informal modes of public transport (paratransit, minibus, matatus) into the process of automated accessibility computation. The extended funding is notable in the payments made to TUB in Table 4.

The reduced amount received by the South African partner is a result of exchange rate deterioration from when the budget was submitted until when the project was approved. The South African Rand (ZAR) value remain unchanged, except for the effect of inflation. At the time of writing this report both the Swiss and Kenyan partners have only received the first of their three payments.

Project spin-offs

Following the work of this project, a new project is funded by the Deutsche Forschungsgemeinschaft (DFG) titled *„Implementation and application of a tightly integrated behavioural land use and transport model“* with the investigators being Prof Rolf Moeckel (Technical University Munich), Prof Kai Nagel (TUB), and Prof Johan W. Joubert (UP).