Transportation Research Part D 118 (2023) 103695

Contents lists available at ScienceDirect

Transportation Research Part D

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

High fidelity estimates of paratransit energy consumption from per-second GPS tracking data

Christopher Hull [a,*], J.H. Giliomee [b], Katherine A. Collett [a], Malcolm D. McCulloch [a], M.J. Booysen [b,*]

[a] *Energy and Power Group at the Engineering Science Department, University of Oxford, United Kingdom*
[b] *Department of E&E Engineering, Stellenbosch University, South Africa*

ARTICLE INFO

Keywords:
Electric vehicles
Energy transition
Energy consumption
Energy planning
Infrastructure planning
Transportation

ABSTRACT

Paratransit, in particular the minibus taxi, is the mainstay of public transport in sub-Saharan Africa. These vehicles are often second-hand, ageing, fuel inefficient, and expensive to operate - issues that electrification can ameliorate. However, modeling and planning large-scale transitions to electric paratransit require reliable estimates of vehicle energy consumption. This paper provides such estimates by applying a vehicle kinetic model to per-second GPS data gathered on minibus taxis. Data include 62 trips across three routes with different driving conditions near Stellenbosch, South Africa. We find a range of energy consumption from 0.29 to 0.51 kWh/km (mean = 0.39 kWh/km). Past estimates in literature relied on per-minute GPS data, which we show leads to inaccurate energy consumption estimates. We recommend new kWh/km values for modeling vehicle operations and grid impact, and discuss how future work can utilize our analysis to advance the transition to electric paratransit sub-Saharan Africa.

1. Introduction

This paper aims to provide high fidelity energy consumption estimates (kWh/km) for electric paratransit vehicles in South Africa. Establishing per-distance energy consumption for electric vehicles (EV) is necessary for planning operations, sizing vehicle batteries, determining appropriate charging infrastructure, and managing vehicle impact on the grid. The availability of high fidelity energy consumption estimates is paramount for the optimization of these models, and thus critical for stakeholders in governments, charge point manufacturers and operators, fleet managers, vehicle manufacturers, utilities companies, network operators, and city planners to effectively plan large-scale transitions to EVs. By providing specific recommendations for kWh/km values to use in these models in various conditions, this paper will improve the efficiency and cost savings of a transition to electric paratransit, which can benefit both operators and passengers.

Transport accounts for roughly third of global energy consumption, and is responsible for about 16% of global emissions, and the development of low-carbon transport in cities is part of the global agenda to mitigate climate change and relates to at least three of the United Nation's Sustainable Development Goals (Zinkernagel et al., 2018). Accordingly, EV sales have seen substantial growth in the Global North and many global vehicle manufacturers and governments seek to stop production of combustion engines altogether as early as 2035 (Motavalli, 2021; Geospatial Commission, 2021; Sunday Times Driving, 2022). In contrast, due to low electricity access and high upfront costs (Pillay et al., 2019), the transition to more expensive EVs has been painstakingly slow in sub-Saharan Africa (SSA). Africa is a major destination for old and used vehicles, which typically consume particularly dirty fuel such as diesel

* Corresponding authors.
E-mail addresses: christopher.hull@eng.ox.ac.uk (C. Hull), mjbooysen@sun.ac.za (M.J. Booysen).

https://doi.org/10.1016/j.trd.2023.103695
Received 25 October 2022; Received in revised form 24 January 2023; Accepted 7 March 2023
Available online 20 March 2023

with high-sulfur content. This not only causes serious air pollution and health problems in African cities, but contributes heavily to greenhouse gas emissions (Odhiambo et al., 2021). The fuel for these vehicles is typically imported, leaving these countries with issues related to fuel quality, energy security and price fluctuations (Odhiambo et al., 2021).

Privately-owned minibus taxis are ubiquitous in the developing cities and rural areas of SSA. They form a substantial part of the "paratransit" system—an informal transport sector common in the region. Paratransit travel comprises at least 90% of road-based public trips in Lagos, Kampala and Dar es Salaam (Behrens et al., 2015; Evans et al., 2018), of which 83% are by minibus taxi (Mccormick et al., 2016; Evans et al., 2018; KCCA, 2016). In general, paratransit is estimated to provide 50%–98% of daily commutes in most major cities around sub-Saharan Africa (Behrens et al., 2015). In South Africa, minibus taxis form the entire paratransit industry, which is worth $3.7bn. Every day, over 300,000 minibus taxis carry out approximately 15 million trips, amounting to more than 73% of daily commuters (Transaction Capital, 2021). In the face of the increasing demand for urban commuting heralded by urbanization and population growth, electrifying minibus taxis becomes even more critical for meaningful transport sector decarbonization and reductions in air pollution in sub-Saharan Africa (SSA) (Ayetor et al., 2021; Lin and Sai, 2022).

It is well recognized that this informal sector is now faced with the need to transform to an electrical energy source (Behrens et al., 2015; Evans et al., 2018; Mccormick et al., 2016; Ehebrecht et al., 2018). However, electricity grids in the region are energy-constrained, which affects the feasibility of converting to EVs. In addition, many stakeholders and decision makers are capital constrained, and thus investment must be carefully considered. Therefore, before commencing a large scale transition, it is crucial to establish high fidelity estimates of minibus taxi energy requirements, based on their micro-mobility patterns across driving conditions. Without high fidelity estimates, it would be difficult for stakeholders to plan operations, size viable vehicle batteries, manage EV impact on the grid, and design appropriate charging infrastructure.

The driving cycles and the concomitant energy requirements of electric passenger vehicles and larger buses in higher-income countries are well established and deeply analyzed (Brady and O'Mahony, 2016; Berzi et al., 2016; Kivekäs et al., 2018; Smith et al., 2011; Wang et al., 2008), and there are energy consumption estimates for minibuses in developed contexts (Cignini et al., 2020). Nevertheless, there is a distinct lack of high fidelity analysis for paratransit in SSA. Given the unique mobility patterns and unconventional driving styles of African paratransit, high-income country conventional driving cycles do not apply to these contexts (Ndibatya and Booysen, 2021).

To estimate vehicle energy consumption, we apply a vehicle kinetic model to per-second GPS data gathered on trips taken on internal combustion engine (ICE) minibus taxis to model how much energy would be required to power an electric equivalent. To date, estimations of energy consumption for electric paratransit in SSA have been either based on data captured at a sampling frequency of once per minute (Abraham et al., 2021; Booysen et al., 2022), or have been forced to a conservative simplifying assumption due to a lack of data (Collett et al., 2021). However, paratransit vehicle drivers have been reported to engage in aggressive maneuvers characterized by sharp acceleration and/or deceleration movements that last mere seconds, and for making numerous quick stops to continuously pick up and drop off passengers (Zeeman and Booysen, 2014). The energy intensity of aggressive micro-mobility patterns have been observed to have a great effect on per-distance energy consumption (Al-Wreikat et al., 2021; Faria et al., 2019; Murphey et al., 2009; Eno Akpa et al., 2019). Therefore, following the Nyquist–Shannon sampling theorem (Shannon, 1949), we hypothesize that a sampling frequency that provides several samples per movement is needed to capture their effect on energy consumption. This paper aims to fill this gap in the literature and improve stakeholder confidence in paratransit vehicle energy consumption estimates by utilizing per-second GPS data that can capture vehicle micro-mobility patterns to provide high fidelity energy consumption estimates. After presenting our main results, we verify that per-second data provides higher fidelity estimates than per-minute data.

Another overlooked aspect in the literature is disaggregating energy consumption estimates by driving conditions. EV energy consumption is known to be highly dependent on driving conditions with different elevation, speed, and micro-mobility patterns (Al-Wreikat et al., 2021; Pan et al., 2017; Jonas et al., 2017), and previous literature has not provided kWh/km estimates for different driving conditions. To fill this gap in the literature and capture the effect of variation in driving conditions on energy consumption, this paper uses data captured on urban, hilly, and inter-city routes. We hypothesized that time of day may also have an effect on traffic patterns, so for each route we captured data in the morning, afternoon, and evening.

Our high fidelity estimates can be used to establish a benchmark for paratransit energy consumption. Once established, such a standard could be used to validate future energy consumption estimates from per-minute data in various locations and driving conditions. This is useful since gathering high-resolution data for vehicles across many different contexts would be time and labor intensive. There are mobility companies such as WhereIsMyTransport and GoMetro that already capture telemetry data for minibus taxis at a resolution of once per-minute in many cities around SSA to create spatial mobility models for operations and route planning, and there is potential for substantial benefit to being able to take advantage of this data that already exists or is being gathered. Thus, while providing high fidelity vehicle energy consumption estimations for minibus taxis in South Africa that can be used immediately, this paper provides a platform for future research to develop accurate methods of estimation based on per-minute data input. For example, developing a micro-traffic simulator, based on the micro-traffic mobility patterns we see in the per-second data, which can intake per-minute data and accurately model taxi driver behavior in between waypoints. With such a tool, per-minute travel data could be fed into a micro-traffic simulator to come up with reliable energy consumption estimations for paratransit in many different locations. Fig. 1 shows how this validation process can feed into large-scale EV and infrastructure design.

2. Methodology

To construct high fidelity energy consumption estimates for minibus taxis: (i) a kinetic model is required to establish energy demand from motion based on GPS data, and (ii) granular GPS data from representative journeys is necessary. Fig. 2 visualizes at a high level the process for using GPS data and a kinetic model to construct energy consumption estimates.

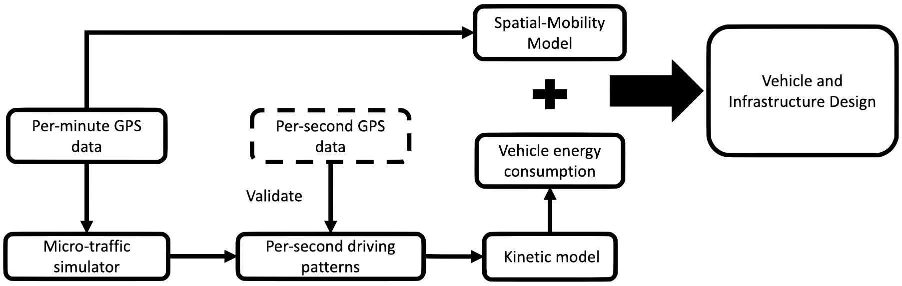


Fig. 1. How per-second GPS data can support EV and infrastructure design by validating energy consumption estimates derived from GPS per-minute data.

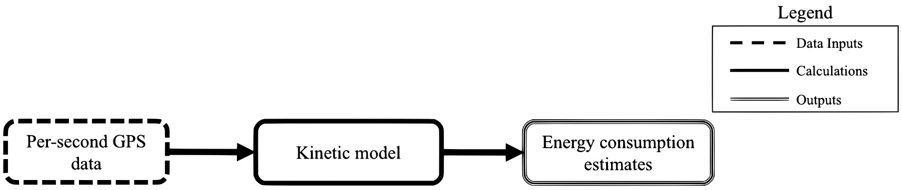


Fig. 2. How per-second GPS data is used to estimate energy consumption via a kinetic model.

2.1. Kinetic model

The objective of the kinetic model is to estimate the energy consumed by the vehicle over the duration of a trip. In this section, we provide background on kinetic models seen in the literature, justify the principles behind the construction of our kinetic model, and describe our model mathematically.

All vehicle kinetic models are derived from the same fundamental principles of physics, which inform us that the instantaneous power output of an EV is a function of its velocity, acceleration, and the slope angle it faces. However, kinetic models in the literature often differ slightly; various kinetic models have been employed to provide estimations of EV energy consumption based on GPS travel data (Kurczveil et al., 2014; Wu et al., 2015; Fiori et al., 2016; Maia et al., 2011; Ilyès et al., 2020). The fundamental laws of physics cannot change, so these variations come from discrepancies based on assumptions from empirical data or fitted model parameters. Validation with ground-truth data has shown that these variations can have an effect on model performance (Sagaama et al., 2019).

For example, Sagaama et al. (2019) show that the accuracy of the kinetic model integrated in the micro-traffic simulator SUMO (Simulation Of Urban MObility) improves when applying a certain dynamic regenerative braking formula in favor of a static regenerative braking factor. We find that adding this dynamic regenerative braking formula into our kinetic mode decreases mean vehicle energy consumption across all trips by 15% (from 0.39 kWh/km to 0.33 kWh/km), a substantial margin. A key assumption in this dynamic formula is that the drivers are making full use of regenerative braking every time they decelerate. However, in this context, given the taxi drivers' lurching driving style and heavy usage of brake pads, this means the dynamic formula will overestimate the energy recuperated back into the battery. For this reason, our kinetic model does not use a dynamic regenerative braking formula. Instead, we apply a static regenerative braking factor that is not subject to being highly influenced by extreme deceleration events.

Sagaama et al. (2019) also show that ambient air temperature can have an effect on power offtake. A limitation of our dataset is that temperature data was not available. However, given that (a) the dynamic power offtake evaluated in Sagaama et al. (2019) was not constructed for paratransit contexts and (b) there is a paucity of low-voltage auxiliary power applications on board paratransit, it is not likely in the authors' estimation that this would have increased model accuracy or significantly affected the results.

While some empirical assumptions and fitted model parameters have been validated in developed contexts with ground-truth benchmark of real per-second EV energy consumption data (Wu et al., 2015; Fiori et al., 2016), these results do not necessarily extrapolate to sub-Saharan African paratransit driving contexts. Furthermore, the motivation of this paper is to provide high fidelity vehicle energy consumption estimations, not a novel kinetic model. Therefore, to avoid concern that our energy consumption estimates are potentially skewed by underlying empirical assumptions or fitted model parameters, the model we employ is based solely on fundamental principles of physics.

While we entirely avoid basing our kinetic model on empirical assumptions or fitted model parameters, any kinetic model must employ some specific vehicle constant parameters to obtain EV energy consumption estimates. These constant parameters along with the relevant references are listed in Table 1. Certain parameters can vary along the duration of a trip, namely: minibus weight, rolling resistance coefficient, powertrain efficiency, regenerative braking efficiency, and power offtake. Justification for avoiding dynamic calculations of the latter two parameters have been explained in this section, but further explanation is required to justify keeping the former three constant: For minibus weight, laden weight for an electric minibus is used. The taxis fill up with passengers

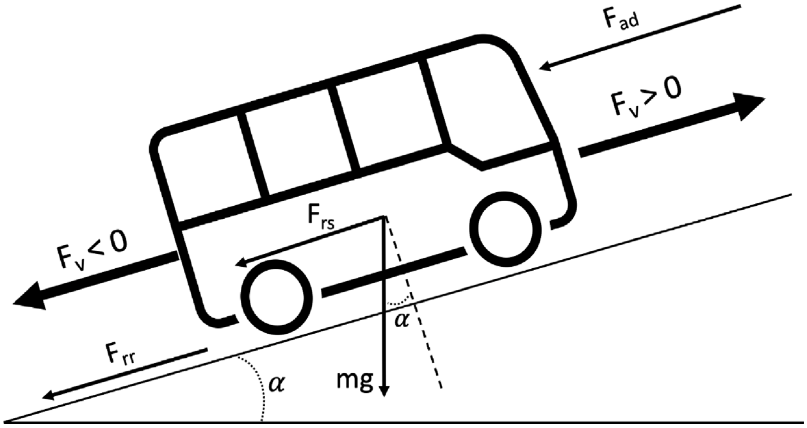


Fig. 3. The forces acting on a vehicle driving on a slope: Aerodynamic drag (F_{ad}), rolling resistance (F_{rr}), slope drag (F_{sd}), and vehicle force (F_v) which can be either positive or negative depending on whether the vehicle is accelerating or decelerating. F_v is shown with thicker lines as it is usually the greatest force, since it must overcome other forces in order to accelerate or decelerate the vehicle.

Table 1
Constants and parameters used in kinetic model.

Parameter	Symbol	Value	Ref.
Gravitational acceleration	g	9.81 m/s^2	–
Density of air at 25 °C	ρ	1.184 kg/m^3	Engineering Toolbox (2008)
Power offtake	p_0	100 W	Fridlund and Wilen (2020)
Minibus weight	m_v	3900 kg	Higer (2020)
Drag coefficient	c_d	0.36	Toyota (2022a)
Rolling resistance coefficient	c_{rr}	0.02	Engineering Toolbox (2008)
Vehicle's front surface area	A	4 m^2	Toyota (2022b)
Powertrain efficiency	μ_v	90%	Renault Group (2021)
Regenerative braking efficiency	μ_{rg}	65%	Tesla (2007)

at the start of the journey at the taxi ranks, so while some pick up and drop off occurs, they are almost always operating at laden weight. For rolling resistance coefficient, the parameter value used represents the average of car tires on tar or asphalt, which is the average road quality upon which these vehicles were traveling. Dynamically modeling powertrain efficiency comes with similar disadvantages as dynamically modeling regenerative braking efficiency. Electric powertrains behave differently than ICE powertrains, and a dynamic model of powertrain efficiency would have been sensitive to the taxis' lurching driving style. Thus, a dynamic model of powertrain efficiency would have been inaccurate in this setting, and the parameter value is chosen based on an estimate of average expected electric powertrain efficiency for the duration of the trips. However, the authors recognize that the kinetic model can be sensitive to these parameters, so a sensitivity analysis is conducted and presented in Section 3.3.2. These parameters can be empirically fitted to electro-kinetic models with a greater degree of confidence when electric taxis are introduced to South Africa in the future and direct measurement is available.

To compute energy consumption over the duration of a trip, we use per-second GPS data to separately compute either the energy out of the battery to turn the wheels or power auxiliary functions, or the energy flow into the battery from regenerative braking at each sample's time step n, $E[n]$. We then sum the estimate at each time step to obtain entire trip's energy consumption $E_{\mathrm{trip}} = \sum E[n]$.

To determine energy flow at each time step, $E[n]$ in (kWh), we implement the following five-step algorithm:

(1) Compute the sum of all external forces on the vehicle for sample n at a sample period of τ, $F_{\mathrm{ext}}[n]$ in (N), using

$$F_{\mathrm{ext}}[n] = -F_{\mathrm{ad}}[n] - F_{\mathrm{rs}}[n] - F_{\mathrm{rr}}[n] \tag{1a}$$

where $F_{\mathrm{ad}}[n]$ is the aerodynamic drag, $F_{\mathrm{rr}}[n]$ is the rolling resistance friction, and $F_{\mathrm{rs}}[n]$ is the slope drag force (positive for incline or negative for decline). The forces are shown in Fig. 3.

The forces in Fig. 3 are individually calculated from fundamental physics using the following equations and the constants listed in Table 1.

$$F_{\mathrm{ad}}[n] = \frac{1}{2}\rho c_d A(v[n])^2 \tag{1b}$$

$$F_{\mathrm{rr}}[n] = \begin{cases} m_{\mathrm{v}} g c_{\mathrm{rr}} cos(\alpha[n]), & \text{if } v[n] > 0.3 \text{ m/s (GPS noise threshold)} \\ 0, & \text{otherwise} \end{cases} \tag{1c}$$

$$F_{\mathrm{sd}} = m_{\mathrm{v}} g sin(\alpha_{\mathrm{t}}) \tag{1d}$$

where

$$\alpha[n] = \begin{cases} \arcsin(\frac{h[n]-h[n-1]}{s[n]}) & \text{if } \mid h[n]-h[n-1] \mid > 0.1 \text{ m (GPS noise threshold)} \\ 0, & \text{otherwise} \end{cases}$$

(2) Compute the expected change in velocity (in m/s) if $F_{ext}[n]$ were the only force applied to the vehicle.

$$d v_{\text{exp}}[n] = \frac{F_{\text{ext}}[n]}{m_{\text{v}}}\tau \tag{2}$$

(3) Any difference between the expected velocity and measured velocity, v_{meas}, is ascribed to the acceleration or deceleration force applied by the vehicle, F_{v}, calculated by:

$$F_{\text{v}}[n] = \frac{v_{\text{meas}}[n] - v_{\text{exp}}[n]}{\tau} m_{\text{v}} \tag{3}$$

(4) When $F_{\text{v}}[n] > 0$, energy in (Ws) was **discharged from** the battery to apply propulsion to the vehicle.

$$E_{\text{prop}}[n] = \frac{F_{\text{v}}[n] v[n] \tau}{\mu} \tag{4a}$$

When $F_{\text{v}}[n] < 0$, energy in (Ws) was **charged into** the battery via regenerative braking:

$$E_{\text{regen}}[n] = \mu_{\text{rg}} F_{\text{v}}[n] v[n] \tau \tag{4b}$$

The energy discharge from the battery simply to keep the vehicle running is defined as power offtake. For every measured time interval, power offtake in (Ws) is calculated by

$$E_{\text{offtake}}[n] = p_0 \tau. \tag{4c}$$

(5) Compute the total energy flow in a given time-step $E[n]$ in (kWh) by summing the propulsive, regenerative, and constant off-take energies.

$$E[n] = \frac{E_{\text{prop}}[n] + E_{\text{regen}}[n] + E_{\text{offtake}}[n]}{3.6e+6} \tag{5}$$

Finally, we can compute the energy consumption over the whole trip in (kWh) by summing over the duration of the trip N.

$$E_{\text{trip}} = \sum_{0}^{N} E[n] \tag{6}$$

2.2. GPS data collection

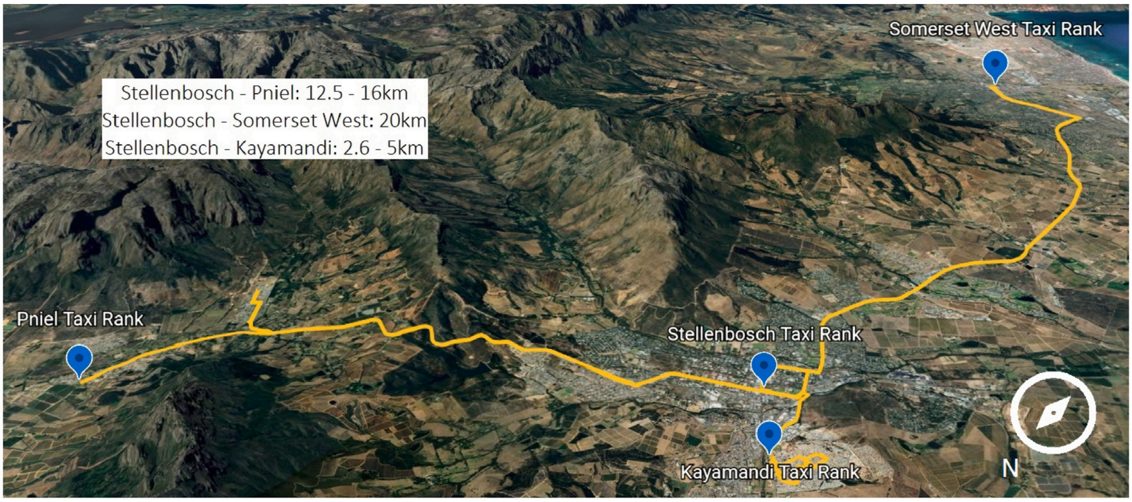
To determine the micro-mobility behavior of the minibus taxis, six tracking devices were used to record GPS data to an SD card at a frequency of 1 Hz. To account for GPS error, the GPS data was filtered before being written on the SD card; only complete and correct signals were taken and the rest were discarded.

The six recording devices are based on the Arduino platform and powered from alkaline battery packs. The device can therefore operate independently of any other device during tests. The acquired data is separately processed after the completion of data recording. Field workers, appointed to take trips on the minibuses as standardized passengers and initiated data capture with the press of a button, which would be terminated once the vehicle had reached the destination. Each recorded trip creates an isolated file. This allows for different routes to be separately investigated and compared to other recordings made on the same route.

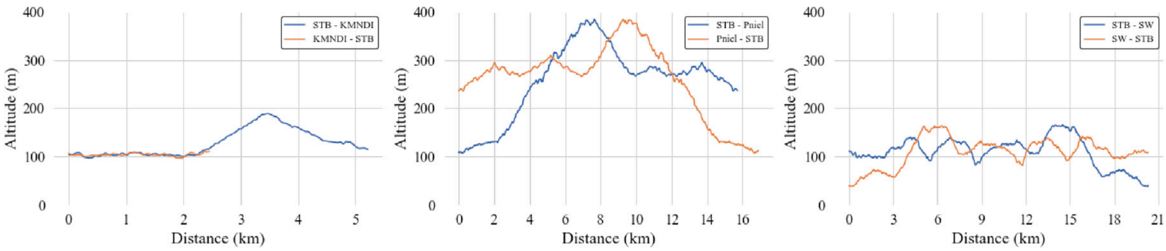
The velocity, $v[n]$, was obtained from the GPS sample's speed value. The elevation, $h[n]$, was obtained from the GPS sample's location value, which was used to lookup the elevation in the Earth Resources Observation and Science Center's dataset (USGS, 2022). This method was preferred because the GPS module's elevation values proved to be less consistent than using the location-based lookup. The displacement, $s[n]$, is calculated as the 3D geodesic distance from subsequent GPS locations. The slope angle (positive for incline and negative for decline), $\alpha[n]$, is calculated from subsequent displacements, $s[n-1]$ and $s[n]$, and corresponding elevations, $h[n-1]$ and $h[n]$.

2.3. Route selection

Vehicle energy consumption was expected to be dependent on driving conditions, specifically changes in elevation with or against gravity, driving speed in inter-city transport, and the stop-start nature of urban driving (Al-Wreikat et al., 2021). Time of travel during the day was also expected to have an impact. Therefore, data was captured on three different types of route: urban, hilly, and inter-city. Data for each route was collected at three distinct times of the day: morning (before 11:30 AM), afternoon (11:30 AM–4:30 PM), and evening (after 4:30 PM). No more than three trips was taken on any individual minibus taxi, and over 40 different minibus taxi drivers were used in data collection to capture variation in minibus taxi driver driving style. All trips are recorded from, or back to, Bergzicht Taxi Rank, the main taxi rank in Stellenbosch, South Africa, and are shown in Fig. 4. The urban route is from Stellenbosch to Kayamandi taxi rank (STB - KMNDI). Kayamandi is a neighborhood of Stellenbosch with a proportion of taxi riders. The routes connecting the taxi ranks have a speed limit of 60 km/h and consist of bidirectional roads only. The to



(a) Bird's eye view of the three routes selected for data capture [47].



(b) Altitude plots for the three routes. Outbound trips are shown in orange, and inbound trips are shown in blue. The trip from Stellenbosch to Kayamandi includes an extra circular segment inside the township, making the trip roughly twice as long as its counterpart.

Fig. 4. The routes used to assess vehicle energy consumption in and around Stellenbosch and their altitude profiles (see Google (2022)).

Table 2
Summary of the number of trips recorded in data acquisition phase.

	Route						
	Stellenbosch Kayamandi	Kayamandi Stellenbosch	Stellenbosch Pniel	Pniel Stellenbosch	Stellenbosch Somerset West	Somerset West Stellenbosch	Total
Distance:	5 km	2.6 km	15 km	15 km	20 km	20 km	
Avg. Time:	18 min	6 min	28 min	23 min	29 min	27 min	
Speed Limit	60 km/h	60 km/h	80 km/h	80 km/h	100 km/h	100 km/h	
Morning	4	4	3	3	3	3	20
Afternoon	3	5	4	4	4	4	24
Evening	3	3	3	2	4	3	18
Total	**10**	**12**	**10**	**9**	**11**	**10**	**62**

and from routes vary slightly, since the taxis pass the Kayamandi taxi rank and completes a circle route through the neighborhood, dropping off passengers near their homes, before returning to the Kayamandi rank for the routine stop. The distances to and from Kayamandi taxi rank are 5 km and 2.6 km respectively.

The hilly route is one with a steep incline from Stellenbosch to Pniel (STB - Pniel) and crosses the Helshoogte mountain pass – a steady climb of approximately 300 m in 7 km. This 12.5 km route, with a small deviation on some trips, features a long accent followed by a short decent, and has a speed limit of 80 km/h.

The inter-city route is a 20 km route from Stellenbosch to Somerset West (STB - SW). The speed limit on this provincial road, which is predominantly a dual carriageway, is 100 km/h.

Table 2 shows the breakdown of data collected after filtering for trips with missing or anomalous data. The goal was to collect at least three trips for each route and time of day combination to establish a reasonable depth of data for each driving context. After data cleaning, this was successful for all route/time of day combinations except for Pniel - STB/'Evening', which was left with 2 trips. The final number of trips used for analysis was 62.

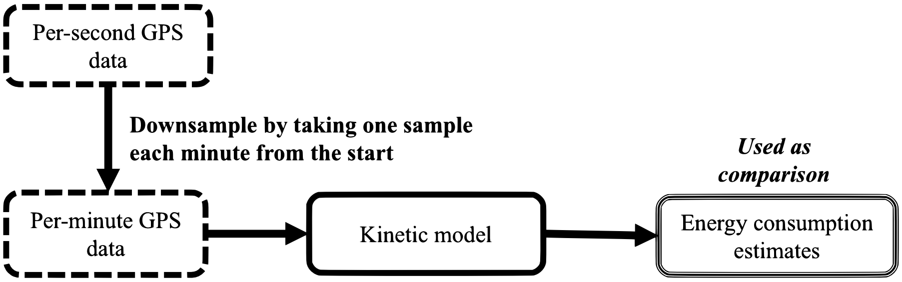


Fig. 5. How an equivalent per-minute dataset created from the raw per-second data and used to construct energy consumption results. These results are then compared with the results generated by the process in Fig. 2.

Table 3
Breakdown of average energy consumption, net elevation change, average speed, maximum speed, and average absolute acceleration by route.

Route	Energy consumption (kWh/km)	Net elevation change (m)	Average speed (km/h)	maximum speed (km/h)	Average acceleration (absolute) (m/s^2)
STB - KMNDI	0.36	12.7	15.9	64.2	0.52
KMNDI - STB	0.37	1.3	26.1	70.7	0.57
STB - SW	0.37	−68.1	42.5	101.4	0.50
SW - STB	0.42	67.5	44.2	98.5	0.49
STB - Pniel	0.47	126.7	33.3	84.0	0.48
Pniel - STB	0.33	−127.2	40.4	90.2	0.52

The dataset is publicly available in this repository at Mendeley data: https://doi.org/10.17632/xt69cnwh56.1. A data-in-brief has been submitted for review.

2.4. Comparing per-second and per-minute GPS data input

To evaluate model performance using per-second versus per-minute data, we must construct an equivalent per-minute dataset. To do so, we first downsample our per-second data to one GPS observation per minute using python pandas resample function. We start from the first sample, and take one sample every minute thereafter. We then feed this per-minute data into our kinetic model to generate energy consumption estimations, and compare that to our original results. This process is shown in Fig. 5.

3. Results

3.1. Evaluation of energy consumption estimates

In this section, we first present our energy consumption estimates in kWh/km and evaluate the differences seen between the various driving conditions and times of day. We then compare our results to values previously seen in the literature.

3.1.1. Impacts of route and time of day

Fig. 6 shows the distributions of energy consumption estimates for the three routes considered, in both directions, for each time of day (yielding a total of $3 \times 2 \times 3 = 18$ distributions). A dashed teal line indicates the overall mean of 0.39 kWh/km. This figure demonstrates how energy consumption is largely determined by the characteristics of the route on which a minibus is driving. The route characteristic with the clearest effect is elevation change, as evidenced by the difference observed between the steep incline and steep decline directions of the hilly route (to/from Pniel). Additionally, it appears that the inter-city travel (to/from SW) was slightly more energy intensive per kilometer on average than urban travel (to/from KMNDI).

Although distinctive patterns are observed between routes, the results do not show a consistent effect of time of day on energy consumption on any route. The estimates are within a fairly tight range for each route, with a mean inter-route standard deviation of 0.02 kWh/km, which supports the hypothesis the per-second sampling strategy leads to high fidelity estimates.

Table 3 summarizes the mean energy consumption and some mobility characteristics observed for each of the routes. This table shows the variation in the physical nature of the routes and some factors that characterize traffic flow (i.e. speed and how often the vehicle must stop). The urban route had the highest average absolute acceleration and lowest speeds, belying the stop-start nature of urban travel. The effect of this driving pattern on energy consumption is explored in later sections. The average of absolute acceleration is used to account for total acceleration and deceleration.

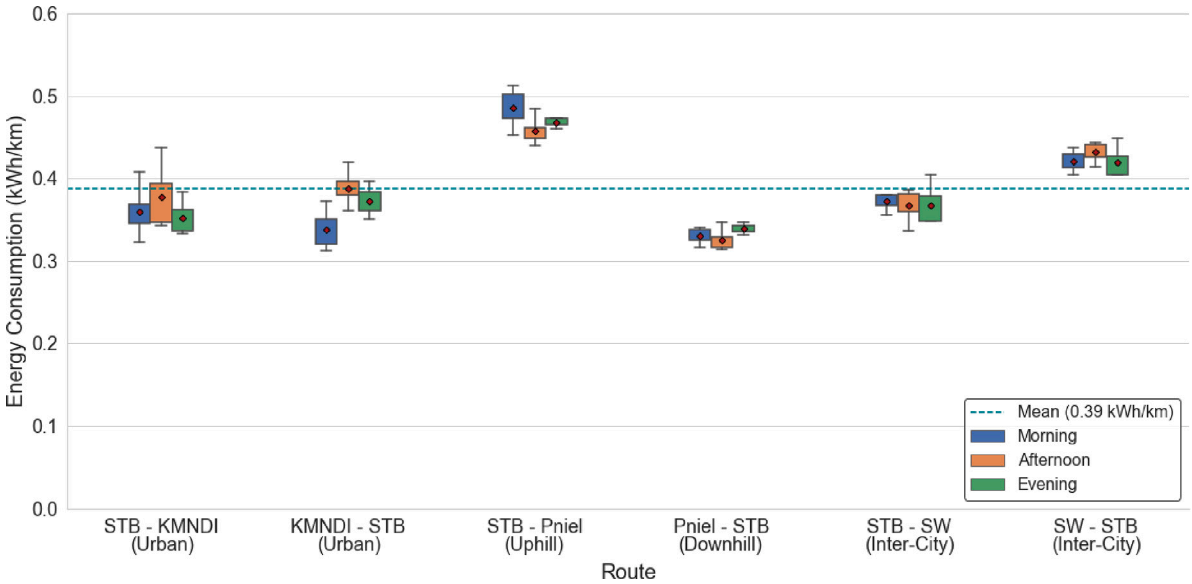


Fig. 6. Distribution of energy consumption for each route in the morning, afternoon, and evening. The overall mean energy consumption of 0.39 kWh/km is indicated with a dashed teal line.

Table 4
Comparative energy consumption (kWh/km) values seen in the literature and from real-world OEM vehicle specs.

Literature source	Energy consumption (kWh/km)	OEM source	Energy consumption (kWh/km)
Abraham et al. (2021)	0.93	Higer H5C EV (Higer, 2020)	0.36
Cignini et al. (2020)	0.70	Dongfeng E-Mini Van (Dongfeng, 2021)	0.27
Collett et al. (2021)	0.50	Ruivii Toano (Ruivii, 2021)	0.26

3.1.2. *Results in context*

To contextualize our energy consumption results, we compared them to values seen in the literature and real world manufacturer-stated kWh/km for similar vehicles. Table 4 lists these values and their sources. While some of these vehicles lack data on certain parameter values in Table 1, they share a physical and mechanical profile to the minibus taxis in question, which is why they were chosen as fair comparison vehicles.

Fig. 7 displays our distributions of energy consumption results for each route in comparison to the values in Table 4. Since time of day was found to have no effect on energy consumption in Section 3.1.1, we aggregate times of day by route for a more concise comparison.

Our results are slightly higher than the energy consumption of small passenger vans on the market, which range from 0.2–0 0.36 kWh/km. This is logical given that taxi drivers are notorious for aggressive driving, which increases energy consumption (Al-Wreikat et al., 2021; Faria et al., 2019; Murphey et al., 2009; Eno Akpa et al., 2019). Seeing values that are similar to but slightly higher than manufacturer given values lends confidence to the accuracy of our estimates.

In contrast, our estimates – barring one uphill trip at 0.51 kWh/km – were lower than the comparative values from literature, which ranged from 0.50–0.93 kWh/km. Notably, the highest results previously in the literature from Abraham et al. (2021) were furthest away from our estimates, despite being the only other estimates constructed from GPS tracking data. Their method uses an EV fleet simulator called 'ev-fleet-sim', which is based on the micro-traffic simulator Simulation of Urban MObility (SUMO) and a matching EV kinetic model developed by Kurczveil et al. (2014). Ev-fleet-sim intakes per-minute GPS waypoints, and uses the SUMO routing function to simulate per-second mobility data on an Open Street Maps virtualization of the road infrastructure. This per-second data is run through the kinetic model to estimate energy consumption. A preliminary assessment of their method has shown the inaccuracy of the physical infrastructure virtualization and of the per-second driving simulation as potential reasons for the high consumption rates. Furthermore, in contrast to our simple first-principle kinetic model, theirs is based on the moment of inertia of internal elements, which could also contribute to the difference. A detailed analysis of their method falls beyond the scope of this paper, and is left to future research.

3.2. *Implementation of results*

The techno-economics of paratransit electrification depend on vehicle energy consumption. Since EV energy demands are known to vary across driving conditions (Al-Wreikat et al., 2021), we aim to provide recommendations in multiple conditions. Specifically,

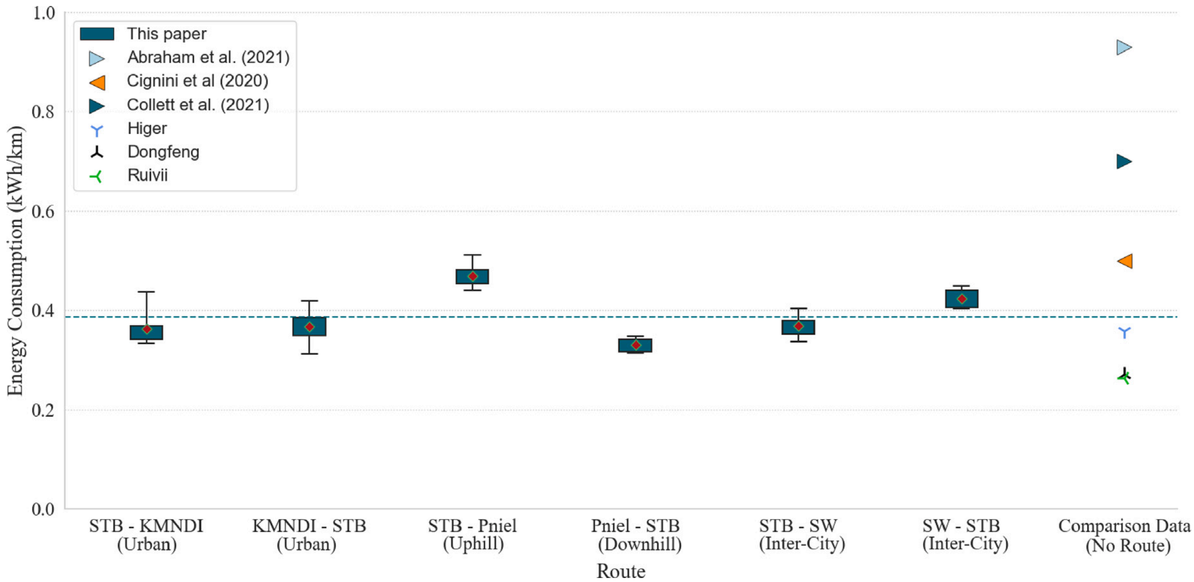


Fig. 7. Distributions of energy consumption, compared to state-of-the-art from literature and real world vehicle specs. The dashed teal line represents the mean of the estimates constructed from the per-second data (0.39 kWh/km).

Table 5
Recommended energy consumption (kWh/km) for calculating range in different driving conditions, based on highest energy consumption observed in each condition. Example ranges (km) using different battery sizes (kWh) are given. Ranges are calculated as 85% of the physical limit that a battery would be able to provide at the given energy consumption.

Energy consumption (kWh/km)	Urban	Uphill	Downhill	Inter-city
	0.44	0.51	0.35	0.45
Range (km) - 50 kWh battery	97	83	122	94
Range (km) - 75 kWh battery	145	125	182	142
Range (km) - 100 kWh battery	193	167	243	189

we provide recommendations for values to use in urban, inter-city, and hilly (uphill and downhill) conditions. Uphill and downhill are treated separately because if a taxi goes up before a charge and down thereafter, you need two separate numbers for the energy budget as the energy demands are significantly different between the two directions.

3.2.1. Operations planning

Electric minibus taxis in South Africa will need to travel up to 250 km before stopping to recharge (Abraham et al., 2021). Typically, taxis will operate continuously during the morning and evening rush hour periods from roughly 6 AM to 10 AM and 5 PM to 9 PM respectively, and operate at a diminished rate throughout the middle of the day (Abraham et al., 2021; Booysen et al., 2022). An electric taxi must be able to sustain operations during the rush hour periods without stopping to charge, lest the driver and owner lose out on profits. Economic loss would discourage uptake of electric taxis and thus the electrification transition, so the highest estimated energy consumption values must be used for planning operations to guarantee continuous operation. However, it is not economic to supply every e-taxi with the largest possible battery pack An excessively large pack would be unnecessarily expensive and impact payload. A taxi that operates over long distances in a hilly region will not necessarily have the same battery requirements as a taxi that operates for short distances in an urban environment. With this in mind, we provide recommendations in Table 5 that are specific to different conditions. Ranges are calculated as 85% of the battery's physical limit for a given rate of energy consumption, since neither overly deep discharge nor maxing out on charge are good for the health of the battery.

3.2.2. Grid impact and emissions

In an energy-constrained and coal-dependent context, the interactions between EVs and the electricity grid must be carefully considered when planning the deployment of an EV fleet. Sufficient charging infrastructure and clean energy generation must be available to sustain the fleet, as EV charging can have a substantial impact on the grid (Buresh et al., 2020). Additionally, policymakers are often interested in the emissions saving of the EV equivalent to an ICE operation, so it is useful for stakeholders to be able to use our estimates to model the greenhouse gas (GhG) emissions impact of transitioning from the status quo an electric fleet.

Since operational (not life-cycle) GhG emissions for EVs depend on their interaction with the electricity grid, the same energy consumption values should be applied for calculating both grid and GhG emissions impact of the fleet. In contrast to operations

Table 6
Recommended energy consumption (kWh/km) to use for calculating e-taxi impact on the electricity grid in different driving conditions, based on average energy consumption estimated for each route type.

	Urban	Uphill	Downhill	Inter-city
Energy consumption (kWh/km)	0.37	0.47	0.33	0.40

planning, using an average energy consumption is appropriate for fleet-wide energy demand calculations. Table 6 breaks down the recommended energy consumption values for e-taxi to grid interaction.

If information on driving conditions is unavailable, we recommend using the overall mean energy consumption of 0.39 kWh/km.

Given approximately 300,000 minibus taxis in South Africa (Transaction Capital, 2021), for an average daily traveling range of 100–200 km, this energy consumption would lead to 11.7–23.4 GWh additional electricity burden per day (1%–3% of national daily electricity production) and 4.3–8.6 Twh annually. Even a local fleet of only 25 taxis that travel 200 km per day, which is typical in the region (Abraham et al., 2021), would require 1.95 MWh per day and 0.71 GWh per year. Such additional loads could impose great burdens on both local grids and the national electricity supply, which is concerning since many cities around sub-Saharan Africa already face electrical supply problems and frequent blackouts. With the goal of facilitating safe and efficient large scale transitions to electric fleets around the region, more research is needed that builds on the values provided in this study for how these taxis can effectively integrated with local power systems.

3.3. Effect of mobility characteristics and physical constants on energy consumption

The analysis in this section aims to enable our energy consumption estimates to be generalized for contexts not explicitly captured in our dataset, and for different sets of vehicle physical constant parameters.

For the former, we quantify the effects of net elevation change (m), average speed (m/s), and average absolute acceleration (m/s^2) on energy consumption using Ordinary Least Squares (OLS) regression. The absolute value of acceleration is taken to retrieve the 'total' effect of acceleration and deceleration. For the latter, we compute the sensitivity of our kinetic model to each of the vehicle parameters from Table 1.

Elevation change, speed, and acceleration were chosen as factors to explore because they directly correspond to forces in the kinetic model, but are also comprehensible real-world measurable variables. Respectively, they correspond to slope drag, aerodynamic drag, and vehicle force. Elevation change is accounted for in slope drag via slope angle (Eq. (1d)), speed enters directly into aerodynamic drag (Eq. (1b)), and acceleration is derived in vehicle force as change in velocity (Eqs. (2) and (3)).[1]

Despite slope angle being the direct measured factor that enters into the slope drag equation, net elevation change was chosen over average slope angle because it is an eminently more useful measurement. It is far easier to retrieve net elevation change than average slope angle for routes around the world, and we found that net elevation change approximates slope angle very well. Quantifying the effects of these factors on energy consumption is only useful if the results can be used by stakeholders and other interested parties. Net elevation change and speed limits are easily obtainable features of a route via a quick google search. Including average absolute acceleration in the regression model controls for the effects of 'driver behavior' and other micro-traffic flows on energy consumption that are difficult to ascertain prior to collecting GPS tracking data.

Not only do elevation change, speed, and acceleration correspond to forces experienced by the vehicle in the kinetic model, they are the only measured factors that do so. Consequently, we can be sure that our regression model captures the majority of the variation seen in energy consumption in the data, giving us confidence in the results of the regression. This is borne out by the R-squared seen in our OLS regression of 0.98. R-squared explains the proportion of variation seen in an outcome variable that is captured by the explanatory variables included in a regression model.

3.3.1. Effects of elevation change, speed, and acceleration

To quantify the effects of net elevation, average speed, and average absolute acceleration on energy consumption, we construct an Ordinary Least Squares (OLS) regression model. OLS regressions estimate model parameters by minimizing the sum of the squared differences between the observed outcome variable (energy consumption in this case), and the outcome predicted by a linear function of the explanatory variables. Our OLS regression results are presented in Table 7.

The coefficients in a regression specify how much the outcome variable changes for a 1 unit increase in each factor:

$$\text{coeff} = \frac{d_{\text{outcome}}}{d_{\text{factor}}}$$

Specifically, the regression in Table 7 shows that an increase of 1 m in net elevation change is predicted to increase energy consumption for a trip by 0.0007 kWh/km, an increase in average speed of 1 km/h is predicted to increase the average energy consumption for a trip by 0.0042 kWh/km, and an increase of 1 m/s^2 in average acceleration during a trip is predicted to increase energy consumption by 0.4540 kWh/km. These coefficients on net elevation change and average speed are the most useful because they are the two factors easiest to ascertain in the absence of GPS tracking data.

[1] The last force in the model, rolling resistance (Eq. (1c)), does not have a direct mobility characteristic analog in our model. While slope angle does enter into the rolling resistance equation, this force is primarily a consequence of the physical parameters of mass and rolling resistance coefficient.

Table 7
OLS Regression model of average speed, net elevation change, and average absolute acceleration on energy consumption. The units on each coefficient are a ratio of the units of the corresponding factor (m, m/s, and m/s^3) to the units on energy consumption (kWh/km). The standard deviation of each coefficient is shown in brackets beneath the coefficient. The percentage of the variation in the data captured by the model is shown in the R-squared statistic.

	Average trip energy consumption (kWh/km)
Net elevation change (m)	0.0007*** (kWh/km per m) [0.0001]
Average speed (m/s)	0.0042*** (kWh/km per m/s) [0.0006]
Average (absolute) acceleration (m/s^2)	0.4540*** (kWh/km per m/s^2) [0.0395]
R-squared	0.9780
R-squared Adj.	0.9769

***Indicates 99.9% confidence that the coefficient is statistically significant.

Table 8
Standardized OLS Regression model of net elevation change, average speed, and average absolute acceleration on energy consumption. Each coefficient represents how many standard deviations (σ) energy consumption changes for a 1 standard deviation change in the corresponding factor. The standard deviation of each coefficient is shown in brackets beneath the coefficient.

	Average trip energy consumption (σ)
Net elevation change (σ)	0.9037*** (σ kWh/km per σ m) [0.0543]
Average speed (σ)	0.3735*** (σ kWh/km per σ m/s) [0.0542]
Average (absolute) acceleration (σ)	0.1229** (σ kWh/km per σ m/s^2) [0.0536]

**Indicate 99% confidence in the statistical significance of the coefficient.
***Indicate 99.9% confidence in the statistical significance of the coefficient.

One potentially confusing aspect of these results is the high coefficient on acceleration compared to the other two factors. This is caused by a discrepancy in units. Where net elevation change and average speed vary by tens and hundreds, average acceleration is typically a value between 0 and 1 m/s^2.

To more clearly understand the relative effect of each factor on energy consumption, we construct an standardized OLS regression model with the same factors as in the previous section, as shown in Table 8. The coefficients in a standardized regression show by how many standard deviations (σ) the outcome variable changes for a 1 σ change in an explanatory variable.

The asterisks next to the regression coefficients indicate the p-value of the coefficient. A p-value less than 0.05 is generally considered statistically significant. We mark $p < 0.001$ by ***, $p < 0.01$ by **, and $p < 0.05$ by *. We find all four factors to be statistically significant to $p < 0.001$. The number in parenthesis below each coefficient is that coefficient's standard deviation. The resultant R-squared of 0.98 demonstrates that the linear model with these four variables captures nearly all of the variation in energy consumption seen in our dataset.

The results from these regressions using the non-standardized model are informative for anyone looking to deploy electric minibuses in other locations with different mobility characteristics. While high sampling frequency GPS data would be ideal for analyzing vehicle energy consumption, it can be difficult or impossible to come by in some circumstances.

One apparent incongruity between the results in the two regression models is the level of significance of the coefficient on acceleration. It shows *** in the non-standardized model, indicating $p < 0.001$, and ** in the standardized model, indicating $p < 0.01$. This is caused by a discrepancy in units between the two models. Specifically, the coefficient on acceleration in the non-standardized model predicts that energy consumption will change by 0.4540 kWh/km from a 1 m/s^2 change in average absolute acceleration, whereas the coefficient in the standardized model predicts the that energy consumption would change by 0.1229 σ given a 1 σ change in average absolute acceleration.

3.3.2. Effects of model parameters

Next, we assess the sensitivity of the kinetic model to the vehicle physical parameters. The results of the sensitivity analysis are summarized in Table 9. The first column shows the relative sensitivity of the model to each parameter as the percentage change in energy consumption for a 1% change in a given parameter. The second column shows the absolute change in energy consumption in kWh/km for a 1% change in a given parameter. The table shows that the model is most sensitive to powertrain efficiency, followed by the minibus weight, then regenerative braking efficiency and rolling resistance coefficient, then drag coefficient and finally vehicle frontal surface area.

The parameters used for estimations in this paper, listed in Table 1, were carefully chosen from various sources in the literature, but reasonable arguments can be made for different values, and they do vary in the literature (Abraham et al., 2021). Similar to the results of the OLS regression, the results of the sensitivity analysis are useful to determine the effect of an individual change

Table 9
Sensitivity of kinetic model to physical parameters.

Parameter	Sensitivity ($\frac{\%\text{ change output}}{\%\text{ change input}}$)	Sensitivity ($\frac{\text{absolute change output}}{\%\text{ change input}}$) (kWh/km)
Minibus weight (m_v)	0.84	0.003
Rolling resistance coefficient (c_{rr})	0.48	0.002
Drag coefficient (c_d)	0.15	0.001
Vehicle's front surface area (A)	0.15	0.001
Regenerative braking efficiency (μ_{rg})	−0.50	−0.002
Powertrain efficiency (μ_v)	−1.48	−0.006

Table 10
Values used to compute sensitivities of the physical parameters of the kinetic model.

Parameter	Values		
Minibus weight (kg)	2900	3900	4900
Rolling resistance coefficient	0.01	0.02	0.03
Drag coefficient	0.24	0.36	0.48
Frontal Area (m^3)	3.50	4.00	4.50
Regenerative braking efficiency (%)	0.50	0.65	0.80
Powertrain efficiency (%)	0.85	0.90	0.95

in each parameter on the energy consumption of taxis. For example, the weight of a taxi could easily change by several hundred kilograms, depending on the weight of passengers and their luggage (aka payload). For example, if the minibus weight parameter was adjusted up by 10% from the value this paper uses (i.e. by 390 kg), to 4290 kg, then overall mean energy consumption would be expected to increase by 8.4% (i.e. 0.03 kWh/km) to 0.42 kWh/km.

To create Table 9, overall mean energy consumption was simulated with three different input values for each parameter. The center value for each parameter are the values from Table 1. The sensitivities were gathered running the model with each parameter varied twice away from the center, while all other parameters were held equal. They varied once above and once below the center value, allowing two equally weighted sensitivities to be calculated for each parameter. The sensitivity reported in Table 9 is the average of the two. The input values are given in Table 10. The values are chosen based on a range of likely operating conditions and values seen in the literature.

3.4. Evaluation of sampling frequency: Per-second vs per-minute data input

This section explores the energy consumption estimates from per-second data in comparison to energy consumption estimates from an equivalent per-minute dataset.

Throughout the section, the per-second data and associated results are colored in teal and the per-minute equivalent in orange.

The differences between the per-second and per-minute datasets are visualized in Fig. 8, which shows the elevation change (m). speed (km/h), and acceleration (m/s^2) profiles for an example trip from Pniel - STB. Each plot is accompanied by a chart that quantifies the percentage of that mobility characteristics' signal that is captured in the per-minute sampling frequency. To construct these charts, we perform a Fast Fourier Transformation (FFT) of the per-second data, retrieve the corresponding power spectral density, and compute the percentage of signal captured in sampling frequencies up to $\frac{1}{60}$ Hz (orange bars). The teal bars show the percentage of signal captured by sampling frequencies greater than $\frac{1}{60}$ Hz, up to $\frac{1}{2}$ Hz (the Nyquist frequency for per-second data (Shannon, 1949)).

These plots show that compared to the per-second data, the per-minute data captures 28.4% of elevation change, 83.4% of speed, and 8.6% of acceleration data. Speed is well approximated in the per-minute data, but the elevation change and acceleration data are much noisier, thus requiring a greater sampling frequency to fully capture. The driving cycle from the per-minute dataset is much smoother, and is unable to account for the micro-mobility movements of the vehicles.

While data for all trips can be found in the Mendeley repository, the particular trip used for this example in Fig. 8 is attached as an addendum along with the code for the kinetic model at: https://github.com/ChullEPG/Bumpy-Ride, so the reader can easily explore the for the example trip at their convenience.

Table 11 characterizes the net elevation change, total absolute elevation change, average speed, and average acceleration profiles by route for the per-minute data in comparison to the per-second data. The table shows that average speed and net elevation change are well characterized in the per-minute data, but crucially, average absolute acceleration and total absolute elevation change are consistently underestimated. Total absolute elevation change was included in this table because, while per-minute data captures overall trends in elevation change, it misses out on the climbs and dips that occur between minutes and contribute to total absolute elevation change. The knock-on effect these oversights have on vehicle energy consumption is explored later in the section. Together, Fig. 8 and Table 11 reinforce the hypothesis that per-minute data does not capture micro-mobility patterns.

The comparison of energy consumption estimates constructed from the per-minute and per-second data is shown in Fig. 9. This figure shows that, compared to the per-second data, the per-minute data (a) consistently underestimates energy consumption

Fig. 8. Comparison of the elevation change, speed, and acceleration profiles in the per-second data (dashed teal lines) and per-minute data (solid orange lines) for an example route from Pniel - STB. Each minutely sample on the orange lines is marked with an X. The bar charts show the portion of signal that is captured in the per-minute sampling frequency (<1/60 Hz), and the portion of signal that it misses (>1/60 Hz) for each factor. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 11
Characteristics of factors known to influence efficiency in per-minute dataset (orange), compared to the analogous measurements in the per-second dataset (teal) from 3.

Route	Net elevation change (m)	Total elevation change (absolute) (m)	Average speed (km/h)	Average acceleration (absolute) (m/s^2)
STB - KMNDI	17.5 (12.7)	25.2 (92.8)	14.0 (15.9)	0.07 (0.52)
KMNDI - STB	−3.5 (1.3)	169.7 (230.8)	24.4 (26.1)	0.09 (0.57)
STB - SW	−70.5 (−68.1)	423.3 (736.4)	40.3 (42.5)	0.14 (0.50)
SW - STB	67.7 (67.5)	409.6 (755.0)	43.1 (44.2)	0.11 (0.49)
STB - Pniel	126.9 (126.7)	481.2 (797.6)	32.6 (33.2)	0.10 (0.48)
Pniel - STB	−127.2 (−127.2)	457.4 (793.2)	40.6 (40.4)	0.10 (0.52)

and (b) estimates with greater variance. The average overall energy consumption is 31% less than the per-second estimates - 0.27 vs. 0.39 kWh/km - and the average inter-route standard deviation was 0.09 kWh/km, greater than three times higher than the 0.02 kWh/km seen in the per-second estimates. Across all trips, the coefficient of variation for the per-minute estimates was 0.48, more than three times higher than the 0.13 seen in the per-second data.

That said, the per-minute data does preserve the relative effect of elevation change, as evidenced by the relatively high energy consumption for the uphill route (STB - Pniel) and low energy consumption for the downhill route (Pniel - STB). This is because per-minute data characterizes net elevation change for the route relatively accurately, despite missing many small dips and climbs that contribute to absolute elevation change and overall energy consumption. The trips with the lowest sample size, the short urban trips to and from Kayamandi (KMNDI), have the greatest variance. Furthermore, the urban route has lower elevation change and speed limits than the hilly or inter-city routes, as well as greater average absolute acceleration. Therefore, following the analysis in Section 3.3, vehicle micro-mobility patterns will have an increased effect on energy consumption in comparison to the other routes. In addition, an analysis of jerk (m/s^3) and number of stops per km along all of the routes also showed that the greatest vehicle jerk and most stop/start driving cycles occurred in the urban context. Therefore, the weakness of the per-minute dataset – its inability to capture micro-mobility patterns – manifests most greatly in the urban environment where environmental characteristics tend to dominate energy consumption less and driver behavior plays a greater role. These findings suggest that per-minute data is least suitable for estimating vehicle energy consumption in urban or residential contexts, which lend themselves to low elevation change, low speed limits, and a larger number of high acceleration/deceleration events (i.e. aggressive driving behavior).

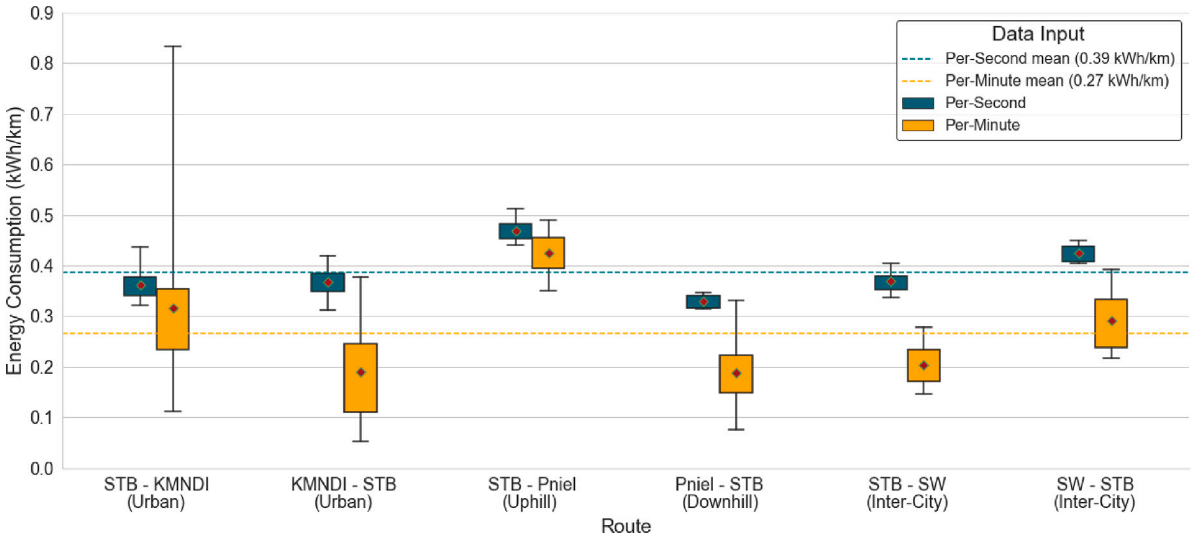


Fig. 9. Distribution of energy consumption estimates for each route when the kinetic model is inputted with per-second data versus per-minute data. The dark red diamond in each box represents the mean for that route. The teal and orange dashed lines represent the mean across all trips for the per-second and per-minute inputs respectively. There is one outlier trip at 0.83 kWh/km in STB - KMNDI. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

One question remains and that is to quantify where the underestimate in the per-minute data comes from. To do so, we must return to the physics of the kinetic model.

There are six physical forces with distinct effects in the kinetic model, four external, dubbed ‘environmental forces’, and two forces generated by the vehicle, dubbed ‘vehicle forces’. The breakdown is as follows:

- Environmental forces: Propulsive slope drag (downhill), resistive slope drag (uphill), aerodynamic drag, and rolling resistance.
- Vehicle forces: Propulsive (motor to wheels), braking (wheels to motor)

At a fundamental level, the vehicle must generate enough propulsive or braking force to overcome the environmental forces it is experiencing and achieve the driver’s desired acceleration or deceleration. Thus, the energy expended by the battery to overcome the environmental forces and achieve the desired acceleration/deceleration can be attributed to environmental forces, and driver behavior. If the per-minute data underestimates the environmental forces and/or the desired acceleration of the driver, then it will underestimate energy consumption.

To quantify where the underestimates in the per-minute model come from, we break down the contribution of each environmental and vehicle force to energy consumption in kWh/km, and perform a FFT to assess how much of this energy is captured by the per-minute sampling frequency. These results are in Fig. 10. The top bars represent how much energy in kWh/km are captured in the per-minute data for each force, and the bottom bars represent how much energy the per-minute data misses. The relative size of the bottom bars is not insignificant, thus belying the source of the underestimates from the per-minute dataset input seen in Fig. 9.

Table 12 further breaks down the in the energy contribution of each force for each dataset, by route. The largest differences are primarily seen in the two vehicle forces, and the slope drags. This reflects how the per-minute dataset underestimates average absolute acceleration and total absolute elevation change as indicated in Table 11. Net elevation change is well captured, which is reflected in how we can see its effect preserved in Fig. 9, in the difference between the uphill and downhill routes. However, missing out on some total absolute elevation change means it underestimates the total contribution of slope drag to energy consumption.

This section has shown that per-minute data fails to capture vehicle micro-mobility patterns, thus leading to underestimates of vehicle energy consumption. Per-minute estimates were also found to be high in variance. This supports the initial hypothesis that motivated gathering per-second GPS data: that the per-second data would provide a more accurate picture of paratransit micro-mobility patterns.

4. Conclusions

The aim of this paper is to provide high fidelity energy consumption estimates (kWh/km) for paratransit minibus taxi vehicles, the mainstay of transport in sub-Saharan Africa, in various driving conditions.

We estimate paratransit vehicle energy consumption to range from 0.29–0.51 kWh/km depending on driving condition. The estimates provided in this paper, based on 1 Hz sampling, are the highest fidelity estimates to date. The lower estimates in this paper compared to previous literature – which provided average estimates from 0.50–0.93 kWh/km – imply that the estimated battery

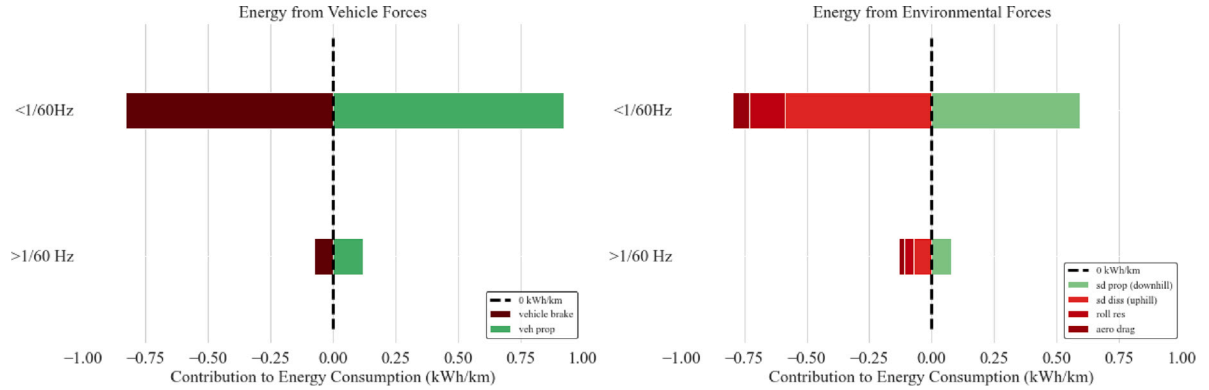


Fig. 10. FFT analysis disaggregated by force type. Shows how much energy for each force is captured by sampling frequencies up to once per-minute (<1/60 Hz) and greater than once per-minute (>1/60 Hz), for all trips. The bottom bar in each graph represents the amount of energy not captured in the per-minute data.

Table 12
Percentage difference between the amount of energy captured in the per-minute and per-second datasets, for each vehicle and environmental force that the vehicle experiences.

Route	Percentage difference between datasets					
	Vehicle braking	Vehicle propulsive	Propulsive slope drag (downhill)	Dissipative slope drag (uphill)	Aerodynamic drag	Rolling resistance
KMNDI - STB	−83.5	−61.8	−51.8	−78.5	6.2	12.2
STB - KMNDI	−52.4	−28.7	8.0	9.9	−1.6	10.0
STB - Pniel	−70.1	−25.9	−34.5	−14.6	15.3	12.3
Pniel - STB	−51.7	−46.8	0.1	−43.7	30.0	22.3
STB - SW	−64.8	−51.1	−28.8	−48.8	1.8	0.6
SW - STB	−75.5	−42.0	−37.0	−46.4	6.4	3.0
Mean	−66.3	−42.7	−24.0	−37.0	9.7	10.1

capacity requirements and thus cost and grid impact of electric minibus taxis are less than previously suggested, an encouraging result.

Our recommendations can help stakeholders determine e-taxi battery capacity requirements, plan operations, and estimate the electricity demand and emissions impact of deploying e-taxis in varying contexts. Depending on the range requirements and driving condition of its operating region, an e-taxi in South Africa could feasibly require batteries from 50 up to over 100 kWh in size—important for range and cost. Given the typical operational range of the taxis, their electrification could impose large burden on local and national electricity grids in this region. Future research based on the values and recommendations provided in this study needs to be done to ensure safe, efficient, and effective transitions to electric fleets of paratransit vehicles.

Energy consumption estimates may be needed for dozens or hundreds of different driving conditions around sub-Saharan Africa for planning transitions to electric paratransit. It is practically impossible to capture every combination of these factors seen across the continent in one dataset; however, a benefit of capturing variations in environmental and mobility profiles with high resolution GPS data is that we can quantify the effect of various mobility characteristics on energy consumption. The coefficients in the non-standardized regression model in Section 3.3.1 can be used to extrapolate our results to routes outside of our dataset that may have different elevation and speed profiles. Similarly, a range of likely operating conditions for e-taxis are possible, and various values are seen in the literature for the parameters on vehicle kinetic models. Accordingly, we quantified the sensitivity of our model to each parameter to allow the effect of adjusting each parameter on our model to be ascertained.

The results show that although per-minute data captures the environmental characteristics of a trip, it misses out on the micro-mobility patterns of a vehicle, leading to underestimates in vehicle energy consumption, and high variance in estimates.

The research done here is critical to planning safe and effective. However, given that the dataset in this study is geographically limited, research is needed to extrapolate the results to other regions in sub-Saharan Africa. Our results lay the foundation for future work to improve micro-traffic simulators to accurately predict paratransit micro-mobility patterns from origin–destination GPS data of a low sampling frequency. Given the commonality of per-minute GPS data in this field, and the cost and labor requirement of gathering quality per-second data, such improvements would be useful for constructing robust estimates of taxi energy consumption all over sub-Saharan Africa with sparse data. Furthermore, once electric minibus taxis come into usage in South Africa, future research could provide a ground-truth for the kWh/km values constructed here, and guide the construction of fitted models that estimate electric minibus energy consumption from ICE minibus mobility data. Lastly, future work on the effect of variation in individual driving style and on representative driving cycles for paratransit vehicles would enhance the literature on paratransit mobility in the developing world, which is understudied and poorly understood compared to analogous mobility sectors in the developed world.

CRediT authorship contribution statement

Christopher Hull: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – original draft, Visualization. **J.H. Giliomee:** Methodology, Software, Formal analysis, Investigation, Data curation, Writing – review & editing, Visualization. **Katherine A. Collett:** Conceptualization, Writing – review & editing. **Malcolm D. McCulloch:** Conceptualization, Methodology, Writing – review & editing, Supervision. **M.J. Booysen:** Conceptualization, Methodology, Formal analysis, Resources, Writing – review & editing, Visualization, Supervision, Project administration.

Data availability

The data is published in a Mendeley Data repo that is linked in the manuscript and in the submission.

Acknowledgments

M.J. Booysen and J.H. Giliomee acknowledge financial support from MTN South Africa (Contract S003061) and Eskom (Tertiary Education Support Programme). All authors acknowledge financial support from World Bank project Paratransit Decarbonization in South Africa.

References

Abraham, C.J., Rix, A.J., Ndibatya, I., Booysen, M.J., 2021. Ray of hope for sub-saharan africa's paratransit: Solar charging of urban electric minibus taxis in South Africa. Energy Sustain. Dev. 64, 118–127. http://dx.doi.org/10.1016/j.esd.2021.08.003.

Al-Wreikat, Y., Serrano, C., Sodré, J.R., 2021. Driving behaviour and trip condition effects on the energy consumption of an electric vehicle under real-world driving. Appl. Energy 297, 117096. http://dx.doi.org/10.1016/j.apenergy.2021.117096, URL: https://www.sciencedirect.com/science/article/pii/S0306261921005444.

Ayetor, G., Mbonigaba, I., Sunnu, A.K., Nyantekyi-Kwakye, B., 2021. Impact of replacing ICE bus fleet with electric bus fleet in Africa: A lifetime assessment. Energy 221, 119852.

Behrens, R., McCormick, D., Mfinanga, D., 2015. Paratransit in African Cities: Operations, Regulation and Reform. Routledge, pp. 1–311. http://dx.doi.org/10.4324/9781315849515.

Berzi, L., Delogu, M., Pierini, M., 2016. Development of driving cycles for electric vehicles in the context of the city of florence. Transp. Res. D 47, 299–322. http://dx.doi.org/10.1016/j.trd.2016.05.010, URL: https://www.sciencedirect.com/science/article/pii/S1361920916303017.

Booysen, M.J., Abraham, C.J., Rix, A.J., Ndibatya, I., 2022. Walking on sunshine: Pairing electric vehicles with solar energy for sustainable informal public transport in Uganda. Energy Res. Soc. Sci. 85, 102403. http://dx.doi.org/10.1016/j.erss.2021.102403, URL: https://www.sciencedirect.com/science/article/pii/S2214629621004904.

Brady, J., O'Mahony, M., 2016. Development of a driving cycle to evaluate the energy economy of electric vehicles in urban areas. Appl. Energy 177, 165–178. http://dx.doi.org/10.1016/j.apenergy.2016.05.094, URL: https://www.sciencedirect.com/science/article/pii/S0306261916306924.

Buresh, K.M., Apperley, M.D., Booysen, M.J., 2020. Three shades of green: Perspectives on at-work charging of electric vehicles using photovoltaic carports. Energy Sustain. Dev. 57, 132–140. http://dx.doi.org/10.1016/j.esd.2020.05.007, URL: https://www.sciencedirect.com/science/article/pii/S0973082620302441.

Cignini, F., Genovese, A., Ortenzi, F., Alessandrini, A., Berzi, L., Pugi, L., Barbieri, R., 2020. Experimental data comparison of an electric minibus equipped with different energy storage systems. Batteries 6 (2), http://dx.doi.org/10.3390/batteries6020026, URL: https://www.mdpi.com/2313-0105/6/2/26.

Collett, K.A., Hirmer, S.A., Dalkmann, H., Crozier, C., Mulugetta, Y., McCulloch, M.D., 2021. Can electric vehicles be good for sub-Saharan Africa? Energy Strategy Rev. 38, 100722. http://dx.doi.org/10.1016/j.esr.2021.100722.

Dongfeng, T., 2021. Dongfeng electric 15 seater mini van (RHD). URL: https://tara-dongfeng.co.uk/listings/dongfeng-ev-15seat-2-2/.

Ehebrecht, D., Heinrichs, D., Lenz, B., 2018. Motorcycle-taxis in sub-Saharan Africa: Current knowledge, implications for the debate on "informal" transport and research needs. J. Transp. Geogr. 69, 242–256. http://dx.doi.org/10.1016/j.jtrangeo.2018.05.006, URL: http://www.sciencedirect.com/science/article/pii/S0966692317303344.

Engineering Toolbox, 2008. Rolling resistance. URL: https://www.engineeringtoolbox.com/rolling-friction-resistance-d_1303.html.

Eno Akpa, N.E., Booysen, M., Sinclair, M., 2019. Fuel savings as an incentive for speed compliance in the informal public transport industry in South Africa. In: 2019 IEEE Intelligent Transportation Systems Conference. ITSC, pp. 492–496. http://dx.doi.org/10.1109/ITSC.2019.8917491.

Evans, J., O'Brien, J., Ch Ng, B., 2018. Towards a geography of informal transport: Mobility, infrastructure and urban sustainability from the back of a motorbike. Trans. Inst. Br. Geogr. 43 (4), 674–688. http://dx.doi.org/10.1111/tran.12239, URL: https://rgs-ibg.onlinelibrary.wiley.com/doi/abs/10.1111/tran.12239.

Faria, M.V., Duarte, G.O., Varella, R.A., Farias, T.L., Baptista, P.C., 2019. Driving for decarbonization: Assessing the energy, environmental, and economic benefits of less aggressive driving in Lisbon, Portugal. Energy Res. Soc. Sci. 47, 113–127. http://dx.doi.org/10.1016/j.erss.2018.09.006.

Fiori, C., Ahn, K., Rakha, H.A., 2016. Power-based electric vehicle energy consumption model: Model development and validation. Appl. Energy 168, 257–268. http://dx.doi.org/10.1016/j.apenergy.2016.01.097, URL: https://www.sciencedirect.com/science/article/pii/S030626191630085X.

Fridlund, J., Wilen, O., 2020. Parameter guidelines for electric vehicle route planning.

Geospatial Commission, 2021. Positioning the UK in the fast lane. URL: https://www.gov.uk/government/news/the-geospatial-commission-announces-transport-innovation-competition-winners-and-outlines-how-to-position-the-uk-in-the-fast-lane.

Google, 2022. Google earth. URL: https://earth.google.com/web/@-33.88614835,18.82391936,130.63246104a,12436.34033288d,35y,350.20430081h,0t,0r.

Higer, 2020. Higer H5C EV specifications. URL: https://en.higer.com/NewEnergy/info.aspx?itemid=6851.

Ilyès, M., Fotouhi, A., Ewin, N., 2020. Electric vehicle energy consumption modelling and estimation - a case study. Int. J. Energy Res. 45 (1), 501–520. http://dx.doi.org/10.1002/er.5700.

Jonas, T., Hunter, C., Macht, G., 2017. Quantifying the impact of traffic on electric vehicle efficiency. World Electr. Veh. J. 13 (1), http://dx.doi.org/10.3390/wevj13010015.

KCCA, 2016. Multimodal Urban Transport Master Plan for Greater Kampala Metropolitan Area. Technical Report, Kampala Capital City Authority (KCCA), Kampala.

Kivekäs, K., VepsäläInen, J., Tammi, K., 2018. Stochastic driving cycle synthesis for analyzing the energy consumption of a battery electric bus. IEEE Access 6, 55586–55598. http://dx.doi.org/10.1109/ACCESS.2018.2871574.

Kurczveil, T., Álvarez López, P., Schnieder, E., 2014. Implementation of an Energy Model and a Charging Infrastructure in SUMO. Springer Berlin Heidelberg, pp. 33–43. http://dx.doi.org/10.1007/978-3-662-45079-6_3.

Lin, B., Sai, R., 2022. Towards low carbon economy: Performance of electricity generation and emission reduction potential in Africa. Energy 251, 123952.
Maia, R., Silva, M., Araújo, R., Nunes, U., 2011. Electric vehicle simulator for energy consumption studies in electric mobility systems. In: 2011 IEEE Forum on Integrated and Sustainable Transportation Systems. pp. 227–232. http://dx.doi.org/10.1109/FISTS.2011.5973655.
Mccormick, D., Schalekamp, H., Mfinanga, D., 2016. The nature of paratransit operations. In: Behrens, R., McCormick, D., Mfinanga, D. (Eds.), Paratransit in African Cities: Operations, Regulation and Reform, first ed. Routledge, New York, pp. 59–78.
Motavalli, J., 2021. Every automaker's EV plans through 2035 and beyond. URL: https://www.forbes.com/wheels/news/automaker-ev-plans/.
Murphey, Y.L., Milton, R., Kiliaris, L., 2009. Driver's style classification using jerk analysis. In: 2009 IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems. pp. 23–28. http://dx.doi.org/10.1109/CIVVS.2009.4938719.
Ndibatya, I., Booysen, M.J., 2021. Characterizing the movement patterns of minibus taxis in kampala's paratransit system. J. Transp. Geogr. 92, 103001. http://dx.doi.org/10.1016/j.jtrangeo.2021.103001, URL: https://www.sciencedirect.com/science/article/pii/S0966692321000545.
Odhiambo, E., Kipkoech, D., Hegazy, A., Hegazy, M., Manuel, M., Schalekamp, H., Klopp, J., et al., 2021. The potential for minibus electrification in three African cities: Cairo, Nairobi, and Cape Town. Volvo Res. Educ. Found..
Pan, C., Dai, W., Chen, L., Chen, L., Wang, L., 2017. Driving range estimation for electric vehicles based on driving condition identification and forecast. AIP Adv. 7 (1), http://dx.doi.org/10.1063/1.4993945.
Pillay, N., Brent, A., Musango, J., 2019. Affordability of battery electric vehicles based on disposable income and the impact on provincial residential electricity requirements in south africa. Energy 171, 1077–1087.
Renault Group, 2021. What is the efficiency of an electric motor? URL: https://www.renaultgroup.com/en/news-on-air/news/the-energy-efficiency-of-an-electric-car-motor/#:~:text=For%20an%20electric%20vehicle%2C%20energy,not%20a%20very%20large%20amount.
Ruivii, 2021. Toano 15 seats electric minibus. URL: https://www.ruviivehicle.com/products/minibus/toano-15-seats-electric-minibus/.
Sagaama, I., Kchiche, A., Trojet, W., Kamoun, F., 2019. Evaluation of the energy consumption model performance for electric vehicles in SUMO. In: 2019 IEEE/ACM 23rd International Symposium on Distributed Simulation and Real Time Applications (DS-RT). pp. 1–8. http://dx.doi.org/10.1109/DS-RT47707.2019.8958704.
Shannon, C., 1949. Communication in the presence of noise. Proc. IRE 37 (1), 10–21. http://dx.doi.org/10.1109/jrproc.1949.232969.
Smith, R., Shahidinejad, S., Blair, D., Bibeau, E., 2011. Characterization of urban commuter driving profiles to optimize battery size in light-duty plug-in electric vehicles. Transp. Res. D 16 (3), 218–224.
Sunday Times Driving, 2022. Car makers' electric vehicle plans for 2022 and beyond. URL: https://www.driving.co.uk/news/new-cars/current-upcoming-pure-electric-car-guide-updated/.
Tesla, 2007. The magic of tesla roadster regenerative braking. URL: https://www.tesla.com/blog/magic-tesla-roadster-regenerative-braking.
Toyota, 2022a. Toyota HiAce specifications. URL: https://toyota.dreamhosters.com/pages/hiace-tmc/spec.php.
Toyota, 2022b. Hiace ses'fikile specfications. URL: https://www.toyota.co.za/specs-compare/hiace-sesfikile?model=CY2.
Transaction Capital, 2021. Transaction capital results presentation. URL: https://thevault.exchange/?get_group_doc=305/1637037219-TCFY21AnalystBooklet.pdf.
USGS, 2022. Earth resources observation and science (EROS) center. URL: https://www.usgs.gov/centers/eros.
Wang, Q., Huo, H., He, K., Yao, Z., Zhang, Q., 2008. Characterization of vehicle driving patterns and development of driving cycles in Chinese cities. Transp. Res. D 13 (5), 289–297.
Wu, X., Freese, D., Cabrera, A., Kitch, W.A., 2015. Electric vehicles' energy consumption measurement and estimation. Transp. Res. D 34, 52–67. http://dx.doi.org/10.1016/j.trd.2014.10.007, URL: https://www.sciencedirect.com/science/article/pii/S1361920914001485.
Zeeman, A., Booysen, M.J., 2014. Public transport sector driver behaviour : measuring recklessness using speed and acceleration. South. Afr. Transp. Conf. URL: http://hdl.handle.net/2263/45581.
Zinkernagel, R., Evans, J., Neij, L., 2018. Applying the SDGs to cities: Business as usual or a new dawn? Sustainability 10 (9), http://dx.doi.org/10.3390/su10093201.