

# The survivability of cycling in a co-evolutionary agent-based model\*

Gerhard Hitge<sup>a,1</sup>, Johan W. Joubert<sup>a,b,2,\*</sup>

<sup>a</sup>Centre for Transport Development, Department of Industrial and Systems Engineering,  
University of Pretoria, Private Bag X20, Hatfield, 0028

<sup>b</sup>KU Leuven, Institute for Mobility, Centre for Industrial Management (CIB), Celestijnenlaan 300, 3001, Leuven, Belgium

---

## Abstract

Many cities plan to grow cycling as a prominent mode to improve accessibility and environmental and financial sustainability. However, relatively few cities have made meaningful headway in this direction. Policymakers would be more inclined to implement the necessary interventions when they have certainty about potential demand, especially knowing where it is located in space.

This paper introduces an approach to estimating potential cycling demand using agent-based modelling to determine who may benefit from switching from their current modes to cycling. People benefit when they obtain a similar or higher travel utility score when cycling between their daily activities than when using their existing modes. The model is based on individual mode selection, that all activities in the trip chain are included and can include detailed road and cycle network elements. The co-evolutionary mechanisms within the agent-based simulation allow us to test the potential for cycling relative to the performance of other modes on the network.

The case for Cape Town, South Africa, shows that about 32% of those that travel would benefit from cycling based on their utility score. Understanding that travel time benefits are not the only criteria for mode selection, we apply a rejection sampling algorithm based on demographic factors to demonstrate that a more realistic, or pragmatic, cycling potential for Cape Town is in the region of 8%. The results also show that more than 46% of the observed pragmatic demand for cycling is concentrated in an area covering less than 7% of the study area. This has practical implications for policymakers to target interventions both in space and towards specific demographic market segments.

*Keywords:* Starter cycling cities, market segmentation, agent-based models, sustainable transport

---

## 1. Introduction

Cycling is increasingly promoted as a policy intervention for sustainable transportation in many cities worldwide. Benefits of cycling include that it does not produce emissions, has substantial health benefits as an active mode, consumes the lowest energy use per kilometre travelled, has a meagre cost, allows for greater social inclusion among communities with limited access to motorised transport, is faster than

---

\*This paper is published. Please cite as follows: Hitge, G., Joubert, J.W. (2024). The survivability of cycling in a co-evolutionary agent-based model. *Transportation*, forthcoming. DOI: [10.1007/s11116-023-10422-z](https://doi.org/10.1007/s11116-023-10422-z).

\*Corresponding author

Email addresses: [gerhardhitge@mweb.co.za](mailto:gerhardhitge@mweb.co.za) (Gerhard Hitge), [johan.joubert@kuleuven.be](mailto:johan.joubert@kuleuven.be) (Johan W. Joubert)

<sup>1</sup>Tel: +27 82 372 2730 (G. Hitge)

<sup>2</sup>Tel: +32 472 57 6264 (J.W. Joubert)

motorised modes over short distances and requires among the least space per person travelling (Blondiau et al., 2016; Copenhagen, 2010).

Many cities adopt policies and plans that declare their intention to promote cycling. Examples are often noted in the media where such intent was followed up with programmes and projects to provide infrastructure and facilities for cycling, but where many such programmes have made very little if any impact on cycling numbers, even after investment over many years (City of Cape Town, 2017). While this could be interpreted as an absence of potential cyclists in that city, it is more likely that interventions such as infrastructure provision were inadequate to address the critical barriers to potential cyclists in the community (Department for Transport, 2020).

Researchers and practitioners have done significant work to estimate the potential for cycling at regional, city and local levels. These methodologies consider a range of factors that affect the propensity to cycle, even though they typically centre on providing a network of cycling infrastructure. Most use zonal-based data for population (origin) and land use (destination). Cycling potential is often defined in nominal terms based on the proximity of a target population to zones with high levels of attraction, based on limiting criteria for the natural (gradients) and built (street network) environments (Fonseca, Ribeiro, and Neiva, Fonseca et al.; Lovelace et al., 2017; Zhang, Magalhães, and Wang, Zhang et al.).

This paper introduces a new methodology to quantitatively define and estimate potential cycling demand for a city, using the agent and activity-based model (ABM), MATSim (Horni et al., 2016). ABM provides several advantages over zonal-based methods for estimating cycling potential: i) it directly compares the utility of cycling for each individual compared to the utility achieved for the same travel plan with his or her prevailing modes; ii) all activities and trip legs can be considered when testing the viability of cycling, rather than only the primary AM peak activity, and iii) the travel decisions and behaviours of all other persons impact the attractiveness of cycling for each person. We define cycling potential as the number of people who would benefit from using a bicycle for their daily trip compared to their prevailing modes. We assert that persons would benefit from adopting cycling when their Travel Utility by bicycle is similar to or higher than that of prevailing modes when executing the same travel plan. The travel utility calculation considers the utility value of participating in activities and the negative utility of travelling between utilities. This method accommodates preferred activity durations and travel speeds between different modes.

We explore two scenarios for cycling potential. The first looks at utility only, assuming that all barriers to cycling had been addressed for the entire population. This provides a best-case scenario to understand the full scope of cycling potential in a city (for conventional bicycles, not e-bikes). The second scenario demonstrates a more pragmatic approach that recognises that demographic, natural and built environment factors are likely to reduce the propensity to cycle for some people – at least in the short to medium term. The second scenario highlights the opportunity to reach a subset of potential cyclists through narrowly targeted interventions to address cycling barriers for selected market segments and geographies.

## **2. Background to estimating cycling potential**

Research has shown that a relatively small set of factors significantly impact the propensity to cycle. Understanding these enables an appreciation of the definition of potential and the need for various interventions to unlock such potential. The literature also demonstrates the techniques and methodologies already available to estimate the potential for cycling from various perspectives.

### *2.1. Factors affecting the propensity to cycle*

Several factors affect the likelihood that someone would choose to cycle for daily utility travel. These have been categorised as demographic factors, including behavioural characteristics, the natural environ-

ment and the built environment (Heinen et al., 2010).

Market segmentation based on *demographics* is a recognised method to associate travel behaviour traits with personal attributes (Damant-Sirois and El-Geneidy, 2015). Most research found that **Age** has a strong correlation with cycling (Goldsmith, 1992; O'Connor and Brown, 2010), even though Wardman et al. (2007) found no correlation. Teenagers and young adults are typically the groups that cycle the most, whereafter cycling rates decrease with an increase in age (Copenhagen, 2010; Wooliscroft and Ganglmair-Wooliscroft, 2014).

While **income** is not strongly correlated with cycling (Heinen et al., 2010), car ownership and usage that is highly correlated with income (de D. Ortúzar and Willumsen, 1994). Behrens et al (2007) found that car drivers and passengers are less likely to shift between modes; in contrast, public transport users are likelier to demonstrate churn between modes. Irlam and Zuidgeest (2018) found that affordability can hinder bicycle ownership among low-income communities. Dill and Rose (2012) observed that electric bicycles, or e-bikes, are increasingly popular among the elderly, for whom it compensates for the loss in muscle strength. However, the cost of e-bikes currently restricts their use to higher-income individuals.

**Gender** was found to be a significant factor in cycling behaviour in many countries and cultures around the world (Taylor et al., 2015). While women cycle much less than men in virtually all reported cases, notable exceptions occur in countries such as the Netherlands and Denmark, where a strong cycling culture exists (Pucher and Buehler, 2008; Aldred et al., 2016). That women cycle less than men in cities where cycling is not a significant mode is attributed to the much higher role of caring for children and being more vulnerable when travelling (Sweet and Kanaroglou, 2016; Reid and Konrad, 2004).

Nello-Deakin and Nikolaeva (2021) introduced a novel factor associated with increased cycling. They describe that some people are likely to take up cycling once cycling reaches adequate levels to create the perception of acceptability and safety as Human Infrastructure.

Factors about the *natural environment* that impact the propensity to cycle include gradient (hilliness), temperature, rain and wind. Gradients up to about 4% are considered ideal for cycling, while gradients above between 5 and 8% are not conducive to cycling (Rodriguez and Joo, 2004; Midgley, 2011). The range is impacted by the factors such as age, fitness and distance over which the gradient applies. Indications are that temperatures below 5° in colder climates and above 30° in warmer climates reduce the propensity to cycle for many people. Strong wind above 10km/h and heavy rain or snow also reduce cycling, especially when these effects are combined (Saneinejad et al., 2012; Goldman and Wessel, 2021). Factors contributing to cyclists' resilience to continued cycling include having breaks that typically overlap with the height of extreme winter or summer conditions and being able to wait out heavy rain due to more flexible time schedules.

*Built environment* factors that increase the likelihood of cycling include higher urban density, a greater mix of land uses and a more dense road network (Ekblad, Svensson, and Koglin, Ekblad et al.; Cervero, Sarmiento, Jacoby, Gomez, and Neiman, Cervero et al.). These factors improve the directness of travel and reduce trip distance. The presence and type of cycling infrastructure significantly impact the propensity to cycle (Young et al., 2020; Buehler and Dill, Buehler and Dill). Providing cycling infrastructure physically segregated from road traffic is arguably the most critical intervention to unlock cycling potential (Buehler and Dill, Buehler and Dill). However, infrastructure is necessary but not sufficient to unlock cycling, especially when poorly designed and constructed (Tortosa, Lovelace, Heinen, and Mann, Tortosa et al.; Jennings et al., 2017). Other types of intervention that have been shown to contribute to increased cycling include bicycle parking, awareness and marketing, bike sharing schemes, intersection safety measures and integrating bicycling with public transport facilities, among others (Dias, Lopez, and Silva, Dias et al.; Hunt and Abraham, 2007). It is often difficult to safely store bicycles in high-rise buildings with small residential

units and informal settlements (Irlam and Zuidgeest, 2018). **Housing type** is also a factor to consider when promoting cycling, especially in the developing world with high rates of informal housing.

Dill and Rose (2012) observed that electric bicycles, or e-bikes, are increasingly popular among the elderly, for whom it compensates for the loss in muscle strength. However, the cost of e-bikes currently restricts their use to higher-income individuals. Cargo bikes may enable many women to transport children and goods as a possible intervention for the impact of dependency (Riggs, 2016). However, As with e-bikes, cargo bikes' cost prohibits them from being adopted widely among lower and middle-income households.

Several studies show that most utility cycling trips are below about 5km, with the largest proportion below 2km (DTU, 2015; Chillón et al., 2016). GTZ (2003) shows that cycling is attractive for up to 10km distances in mature cycling cities. Cycling competes with motorised modes over short distances before personal effort and lower speeds reduce its appeal as a utility mode. However, the range over which cycling can be competitive is extended during peak periods as it is typically not subject to the same levels of congestion experienced by traffic or waiting and transfer times associated with public transport travel.

Cycling potential would be higher for and within a city where the parameters of the natural environment, built environment and population demographic favour the propensity to cycle. For instance, cities with moderate temperatures and gentle gradients are more likely to be attractive to cyclists. Higher-density mixed land-use areas with safe cycle lanes are more likely to attract cycling than other parts of the same city. Populations with a higher proportion of younger people, especially students or higher-income households with the ability to access e-bikes, have a higher potential for cycling than other scenarios, all else being equal.

## 2.2. Methods to estimate cycling potential

The growing interest in promoting cycling as an urban utility mode has led to a growing body of literature on methods to estimate cycling demand in various parts of the world. Examples range from academic research to practitioner-led models and from purely methodological to planning support tools (Lovelace et al., 2017).

*Scale:* Research and measures to estimate cycling potential have been developed nationally, for example, for England and Wales (Parkin et al., 2008; Lovelace et al., 2017). While many studies focus on city-wide estimations (Fonseca, Ribeiro, and Neiva, Fonseca et al.; Young et al., 2020; Hitge and Joubert, 2021), approaches that focus on nodes or network segments within a city include work from Mahfouz, Arcaute, and Lovelace (Mahfouz et al.); Kuzmyak, Walters, Bradley, and Kockelman (Kuzmyak et al.).

*Measures:* In line with the three categories that influence the propensity to cycle (Demographics, Natural and Built Environment), cycling potential research mostly follows area-based, population-based or route-based approaches (Lovelace et al., 2017). However, natural environment factors such as rainfall were included (Parkin et al., 2008), while Silva et al. (2019) included a factor to test how Political Commitment influence cycling potential in a city. Area-based factors include slopes, gradient or hilliness, street hierarchy, the geometry of the street network and land use mix, or proximity to activities (Lovelace et al., 2017; Fonseca, Ribeiro, and Neiva, Fonseca et al.). Population-based factors include age, population density, car ownership, and household income (Fonseca, Ribeiro, and Neiva, Fonseca et al.; Hitge and Joubert, 2021). Route-based approaches include desire lines between OD pairs, or cycle network permeability and contrasting directness with quietness of the route (Mahfouz, Arcaute, and Lovelace, Mahfouz et al.).

*Methods:* In line with conventional transportation research methods, most research uses data from Administrative or Traffic Analysis Zones. Population, land use and travel data are typically aggregated into zones from Origin and Destination (OD) pairs and are derived (Lovelace et al., 2017; Silva et al., 2019). Hitge and Joubert (2021) use a synthetic population that includes travel plans for individual household members at the individual property level.

In most cases, potential estimation is based on proximity and regression models that employ distance decay functions (Lovelace et al., 2017; Olmos et al., 2020), and demographic exclusion factors (Fonseca, Ribeiro, and Neiva, Fonseca et al.; Hitge and Joubert, 2021). Several studies recognise the limitation of using historical survey data only and that new sources of OD data and advances in computational capabilities may benefit projections of cycling potential (Lovelace et al., 2017; Mahfouz, Arcaute, and Lovelace, Mahfouz et al.).

*Scenarios:* The research recognises that the propensity to cycle is influenced by various factors that enable and deter individuals from cycling, as discussed above. Research methodologies often assess potential based on scenarios that address one or more enablers or deterrents to cycling. Olmos et al. (2020) calibrated their model with existing cycling data, which they can use to extrapolate potential demand when similar conditions are created. Lovelace et al. (2017) use four scenarios that i) estimate the required interventions to double current cycling demand, ii) estimate potential when women cycle as much as men, iii) a Go-Dutch scenario if their study area population exhibited the same behaviour as the Dutch, who have among the highest uptake of cycling in the world, and iv) any e-bike scenario, for which, among others, hilliness plays a minor role that for conventional bicycling.

*Results / Outputs:* The majority of research produces quantitative outputs in that it shows where cycling potential is likely to exist within a city or region (Lovelace et al., 2017; Young et al., 2020) and which cycle network elements to improve to unlock this potential Mahfouz, Arcaute, and Lovelace (Mahfouz et al.); Fonseca, Ribeiro, and Neiva (Fonseca et al.).

Goodman, Rojas, Woodcock, Aldred, Berkoff, Morgan, Abbas, and Lovelace (Goodman et al.) provided a quantitative assessment that showed the percentage of children in the UK cycling to school might increase from 1.8% to 41% if Dutch levels of cycling could be achieved. Hitge and Joubert (2021) found that up to 15% of pre-defined demographic groups in Cape Town live within a comfortable cycling range of destinations they are likely to attend.

### 3. Methodology and case study

#### 3.1. Methodology

This paper proposes a travel utility-based methodology to estimate potential cycling demand in a city. We assert that potential cyclists would be similar or better off cycling than using their existing modes to perform their daily travel plan. We measure the utility of travel in the Agent-Based Model (ABM) MATSim (Horni et al., 2016). The key feature of this approach is that cycling's competitiveness is measured to be relative to other modes while using a congested road network.

A base scenario is created that measure travel utility for the population based on information obtained from a travel survey, consisting of activity chains and modes serving each trip between consecutive activities. For each person's travel plan, MATSim calculates a utility score, where the score increases when participating in productive activities and decreases while travelling. For the utility *score*, a generalised cost that accounts for both time and monetary expenses, this paper uses the default MATSim parameters from Charypar and Nagel (2005), which translates into marginal utilities of 20€/h for performing an activity, -12€/h for travel time, -6€/h for waiting and -18€/h for arriving too late at one's activity. The result is that the overall score reported for the scenarios could be interpreted as the *experienced value (in €) of the daily activities*. We neither converted these parameters from Euro-based values to South African Rand (ZAR) nor calibrated the model to account for different travel behaviour in South Africa. Consequently, when interpreting the reported results, evaluating the cycling scenario *relative to* the base scenario is essential, and not interpreting the absolute values.

Like other methods, we test cycling potential for different scenarios [Lovelace et al. \(2017\)](#). In the first cycling scenario, we enable each member of the population to select a bicycle as an alternative to their existing mode or mode mix. Bicycles are typically competitive according to the range defined by a trip distribution for cyclists. However, the travel time competitiveness of a bicycle is relative to the travel time in other modes, which may or may not be subjected to congestion or waiting time in the case of transit. Through its co-evolutionary algorithm, where cycling's performance is co-dependent on the mode choice of all other persons that travel, MATSim retains cycling as the preferred mode for those for whom cycling yields a higher utility score. Since all other model parameters are kept the same, the higher utility score reflects a travel time advantage of cycling over prevailing modes.

It is accepted that travel time is not the only criterion for mode selection, and not everyone with a travel time benefit would switch to cycling. We create a second cycling scenario that employs a rejection-sampling algorithm based on a set of demographic characteristics that have been shown to influence individuals' propensity to cycle. This provides a more pragmatic potential for cycling in the short to medium term, especially for a Starter Cycling City, where a cycling culture is yet to be established.

Finally, we demonstrate how the methodology could be applied to test specific policy questions. In this third cycling scenario, we test the potential for cycling among children to cycle to school, assuming they attend the nearest suitable school.

The outcomes assume that all enabling measures have been provided and that all barriers to cycling had been removed for the entire population. For this study we use the parameters of a pedal bicycle without electrical assistance. The impact that e-bikes may have on the potential for cycling could be tested as a new scenario, as these have been shown to extend the range and reduce the impact of gradient on cycling ([Van Cauwenberg et al., 2019](#)). The model outputs indicate where trip destinations are concentrated as an indication of higher potential demand for cycling. Policymakers can use this information to target interventions spatially to create an environment conducive to cycling, which is likely to enable the potential cyclists to make the shift.

### 3.2. Approach and main contribution

Cape Town, South Africa, was selected as the case study. The municipality declared its intention to increase the share of people cycling to work from less than 1% in 2016 to 8% by 2030 ([City of Cape Town, 2017](#)). In addition, a MATSim model exists for the city, complete with publicly available data that includes a detailed synthetic population ([City of Cape Town, 2013](#); [Joubert, 2021](#)).

Section 5 describes the results emanating from the modelling process, while Section 6 discusses the policy implications of the results and suggests further research to expand this knowledge.

The paper is structured as follows. The next section describes the attributes of the base scenario of the model that is used to determine the baseline utility for each person's travel plan using their prevailing modes. Section 4 describes the two *cycling* scenarios, defined as *optimistic potential* and *pragmatic potential*. Section 5 describes the results emanating from the modelling process, while Section 6 discusses the policy implications of the results and suggests further research to expand this knowledge.

## 4. Building a baseline scenario in MATSim

This section provides an overview of MATSim ([Horni et al., 2016](#)), the agent-based modelling platform used for the analysis. The outline of the simulation process can be described using Figure 1. Two main inputs are required to model a scenario as *initial demand*. Firstly, the transport *supply* in the form of the road network (Section 4.1) and, if available, the detailed transit network in the form of routes and schedule.

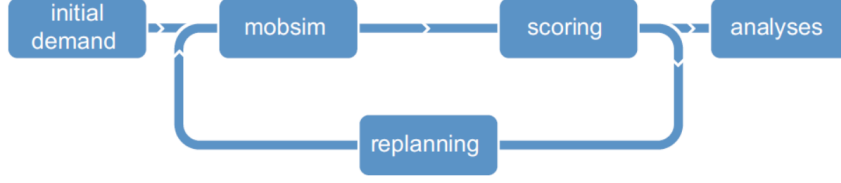


Figure 1: The overall simulation process in MATSim.

The second input is the *travel demand*, addressed here in two sections: the synthetic population (Section 4.2) and its activity-based travel demand (Section 4.3).

#### 4.1. Transport network

At a minimum, the transport system in MATSim is described in terms of the road infrastructure or network. The Cape Town road network is extracted from *OpenStreetMap* (OSM, 2021) as a series of interconnected links joined at nodes. Links have inferred capacity and speed parameters that are used during the simulation to determine the travel time delay for people travelling by car, either as the driver (mode car) or as a passenger (mode ride).

Without publicly available and curated transit data, transit trips are not explicitly modelled on the (road or rail) network. Still, they are teleported between activity locations, with the travel time being a function of the specific modes' speed, access and egress times. Hitge and Vanderschuren (2015) reported average walking speed and typical in-vehicle travel speeds for different transit modes in Cape Town. They also said the distance to access the main transit modes from home. Average travel speed by bicycle varies from 12 to 16 km/h for ordinary bikes and higher for e-bikes (KiM, 2019; Saneinejad et al., 2012). A cycling speed of 15 km/h is used in this study.

For this paper, we break the transit trip into two parts: 1) the access and egress walk, denoted by  $a_m$  for mode  $m$ , and 2) the in-vehicle travel time, denoted by  $v_m$ . The parameters for the different modes are reported in Figure 2 along with a graph that indicates the different modes' performance (total travel time) over various distances. To calculate the travel time between two consecutive activities, we need the total

Mode $m$	Access & egress walk time (min) $a_m$	In-vehicle speed (km/h) $v_m$
Cycling	–	15.0
Bus	22	21.5
BRT	15	30.0
Train	29	22.6
Other	10	10.0

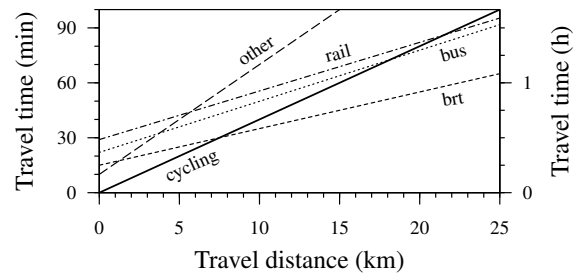


Figure 2: Mode specific parameters (left) to estimated total travel time (adapted from (Hitge and Vanderschuren, 2015)). The total travel time comparison of the teleported modes is compared in the graph on the right.

distance,  $d$ , expressed in kilometres (km). The travel time by mode  $m$ , denoted by  $t_m$  and expressed in minutes, is then calculated using (1).

$$t_m = a_m + \frac{d}{v_m} \times 60 \quad (1)$$

The distance to and from the rail is typically longer than to and from a bus, which, in turn, is longer than access to and from *minibus taxis*, the dominant paratransit mode in South Africa (StatsSA, 2013). Having different operating entities, this paper distinguishes between the more traditional bus service provided by *Golden Arrow Bus Services* and the municipal *MyCiTi* Bus Rapid Transit (BRT) service.

#### 4.2. Population

Due to the sensitive nature of personal data, accurate, detailed population data is frequently not readily available. At a fine-grained spatial level, census data is usually aggregated to some zonal level. Consequently, you may know that there are 425 females in a zone and 143 persons aged 25–30. Also, you may know that 765 persons are black African, but you do not know how many 28-year-old black African females are in a particular zone. To partially address the problem, authorities make Public Use Micro Samples (PUMS) of the census data available, representing the detailed records of a subset of persons. However, the spatial granularity of the PUMS is much coarser to ensure data privacy. In South Africa, the PUMS represents 10% of census records.

To generate a high-resolution population with detailed household structure, demographic and socioeconomic attributes, researchers turn to a process called *population synthesis*, which focuses on fitting a contingency table constructed from the micro samples to marginal constraints from aggregated census data (Sun and Erath, 2015). This paper benefits from a publicly available synthetic population developed for Cape Town that is controlled at the household level using (household) income, and at individual levels using gender and population group characteristics. The result provides a complete stock of individuals while accounting for detailed demographic, socioeconomic information, and household structure (Joubert, 2021).

Each household in the original synthetic population has a home coordinate that is randomly scattered in the *sub place*, the lowest level of spatial demarcation available in the census data. However, this paper benefits from a detailed parcel data set for Cape Town. Since one of the household attributes in the synthetic population is the main dwelling type in which the household resides, we adjust the household’s home coordinate to that of a sampled parcel. Some land use types allow multiple households while others can only accommodate a single household, in which case the parcel is removed after being sampled.

#### 4.3. Travel demand

The synthetic population includes socioeconomic and demographic attributes of individuals and their households. This section describes the process of assigning activity-based travel demand to each individual. Travel demand here refers to a detailed sequence of activities and trips, referred to as a person’s daily *travel plan*. Travel plans are derived from a comprehensive Household Travel Survey (HHTS) conducted for Cape Town in 2013 (City of Cape Town, 2013). More specifically, this paper relies on those households and individuals that also completed the travel diary. For each activity in a person’s reported diary, we know the start and end time, activity type, and location, which is expressed as a transport analysis zone (TAZ). In this paper, we distinguish between *primary* and *secondary* activity types. Primary activities include home (denoted by *h*), primary and secondary education (*e1*), tertiary education (*e2*) and work (*w*). Secondary activities include shopping (*s*), leisure (*l*), accompanying another person to their primary or secondary education activity (*e3*), visiting another person (*v*), medical and health-related activities (*m*) and all other activities (*o*).

For each trip in the diary, we know the mode of travel. Here we distinguish between walking (denoted by *walk*), private car as a driver (*car*), private car as a passenger (*ride*), bus (*bus*), BRT (*brt*), train (*train*, minibus taxi (*taxi*) and all other modes (*other*). Cycling was not reported in the travel diary, and we acknowledge that *brt* is likely underrepresented because the *MyCiTi* service was in its infancy at the time of the survey.



Of the 9 248 respondents, 4 737 (51.2%) reported travelling to at least one activity away from home, while 4 511 (48.8%) reported no travel in the reporting period. Approximately 300 unique travel plan configurations were extracted from the survey data. For example, if a person travelled from home to work and then stopped at a shop to buy groceries before returning home, the travel plan could be expressed as h-w-s-h. The seven plans shown in Table 1 make up more than 75% of all plans of the base population, with a further 27 plans contributing to more than 90% of all reported plans. One example of the 167 plans

Table 1: Seven plans occurring 75% of the time.

Travel plan	Reported instances	Percentage of total	Cumulative percentage
h-w-h	1 346	27.5%	27.5%
h-e1-h	812	16.6%	44.1%
h-s-h	729	14.9%	59.0%
h-o-w-o-h	255	5.2%	64.2%
h-v-h	197	4.0%	68.2%
h-o-h	172	3.5%	71.7%
h-l-h	166	3.4%	75.1%

that were only reported once is h-w-o-h-v-o-h. For this paper, we removed chains with more than five activities. Of those remaining, the 95% most popular (highest number of occurrences) chains were used to sample from.

Respondents to the travel survey reported socioeconomic and demographic characteristics at the household and individual levels. This paper used these attributes to sample a representative travel plan from the travel diary for each person in the synthetic population. The assumption (discussed above) is that people with similar characteristics will demonstrate similar travel behaviour. We find the  $N$  geographically closest survey respondents with similar socioeconomic and demographic attributes for each person in the synthetic population. A random respondent is picked from these, and a travel plan is assigned to the person in the synthetic population. In future work, this approach may be enhanced to instead depend on a more formally estimated behaviour model. One such behavioural model using Bayesian networks to synthesise activity-based travel plans is by [Joubert and de Waal \(2020\)](#).

Table 2 shows that 52.8% of the synthetic population received travel plans compared to the 51.2% reported in the travel survey.

Table 2: Number of Agents travelling in base scenario.

Category	Number of agents	Percentage of population
All agents in scenario population	375 951	
Travelling agents	198 441	52.8%
Non-travelling agents	177 510	47.2%

The travel plan assigned to each person still reflects the original activity locations. The next step is to relocate each activity so travel distances remain aligned with the sampled activity chain. First, we relocate

all home activities to the same coordinate as the synthetic person’s household home. All other activity locations rely on facilities parsed again from *OpenStreetMap*. As part of this research, we explicitly captured the level of education for all facilities in the study area according to the International Standard Classification of Education (ISCED) levels (UNESCO, 2012). Since schools (e1) are relatively uniformly spaced across a city, we assume that a scholar (their parents) will prefer a school closer to their home location.

Without reliable workplace data, we performed a double-constrained matrix balancing to not only rely on trip distance from the travel survey. Once the primary activities (home, work and education) are placed, secondary activity locations between primary activities are sampled in the ellipsoid anchored by the two primary activities.

Finally, we describe the simulation process. This includes simulating the initial plan, which (1) sees each agent depart from the prevailing activity location at a specified time using a given mode and travelling to the allocated destination to participate in its planned next activity; (2) describes how this iteration is scored; that some agents get to replan and what the conditions are for replanning; how the scores are assessed after each iteration until the total score of all plans converges, signifying a stabilisation of the given land-use and transport system.

#### 4.4. Simulation and scoring

The initial demand data is provided to the MATSim scenario as input. The first action in the iterative MATSim run is the mobility simulation (Figure 1). At the start of an iteration, each person selects one travel plan from their memory, noting that, initially, each agent only has a single plan. The better a plan performed in a prior iteration, the higher the likelihood that an agent will select it again in the next iteration. This principle is supported by the Habit Formation Theory of travel behaviour, which suggests that a person gravitates to decisions that have a positive outcome (Adjei and Behrens, 2010). All agents then execute their travel plan on the network, tending to the activities as stipulated in the plan and moving between them using the specified mode. The travel times for the car and ride modes are derived from the experienced congestion on the network. For other modes, travel times are derived from Figure 2.

Using only an average speed may be adequate for comparing transit time with other motorised modes. However, a bicycle moves much further than a person walking to, waiting for and transferring between transit lines, but is eventually overtaken by higher in-vehicle speeds of transit modes. The fixed time component of access and egress times allows this current paper to determine for which trips cycling holds a time advantage.

After the mobility simulation, each agent scores its plan as the sum of all activity utilities plus the (negative) sum of all travel *dis*utilities. Scores increase when activities are performed above a threshold duration, while scores decrease for travel based on the generalised cost. In this current research, *generalised cost* only accounts for the time component of travel and not the monetary cost associated with owning and operating a vehicle, bicycle, or transit fares. All plans are scored in every iteration, which means that even the same plan could score differently due to changes in other agents’ plans.

At the end of the scoring phase, some agents can adjust their travel plans for the next iteration. For the configuration used in this study, 70% of the agents picked a plan from their memory. Agents can save up to five plans to select for each iteration. Once replanning results in five plans in the agent’s basket, the lowest-scoring plan is removed and unavailable for future selection. The selected plan is not necessarily the best based on the selection strategy, and the default selection strategy uses a higher probability of choosing a better scoring plan. Some 15% of agents can pick a plan and adjust the departure times to improve their scores. The balance of agents picks a plan and is allowed to reroute in search of less congested routes. This method satisfies the principle of churn, where individuals make daily variations in their real-world trips,

such as switching between different modes, routes or departure times (Chatterjee, 2001; Saleh and Farrell, 2007).

Each agent then executes its newly selected plan, and the process repeats. The autonomy of each agent allows for a co-evolutionary search where the actions of other agents, along with the agent's own choice, impact future performance. The combined utility of all agents' plans stabilises after 200 iterations as no agent can consistently improve their own plans' performance using the available innovation strategies.

## 5. The cycling scenarios

This section describes the MATSim scenarios where cycling is introduced as a potential mode to a population subset. It shows which of these agents' cycling survives by yielding similar or higher utility scores than their prevailing modes. The population characteristics, transport network and activity locations remain the same as in the base scenario.

### 5.1. *Optimistic scenario: Everyone can test cycling*

The *optimistic scenario* tests the number of people that could benefit from cycling, purely based on their cycling utility compared to their existing modes and provides an indication of the upper limit of cycling potential for the population's prevailing activity locations and travel plans. It is deemed optimistic since people choose modes on many more criteria than travel time considerations. Therefore, not everyone that may benefit will likely shift to cycling - at least not in the short term.

For this scenario, all agents retain their highest scoring plan from the base scenario as the initial demand of the cycling scenario. All Agents are "handed" an additional *cycling plan*, defined by replacing all modes in their highest scoring plan with the bicycle as mode. This plan retains the activities, locations, start times, and route choices. To simplify the model, we exclude multi-modal trips with a bicycle, such as cycling as a feeder to BRT or rail. The cycling plan is selected by all agents in the first iteration of the Cycling scenario and scored according to the same configuration rules as before. The results for both scenarios are demonstrated in Figure 3.

### 5.2. *Pragmatic scenario: Those more likely to cycle only*

The literature described in Section 2 shows that demographic characteristics such as age and income are typically correlated with the propensity to cycle. In this scenario, we apply a rejection-sampling algorithm based on five demographic characteristics that have been shown to influence individuals' propensity to cycle. Alternative scenarios may be constructed to test only the potential for cycling among specific market segments, for specific activities, or selected origin or destination areas. This scenario applies a variation of the profile of potential cyclists described for Cape Town (a Starter Cycling City) by Hitge and Joubert (2021). It demonstrates, among other things, the different outputs between the proximity-based approach to estimating cycling potential and a travel plan-based approach of this paper.

While the scenario is believed to be plausible and, therefore, suitable to demonstrate the methodology, it does not claim to be probable. To do so would require calibration of the parameter values and possible inclusion of new parameters that may have a stronger correlation with the propensity to cycling for some or all community segments in a particular city. The authors assert that creating a desktop demographic profile at a relatively low cost enables a constructive discussion with policymakers. Promising high-level results are more likely to convince policymakers to collect relevant data for implementation.

### 5.2.1. Age

Data from Copenhagen shows that persons aged 18 to 25 cycle the most, with the age group 26 to 35 a little less. While a meaningful proportion of people cycle in the age group from 36 to 59, cycling declines fast for higher ages (Copenhagen, 2010). It is assumed that persons are more likely to try and shift to cycling at a younger age.

For this study, we assign a likelihood to cycle within age brackets because the probability declines. We assume all high school learners (ages 15 to 19) and young adults up to 25 are eligible cyclists. We presume that 80% of people aged 26 to 35 would cycle and that 40% of people aged 36 to 60 would shift to cycling if conditions were favourable. It is doubtful that people over 60 would start cycling in a starter cycling city such as Cape Town.

Cape Town's population in 2011 consisted of 25% persons in the age group under 15, 8% persons in the age group 15 to 19, 30% from 20 to 34, 29% from 35 to 64 and 8% above 64 years of age (StatsSA, 2012b). This means that at least 34% of the population will not be considered potential cyclists for this study based on age alone.

### 5.2.2. Household income

While the literature does not describe a clear correlation between cycling and income (in mostly first-world cities), we assert that household income (HH income) affects the likelihood of shifting to cycling in two ways: firstly, in terms of the affordability of a bicycle and related equipment and secondly in terms of attractiveness of cycling compared to exiting modes. For instance, while high-income persons could afford suitable bicycles, even electric bicycles, many would not be willing to forgo the comfort and usefulness of a car as the marginal monetary cost per trip is negligible to them. Middle to higher-income workers are also more likely to use their vehicles for business and recreational trips. On the other hand, members from lower-income households may benefit significantly from a bicycle's meagre running cost. Still, they may need help to afford the purchase and maintenance costs.

While several programmes exist to distribute subsidised or free bicycles to communities where affordability is a barrier to cycling (National Department of Transport, 2020; BEN, 2020; Qhubeka, 2020), this intervention would benefit from first providing adequate infrastructure. This serves as an example of investments in the public realm vs that aimed at individuals and that the latter should only be considered once the former has reached adequate levels.

In the absence of data, Table 3 shows the impact that household income is expected to have on the likelihood that a person would cycle. The *Income factor* is the product of the *Affordability* and *Attractiveness* factors.

Table 3: HH Income influence of cycling potential.

Income category	Affordability	Attractiveness	Income factor
High	1.0	0.4	0.4
Middle-High	1.0	0.9	0.9
Low-Middle	0.8	1.0	0.8
Low	0.6	1.0	0.6

Only 14% of the Cape Town population was in the High and High-Middle income categories in 2011. While 39% were in the Low-Middle category, a substantial 47% were in the Low-income category, including the majority of unemployed (StatsSA, 2012a).

### 5.2.3. Gender

In Copenhagen, considered a gender- and income-equal society, roughly the same number of women and men cycle, with slightly more women than men (53% : 47%) (Copenhagen, 2010). South Africa is regarded as a patriarchal society where the roles of men and women mirror those described above more closely than in Scandinavian countries. While there is no legal constraint against women cycling, as there are in some countries, many women suffer restrictions on their travel decisions through cultural and social norms. Similar cultural and social norms result in men experiencing more peer pressure to obtain and use a car. It could therefore be expected that the uptake of cycling among women would lag behind that of men in an Starter Cycling City (SCC).

If the study's premise is to explore the potential market for cycling once adequate interventions have been implemented, women would be as much, if not more likely, to cycle once the environment is conducive to cycling. We therefore test the effect where 90% of women and 80% of men over the age of 19 would cycle; assuming school children are not affected by gender-related factors.

### 5.2.4. Household composition

A household (HH) consists of a combination of interdependent persons that typically share specific tasks. One person can do the shopping for all members, so only some have to allow time for shopping. The e3 activity demonstrates the role of some person in the household to accompany others to school. Other examples include the function of care, typically performed by a non-working adult in the home and looking after small children and elderly parents or relatives. The dependency ratio for Cape Town was more than 43% (StatsSA, 2012b), which mainly consists of people aged below 15 and above 65 that are being looked after by those between the ages of 15 and 65.

Interdependence affects the travel decisions of all household members, including the likelihood of shifting to cycling. Anecdotal evidence among staff at the University of Cape Town showed that the need to escort children to school was a leading reason for choosing to use a car rather than public transport or cycling to work. Accordingly, persons in households without dependent persons have the greatest freedom to choose their modes. We assume that members from families where no members are below 15 or above 65 have a 100% chance of cycling, while those with dependents have a 50% chance. This excludes children aged 15 to 19 who are assumed to go to school and do not care for dependents.

The household size in Cape Town reduced from 3.8 persons per household in 1996 to 3.3 in 2011. The fact that population growth exceeded household growth and that the number of small families increased substantially could have significant implications for transport planning and the potential for cycling and is worth further investigation.

### 5.2.5. Dwelling type

There is an inverse correlation between household size and income, but a positive correlation between dwelling size and income (StatsSA, 2011). People living in informal settlements are much more likely to be unemployed and live in larger families with little room to store bicycles. We assume that only 20% of persons living in informal housing would be able to shift to cycling due to the difficulty in safely keeping a bike at home, and that this will not be a constraint on any other housing type.

## 5.3. Creating the cycling population

Only persons between the ages of 15 and 65 are considered potential cyclists for this scenario. The overall probability of being a cyclist is the product of the likelihood of all five factors we consider. For instance, a 28-year-old person has an 80% chance of being retained based on age alone. If that person lives in a Middle-High income household, they have a further 90% of being retained (or a 72% chance in total).

Should that household consist of more than two members, their likelihood of being retained in the cycling population halves. If they are kept at this point, 90% of women and 80% of men will be included. While all the remaining persons living in formal dwellings are retained, those living in an informal settlement will have only a 20% chance of inclusion in the final cycling population.

Table 4 shows the five elements we consider in this paper, the categories and factor values for each. The above criteria are applied to the base scenario’s synthetic population to determine the potential cyclists’ subpopulation.

Table 4: Diminishing factor values.

Diminishing factor		Factor values	
$j$	Description	Value	Description
1	Age	1.0	aged 15–25
		0.8	aged 26–35
		0.4	aged 36–60
		0.0	any other age
2	HH Income	0.4	High Income
		0.9	High-Middle Income
		0.8	Middle-Low Income
		0.4	Low Income
3	HH composition	1.0	No dependents in HH, or aged 15–19
		0.5	Dependents in HH, except aged 15–19
4	Gender	0.9	female
		0.8	male
5	Dwelling type	1.0	Formal housing
		0.2	Informal dwelling

The number of agents likely to cycle depends on the random seed value used in the above process. To avoid an unrealistic small or large cycling population, 1 000 different populations were generated, each using a unique random seed. The median percentage of individuals sampled as potential cyclists were 12.50%, with a 95% confidence level that the percentage is between 12.38% and 12.61%.

A population close to the distribution’s median was randomly selected for the cycling scenario as a highly likely and, therefore, representative subpopulation.

#### 5.4. Simulating the cycling scenarios

All agents retain their highest scoring plan from the base scenario as the initial demand of both cycling scenarios. The potential cyclists identified in the process above are “handed” an additional *cycling plan*, which is defined by replacing all modes in their highest scoring plan with the bicycle as mode. These plans retain the activities, locations, start time, and route choice, where applicable. To simplify the model, we exclude multi-modal trips with a bicycle, such as cycling as a feeder to BRT or rail. The cycling plan is selected by all affected agents in the first iteration of both the Optimistic and Pragmatic scenarios and scored according to the same configuration rules as before.

Figure 3 demonstrates how the number of cycling plans starts high, at 100% and 12.47% for the optimistic and pragmatic scenarios, respectively.

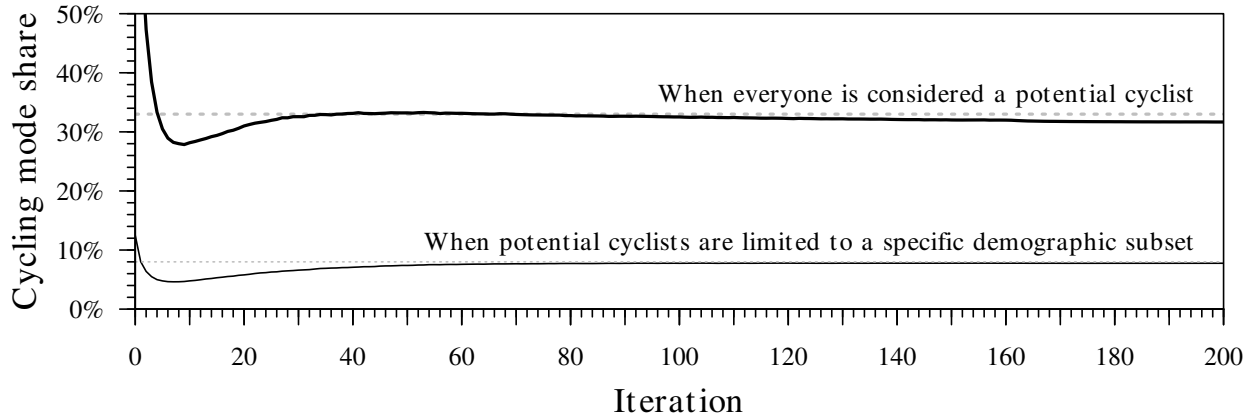


Figure 3: Convergence in the selection of bicycle plans as the number of model iterations increases.

It takes 7 to 10 iterations for all agents with cycling plans to have selected their original mode plan, which accounts for the sudden drop in the graph. After that, the number of cycling plans and replanned versions slowly increased as their scores consistently exceeded that of their base plans (as amended through replanning). The graph shows that the number of cycling plans stabilises at 31.7% and 7.8% for the two scenarios.

## 6. Results

In this section, we discuss the number of agents for whom cycling survived and their profile compared to the non-cycling agents. Table 5 shows the summarised results of the Cycling scenarios. Of the 1.98 million travelling agents, all were given a cycling plan in the *Optimistic scenario* and of these, cycling survived as a mode for 628 140 agents. 248 200 agents were given a cycling plan in the *Pragmatic scenario* and of these, cycling survived as a mode for 154 560 agents.

Table 5: Agents for whom cycling survives.

Category	Number of agents	% of agents
All travelling agents	1 984 410	100.0%
<b>Optimistic scenario</b>		
Agents with cycling plan	1 984 410	100.0%
Cycling survives ( <i>cyclists</i> )	628 140	31.7%
<b>Pragmatic scenario</b>		
Agents with cycling plan	248 200	12.5%
Cycling survives ( <i>cyclists</i> )	154 560	7.8%

Since we know the individual attributes of each person in the simulation, one can analyse the results more personally. The average and median ages of the population are 29.7 and 28 years, respectively, while

the median ages of cyclists in the Optimistic and Pragmatic scenarios are 25 and 22 years, respectively. The lower age of cyclists in the Pragmatic scenario can be expected based on the selection criteria shown in Table 4.

The average and median households incomes of the population are R226 377 ( $\approx$ US\$ 11 900 or 11 000€) and R76 800 ( $\approx$ US\$ 4 000 or 3 700€) per year, respectively. The median income of potential cyclists in the *optimistic* scenario is R38 400 ( $\approx$ US\$ 2 000 or 1 850€), or only 17% of the population. This may reflect the poor travel time experienced by lower-income household members in Cape Town, who are mostly dependent on public transport, for which the average travel time is more than double that of the car (Hitge and Vanderschuren, 2015). Given that the selection criteria for the *pragmatic* scenario mostly excludes persons from lower-income households, the median income of these potential cyclists is R153 600 ( $\approx$ US\$ 8 100 or 7 500€), or 62.9% of the population.

### 6.1. Compare scores between cyclists and non-cyclists

Figure 4a shows the comparison between the initial ( $x$ -axis) and final scores ( $y$ -axis) of the *pragmatic* scenario for all agents for whom cycling survived, the 154 560 *cyclists* in the final iteration. Each dot

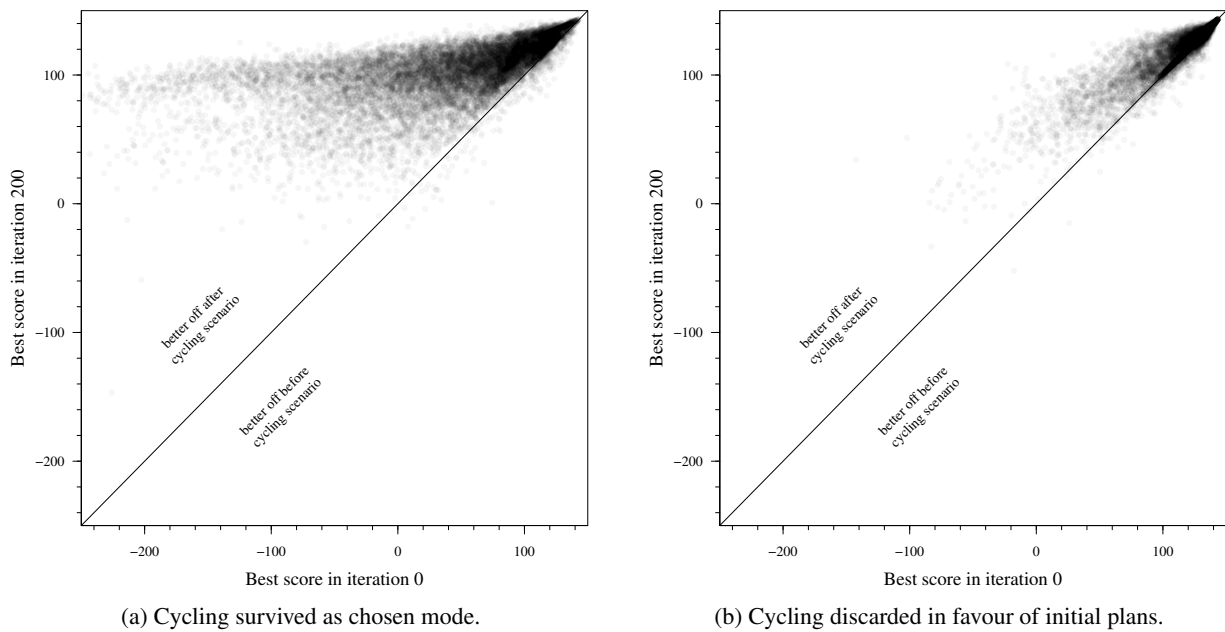


Figure 4: Best plan scores in *cycling* scenario.

represents one individual agent.

Figure 4b shows the same result for the 93 640 *non-cyclists*, or those with potential for whom cycling did not survive. The  $x$ -axis shows the original scores, which indicates many agents with scores below -200, the lighter cloud to the left of the graph. The  $y$ -axis shows that only some Agents (less than 1%) ended with scores below -100. Virtually all cyclists and non-cyclists ended with positive scores.

A score on the diagonal indicates no change, while below the diagonal indicates a worse score in the cycling scenario. The few scores below the diagonal could be attributed to the selection strategy. Some agents would have selected a marginally suboptimal plan from their basket in the final iteration. This is



expected in a co-evolutionary model environment which does not aim to reach an optimised state for the individual.

The data behind the graph provides important insight. After the 200<sup>th</sup> iteration of the *baseline* scenario, the median utility score was 93.1. This improved to 113.4 and 121.3 for the *pragmatic* and *optimistic* scenarios, respectively. The median scores for *non-cyclists* increased from 121.6 to 125.7 in the *pragmatic* scenario and from 103.7 to 125.2 in the *optimistic* scenario. *Potential cyclists* increased their median scores from 59.8 to 107.3 in the *pragmatic* scenario and from 70.7 to 111.5 in the *optimistic* scenario. It is deemed significant that the scores for *potential cyclists* is similar to that of *non-cyclists* before cycling is introduced. This suggests that cycling can potentially reduce transport poverty for many marginal communities.

For context: The default utility value for participation in an activity is +6, while the (dis)utility for travel is -10. Therefore, the theoretical maximum score is  $24 \times 6 = 144$  if an agent attends an activity for the entire day. Since we only calculate utility for those with away-from-home activities, the maximum cannot be reached due to the time travelling while away from home.

The marginally improved scores of non-cyclists could be attributed to the reduction in congestion from some *cyclists* that are no longer driving. The resultant lower travel times translate into a lower disutility for *non-cycling car* and *ride* agents. This effect, where a shift to cycling also benefits motorists, has been observed and described in several studies (FLOW, 2016; Nordic Council of Ministers, 2005). However, the effective roadway capacity created by reducing traffic demand could induce new trips (latent demand) since the decrease in travel time now reduces the marginal cost of travel for some users below their threshold to make new trips (OECD, 2007).

## 6.2. Modes cyclists shifted from

Assessment of the modes *potential cyclists* shifted from provides an opportunity to test the validity and plausibility of MATSim for the purpose it is applied to in the study. The data from the *optimistic* scenario is used for this analysis. Only trip chains selected by more than 100 agents were included in the analysis. These 54 mode combinations account for 99.3% of all agents' trips. Each permutation of a mode chain is unique. For example, car-walk-car is distinguished from car-car-walk.

It is not surprising that 41% of *potential cyclists* shifted from car-car, since this mode combination accounts for 47.3% of all mode combinations in the population. What is significant is to observe which mode combinations lost more trips to cycling. 20 mode combinations could not retain a single *potential cyclist*, while another two retained one each. These 22 mode combinations were selected by two agents from among 16 077, or 25.8% of all *potential cyclists*. Of these, only 1 648 included a car trip on one leg of the trip chain. Only 27.3% of the predominant car-car mode combination could not compete with cycling.

In the *pragmatic* scenario, only one agent with a walk leg in the mode combination did not benefit from cycling, or at least did not select cycling in the final iteration of the scenario.

## 6.3. Distribution of cycling destinations

Learning that between 7 and 33% of the travelling population (pragmatically and optimistically) are likely to benefit from cycling should encourage policymakers to design and prioritise interventions to unlock this potential. Knowing the geographic concentration of potential demand enables spatial targeting of various interventions for a more significant impact. Also knowing the profile of likely cyclists in geographic space could assist policymakers in designing appropriate targeted interventions to maximise early gains. For instance, the marketing strategy used for students may differ from that designed for office or industrial workers, even within the same geographic area.

Figure 5 shows the destinations with the greatest concentration of cycling trips for Cape Town's population, based on the *pragmatic* scenario.

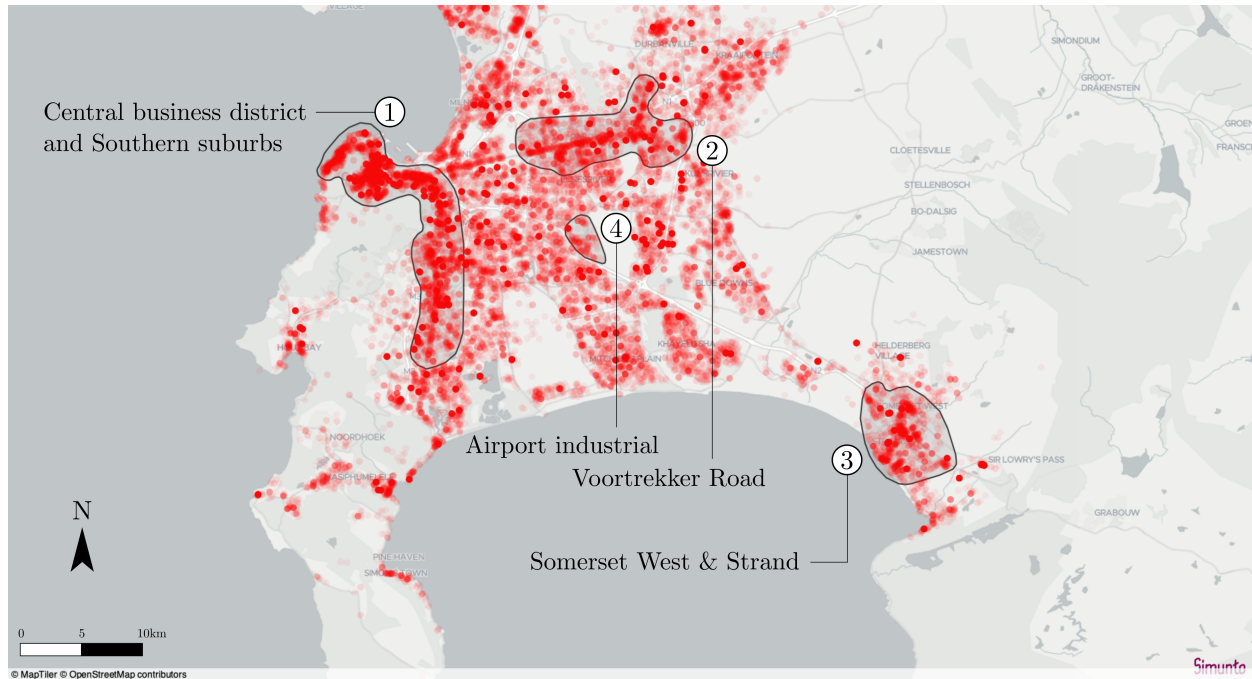


Figure 5: Map of popular destinations for Cyclists. Source: own image using *MapTiler*.

Table 6: Popular destinations for cyclists.

Destination	Area	All	Demographic
CBD and Southern suburbs	3.0%	31.8%	32.0%
Voortrekker Road	2.0%	8.7%	8.5%
Somerset West & Strand	1.6%	4.8%	5.5%
Airport industrial	0.3%	0.7%	0.5%
Total	6.9%	46.0%	46.5%

The map indicates that 46.5% of destinations (highlighted as polygons in Figure 5) fall within areas that represent only 6.9% of the city’s footprint. The area that represents the central business district and Southern suburbs along Main Road ① accounts for 32% of destinations; the dense business corridor along Voortrekker Road ② for 8.5%; the adjacent suburbs of Somerset West and Strand ③ for 5.5%. For comparison, we also included the industrial area adjacent to the Cape Town International Airport ④, which accounts for only 0.5% of destinations. The map shows where approximately 3.6% of the total travelling population could benefit from cycling if interventions were concentrated in these relatively small targeted areas. This outcome expands on the work of [Hitge and Joubert \(2021\)](#), who argue for a nodal approach to unlock cycling potential in a *starter cycling city* based on proximity to likely activities alone.

## 7. Discussion

This paper demonstrates how an activity-based transport model can be applied to estimate the size and spatial distribution of potential city cyclists, even without reported cycling usage. The methodology holds

particular promise for Starter Cycling Cities where decision-makers have little or no data that indicates if, how much and where the potential for the uptake of cycling may exist. The authors believe such information may contribute significantly to the willingness to invest scarce resources to unlock the benefits of cycling and move towards a more sustainable transportation system.

The optimistic scenario shows a probable maximum cycling potential for the existing population's prevailing land use distribution and travel plans. While this may be viewed as an ultimate longer-term target, the proportion of people that may benefit from cycling could be increased by promoting higher-density mixed land use development and shorter-term interventions to unlock existing potential. (Jarass and Scheiner, Jarass and Scheiner) describes that creating environments conducive to cycling in mixed-use areas is likely to attract more people that are already biased towards cycling as a form of residential self-selection.

The Pragmatic scenario demonstrates how the methodology can be applied to a wide range of scenarios that target variations in any or all the factors that affect the propensity to cycle. Examples include: *Network* - e.g. adding dedicated cycle lanes to reduce distances relative to traffic or providing communal bike parking in informal settlements; *Population* - e.g. promoting e-bikes among elderly or high-income households); *Travel behaviour* – e.g. subsidising e-bikes among students, and; *Activity chains* - e.g. latent demand among non-travellers in a specific region.

A relatively accurate model (not fully calibrated) can already demonstrate outcomes for different scenarios relative to the base case or each other. When calibrated, the model may be able to guide investment decisions since cost benefits may be estimated at a suitable level of accuracy. Alternatively, surveys can be designed to target specific geographies of market segments to determine specific preferences, needs or barriers to cycling, which may lead to the design and phased implementation of interventions.

In this study, we do not account for secondary personal benefits, such as cost savings and health benefits, to reduce the disutility of cycling. Expanding the model to consider such factors would increase the pool of potential cyclists in an SCC or may be used to test the growth potential in mature cycling cities.

The model could be used to screen several cities within a country or region for their readiness to implement cycling programmes. This ability may be helpful to governments or funding agencies that wish to improve the sustainability of urban transport systems on the back of cycling. The same methodology could be applied to introducing modes such as Bus-Rapid Transit (BRT) or any intervention targeted to specific user groups or city regions.

The study demonstrates that a significant shift towards cycling would benefit other road users and potential cyclists. The benefit for car drivers is explained through basic traffic flow theory, where the marginal decrease in traffic volume on a congested road network leads to an exponential reduction in delay. Such a decrease in delay may result in induced traffic if road capacity is not concurrently restricted. However, induced traffic may have positive socio-economic benefits in a society where poor accessibility constrains social and economic activity.

The authors plan to expand the research to include metrics that estimate greater societal benefits from a significant shift to cycling. This would be achieved by having health benefits and lower transport costs for cyclists, reduced emissions from reduced traffic and congestion, financial and environmental savings on reduced infrastructure spending and others. The information generated in this way could inform the extent of budgets and other resources that can be applied to achieve the benefits by a chosen target date. The modelling process output should provide policymakers confidence that interventions will likely yield the anticipated benefits associated with a cycling city.

## References

- Adjei E, Behrens R (2010) Travel behaviour change theories and experiments: A review and synthesis. *Proceedings of the 29th Southern African Transport Conference (SATC)*.
- Aldred, R., J. Woodcock, and A. Goodman. 2016. Does more cycling mean more diversity in cycling? *Transport Reviews* 36(1): 28–44. Doi: [10.1080/01441647.2015.1014451](https://doi.org/10.1080/01441647.2015.1014451).
- Behrens R, Del Mistro R, Lombard M, et al (2007) The pace of behaviour change and implications for TDM response lags and monitoring: Findings of a retrospective commuter travel survey in Cape Town. Proceedings of the 26th Southern African Transport Conference (SATC)
- BEN. 2020. <https://www.benbikes.org.za/>. Available online at <https://www.benbikes.org.za/>.
- Blondiau, T., B. van Zeebroek, and H. Haubold. 2016. Economic benefits of increased cycling. *Transportation Research Procedia* 14: 2306–2313. Doi: [10.1016/j.trpro.2016.05.247](https://doi.org/10.1016/j.trpro.2016.05.247).
- Buehler, R. and J. Dill. Bikeway networks: A review of effects on cycling. 36(1): 9–27. Doi: [10.1080/01441647.2015.1069908](https://doi.org/10.1080/01441647.2015.1069908).
- Cervero, R., O. Sarmiento, E. Jacoby, L. Gomez, and A. Neiman. Influences of built environments on walking and cycling: Lessons from Bogotá. 3(4): 203–26. Doi: [10.1080/15568310802178314](https://doi.org/10.1080/15568310802178314).
- Charypar, D., K. Nagel. 2005 Generating complete all-day activity plans with genetic algorithms. *Transportation* 32: 369–397. Doi: [10.1007/s11116-004-8287-y](https://doi.org/10.1007/s11116-004-8287-y).
- Chatterjee, K. 2001. Asymmetric churn-academic jargon or a serious issue for transport planning? *Transport Planning Society Bursary Paper*.
- Chillón, P., J. Molina-García, I. Castillo, and A. Queralt. 2016. What distance do university students walk and bike daily to class in Spain. *Journal of Transport & Health* 3: 315–320. Doi: [10.1016/j.jth.2016.06.001](https://doi.org/10.1016/j.jth.2016.06.001).
- City of Cape Town 2013. Household survey report - final draft. Technical report, City of Cape Town.
- City of Cape Town 2017. Cycling strategy for the city of Cape Town. Technical report, Transport and Urban Development Authority (TDA).
- Copenhagen 2010. Good, better, best: The city of Copenhagen’s bicycle strategy 2011-2025. Technical report, City of Copenhagen.
- Damant-Sirois, G. and A.M. El-Geneidy. 2015. Who cycles more? determining cycling frequency through a segmentation approach in Montreal, Canada. *Transportation Research Part A: Policy and Practice* 77: 113–125. Doi: [10.1016/j.tra.2015.03.028](https://doi.org/10.1016/j.tra.2015.03.028).
- de D. Ortúzar, J. and L.G. Willumsen. 1994. *Modelling transport* (Second ed.). Chichester, UK: John Wiley & Sons, Ltd.
- Department for Transport. Gear change: A bold vision for cycling and walking.
- Dias, A.M., M. Lopez, and C. Silva. More than cycling infrastructure: Supporting the development of policy packages for starter cycling cities. 2676: 785–797. Doi: [10.1177/03611981211034732](https://doi.org/10.1177/03611981211034732).
- Dill, J. and G. Rose. 2012. Electric bikes and transportation policy: Insights from early adopters. *Transportation Research Record* 2314. Doi: [10.3141/2314-01](https://doi.org/10.3141/2314-01).
- DTU. 2015. The Danish National Travel Survey-fact sheet about cycling in Denmark. [https://cyclingsolutions.info/embassy\\_category/cycling-facts-figures/](https://cyclingsolutions.info/embassy_category/cycling-facts-figures/).
- Ekblad, H., A. Svensson, and T. Koglin. Bicycle planning in an urban context: A literature review. Technical report, Bulletin 300, Transport and Roads, Dept of Technology and Society, Lund University, Lund.
- FLOW 2016. The role of walking and cycling in reducing congestion: A portfolio of measures. Technical report, FLOW, Brussels.
- Fonseca, F., P. Ribeiro, and C. Neiva. A planning practice method to assess the potential for cycling and to design a bicycle network in a starter cycling city in Portugal. pp. 4534. Doi: [10.3390/su15054534](https://doi.org/10.3390/su15054534).
- Goldman, K. and J. Wessel. 2021. Some people feel the rain, others just get wet: An analysis of regional differences in the effects of weather on cycling. *Research in Transportation Business & Management* 40. Doi: [10.1016/j.rtbm.2020.100541](https://doi.org/10.1016/j.rtbm.2020.100541).
- Goldsmith, S. 1992. Case study no. 1: Reasons why bicycling and walking are not being used more extensively as travel modes. In *National bicycling and walking study*.
- Goodman, A., I.F. Rojas, J. Woodcock, R. Aldred, N. Berkoff, M. Morgan, A. Abbas, and R. Lovelace. Scenarios of cycling to school in England, and associated health and carbon impacts: Application of the ‘propensity to cycle tool’. 12(2): 263–278. Doi: [10.1016/j.jth.2019.01.008](https://doi.org/10.1016/j.jth.2019.01.008).
- GTZ. 2003. Preserving and expanding the role of non-motorised transport, GTZ sustainable urban transport project. Available online at [www.sutp.org](http://www.sutp.org).
- Heinen, E., B. van Wee, and K. Maat. 2010. Commuting by bicycle: An overview of the literature. *Transport Reviews* 30(1): 59–96. Doi: [10.1080/01441640903187001](https://doi.org/10.1080/01441640903187001).
- Hitge, G. and J.W. Joubert. 2021. A nodal approach for estimating potential cycling demand. *Journal of Transport Geography* 90, 102943.
- Hitge, G. and M. Vanderschuren. 2015. Comparison of travel time between private car and public transport in Cape Town. *SAICE Journal* 57(3): 35–43. Doi: [10.17159/2309-8775/2015/V57N3A5](https://doi.org/10.17159/2309-8775/2015/V57N3A5).
- Horni, A., K. Nagel, and K.W. Axhausen. 2016. *The Multi-Agent Transport Simulation, MATSim*. London: Ubiquity press.
- Hunt, J.D. and J.E. Abraham. 2007. Influences on bicycle use. *Transportation* 34: 453–470. Doi: [10.1007/s11116-006-9109-1](https://doi.org/10.1007/s11116-006-9109-1).

- Irlam, J.H. and M. Zuidgeest. 2018. Barriers to cycling mobility in a low-income community in Cape Town: A best-worst scaling approach. *Case studies on Transport policy* 6: 815–823. Doi: [10.1016/j.cstp.2018.10.003](https://doi.org/10.1016/j.cstp.2018.10.003) .
- Jarass, J. and J. Scheiner. Residential self-selection and travel mode use in a new inner-city development neighbourhood in Berlin. 70: 68–77. Doi: [10.1016/j.jtrangeo.2018.05.018](https://doi.org/10.1016/j.jtrangeo.2018.05.018) .
- Jennings, G., B. Petzer, and E. Goldman 2017. *When bicycle lanes are not enough: Growing mode share in Cape Town, South Africa: An analysis of policy and practice*, Chapter 13. Cornwall: Routledge.
- Joubert, J.W. 2021. Synthetic populations of South African urban areas. *Mendeley Data*. Doi: [10.17632/dh4gcm7ckb.2](https://doi.org/10.17632/dh4gcm7ckb.2) .
- Joubert, J.W. and A. de Waal. 2020. Activity-based travel demand generation using Bayesian networks. *Transportation Research Part C: Emerging Technologies* 120: 102804. Doi: [10.1016/j.trc.2020.102804](https://doi.org/10.1016/j.trc.2020.102804) .
- KiM 2019. Mobiliteitsbeeld 2019 (mobility report 2019). Technical report, Ministry of Infrastructure and Water Management, The Hague. Netherlands Institute for Transport Policy Analysis (KiM) Het Kennisinstituut voor Mobiliteitsbeleid.
- Kuzmyak, R., J. Walters, M. Bradley, and K. Kockelman. Estimating bicycling and walking for planning and project development: A guidebook. Doi: [10.17226/22330](https://doi.org/10.17226/22330) .
- Lovelace, R., A. Goodman, R. Aldred, N. Berkoff, A. Abbas, and J. Woodcock. 2017. The propensity to cycle tool: An open source online system for sustainable transport planning. *Journal of Transport and Land Use* 10(1). Doi: [10.5198/jtlu.2016.862](https://doi.org/10.5198/jtlu.2016.862) .
- Mahfouz, H., E. Arcaute, and R. Lovelace. A road segment prioritization approach for cycling infrastructure. Doi: [10.48550/arXiv.2105.03712](https://doi.org/10.48550/arXiv.2105.03712) .
- Midgley, P. 2011. Bicycle-sharing schemes: enhancing sustainable mobility in urban areas. *United Nations, Department of Economic and Social Affairs* 8: 1–12.
- National Department of Transport. 2020. Shova Kalula. Available online at <https://www.transport.gov.za/>.
- Nello-Deakin, S. and A. Nikolaeva. 2021. The human infrastructure of a cycling city: Amsterdam through the eyes of international newcomers. *Urban Geography* 42(3): 289–311. Doi: [10.1080/02723638.2019.1709757](https://doi.org/10.1080/02723638.2019.1709757).
- Nordic Council of Ministers. 2005, Accessed 22 December 2021. *CBA of Cycling*. Copenhagen: ProQuest Ebook Central.
- OECD 2007. Managing urban traffic congestion. Technical report, OECD, Paris.
- Olmos, L.E., M.S. Tadeo, D. Vlachogiannis, F. Alhasoun, X. Espinet Alegre, C. Ochoa, F. Targa, and M.C. González. 2020. A data science framework for planning the growth of bicycle infrastructures. *Transportation Research Part C: Emerging Technologies* 115: 102640. Doi: [10.1016/j.trc.2020.102640](https://doi.org/10.1016/j.trc.2020.102640).
- OSM. 2021. *OpenStreetMap*. Available online at <http://www.openstreetmap.org>.
- O'Connor, J.P. and T.D. Brown. 2010. Riding with the sharks: Serious leisure cyclist's perceptions of sharing the road with motorists. *Journal of Science and Medicine in Sport* 13(1): 53–58. Doi: [10.1016/j.jsams.2008.11.003](https://doi.org/10.1016/j.jsams.2008.11.003).
- Parkin, J., M. Wardman, and M. Page. 2008. Estimation of the determinants of bicycle mode share for the journey to work using census data. *Transportation* 35(1): 93–109. Doi: [10.1007/s11116-007-9137-5](https://doi.org/10.1007/s11116-007-9137-5).
- Pucher, J. and R. Buehler. 2008. Cycling for everyone: Lessons from the Netherlands, Denmark and Germany. *Transportation Reviews* 28(4): 495–528. Doi: [10.1080/01441640701806612](https://doi.org/10.1080/01441640701806612).
- Qhubeka. 2020. Qhubeka. Available online at <https://qhubeka.org/>.
- Reid, L.W. and M. Konrad. 2004. The gender gap in fear: Assessing the interactive effects of gender and perceived risk on fear of crime. *Sociological Spectrum* 24(4): 399–425. Doi: [10.1080/02732170490431331](https://doi.org/10.1080/02732170490431331).
- Riggs, W. 2016. Cargo bikes as a growth area for bicycle vs. auto trips: Exploring the potential for mode substitution behavior. *Transportation Research Part F: Traffic Psychology and Behaviour* 43: 48–55. Doi: [10.1016/j.trf.2016.09.017](https://doi.org/10.1016/j.trf.2016.09.017).
- Rodriguez, D.A. and J. Joo. 2004. The relationship between non-motorized mode choice and the local physical environment. *Transportation Research Part D: Transport and Environment* 9(2): 151–173. Doi: [10.1016/j.trd.2003.11.001](https://doi.org/10.1016/j.trd.2003.11.001).
- Saleh, W. and S. Farrell. 2007. Investigation and analysis of evidence of asymmetric churn in travel demand models. *Transportation Research Part A: Policy and Practice* 41(7): 691–702. Doi: [10.1016/j.tra.2006.09.016](https://doi.org/10.1016/j.tra.2006.09.016).
- Saneinejad, S., M.J. Roorda, and C. Kennedy. 2012. Modelling the impact of weather conditions on active transportation travel behaviour. *Transportation Research Part D: Transport and Environment* 17: 129–137. Doi: [10.1016/j.trd.2011.09.005](https://doi.org/10.1016/j.trd.2011.09.005).
- Silva, C., J. Teixeira, and A. Proença. 2019. Revealing the cycling potential of starter cycling cities. *Transportation Research Procedia* 41: 637–654. Doi: [10.1016/j.trpro.2019.09.113](https://doi.org/10.1016/j.trpro.2019.09.113) .
- StatsSA. 2011. South African Census 2011. Available online at <https://www.statssa.gov.za/publications/P03014/P030142011.pdf>. Accessed 20200228.
- StatsSA 2012a. 2011 census data. Technical report, Statistics South Africa.
- StatsSA 2012b. Census 2011 municipal report - Western Cape. Technical report, Statistics South Africa.
- StatsSA 2013. National household travel survey february to march 2013. Technical report, Statistics South Africa, Pretoria.
- Sun, L. and A. Erath. 2015. A Bayesian network approach for population synthesis. *Transportation Research Part C: Emerging Technologies* 61: 49–62. Doi: [10.1016/j.trc.2015.10.010](https://doi.org/10.1016/j.trc.2015.10.010).
- Sweet, M. and P. Kanaroglou. 2016. Gender differences: The role of travel and time use in subjective well-being. *Transportation Research Part F: Traffic Psychology and Behaviour* 40: 23–34. Doi: [10.1016/j.trf.2016.03.006](https://doi.org/10.1016/j.trf.2016.03.006).

- Taylor, B.D., K. Ralph, and M. Smart. 2015. What explains the gender gap in schlepping? testing various explanations for gender differences in household-serving travel. *Social Science Quarterly* 96(5): 1493–1510. Doi: [10.1111/ssqu.12203](https://doi.org/10.1111/ssqu.12203).
- Tortosa, E.V., R. Lovelace, E. Heinen, and R. Mann. Infrastructure is not enough: Interactions between the environment, socio-economic disadvantage, and cycling participation in England. *I4*(1): 693–714. Doi: [10.5198/jtlu.2021.1781](https://doi.org/10.5198/jtlu.2021.1781).
- UNESCO 2012. International standard classification of Education-ISCED 2011. Technical report, UNESCO Institute for Statistics.
- Van Cauwenberg, J., I.D. Bourdeaudhuij, P. Clarys, B. de Geus, and B. Deforche. 2019. E-bikes among older adults: benefits, disadvantages, usage and crash characteristics. *Transportation* 46: 2151–2172. Doi: [10.1007/s11116-018-9919-y](https://doi.org/10.1007/s11116-018-9919-y).
- Wardman, M.R., M.R. Tight, and M. Page. 2007. Factors influencing the propensity to cycle to work. *Transport Research Part A* 41(4): 339–350. Doi: [10.1016/j.tra.2006.09.011](https://doi.org/10.1016/j.tra.2006.09.011).
- Wooliscroft, B. and A. Ganglmair-Wooliscroft. 2014. Improving conditions for potential New Zealand cyclists: An application of conjoint analysis. *Transportation Research Part A: Policy and Practice* 69: 11–19. Doi: [0.1016/j.tra.2014.08.005](https://doi.org/10.1016/j.tra.2014.08.005).
- Young, M., B. Savan, K. Manaugh, and J. Scott. 2020. Mapping the demand and potential for cycling in Toronto. *International Journal of Sustainable Transportation* 15(4): 285–293. Doi: [10.1080/15568318.2020.1746871](https://doi.org/10.1080/15568318.2020.1746871).
- Zhang, D., D.J.A.V. Magalhães, and X.C. Wang. Prioritizing bicycle paths in belo horizonte city, brazil: Analysis based on user preferences and willingness considering individual heterogeneity. *I4*: 268–278. Doi: [10.1016/j.tra.2014.07.010](https://doi.org/10.1016/j.tra.2014.07.010).

## Biographies

**Gerhard Hitge** has been a transportation engineer in the public and private sectors for the past 20 years. He has a Master’s Degree in Transportation Engineering (University of Pretoria). He was employed by the City of Cape Town (CCT) and now operates as an independent consultant. His current focus is on developing integrated, multi-modal transport systems sustained by higher-density, mixed-use urban development. He is currently a doctoral student in Industrial Engineering at the University of Pretoria, South Africa.

**Johan W. Joubert** is Professor at KU Leuven in the Centre for Industrial Management (CIB). He is also associated with the Leuven Institute for Mobility. His research focuses on behaviourally-sensitive models, where he prefers agent-based simulation to capture the dynamic and emergent phenomena in the mobility of both people and freight stakeholders.

## CRedit statement

**Gerhard Hitge** Conceptualisation; formal analysis; validation; writing – original draft; writing — review and editing.

**Johan W. Joubert** Conceptualisation; data curation; formal analysis; methodology; resources; software; supervision; visualisation; writing – original draft; writing — review and editing.

## Data availability

The datasets generated and analysed during the current study are not publicly available because they constitute an excerpt of research in progress but are available from the corresponding author upon reasonable request.