

OPEN  
ARTICLE

# An open tool for creating battery-electric vehicle time series from empirical data, *emobpy*

Carlos Gaete-Morales<sup>1</sup>, Hendrik Kramer<sup>2</sup>, Wolf-Peter Schill<sup>1</sup>✉ & Alexander Zerrahn<sup>1</sup>

There is substantial research interest in how future fleets of battery-electric vehicles will interact with the power sector. Various types of energy models are used for respective analyses. They depend on meaningful input parameters, in particular time series of vehicle mobility, driving electricity consumption, grid availability, or grid electricity demand. As the availability of such data is highly limited, we introduce the open-source tool *emobpy*. Based on mobility statistics, physical properties of battery-electric vehicles, and other customizable assumptions, it derives time series data that can readily be used in a wide range of model applications. For an illustration, we create and characterize 200 vehicle profiles for Germany. Depending on the hour of the day, a fleet of one million vehicles has a median grid availability between 5 and 7 gigawatts, as vehicles are parking most of the time. Four exemplary grid electricity demand time series illustrate the smoothing effect of balanced charging strategies.

## Introduction

We introduce *emobpy*. It is an open-source, python-based tool that creates profiles of battery-electric vehicles (BEV), based on empirical mobility statistics and customizable assumptions. We additionally provide a first application of the tool and create vehicle profiles based on representative German mobility data. An *emobpy* profile consists of four time series: (i) *vehicle mobility* containing the vehicle's location and distance travelled, (ii) *driving electricity consumption*, specifying how much electricity is taken from the battery for driving; (iii) *BEV grid availability*, providing information whether and with which power rating a BEV is connected to the electricity grid at a certain point in time; and (iv) *BEV grid electricity demand*, specifying the actual charging electricity drawn from the grid, based on different charging strategies.

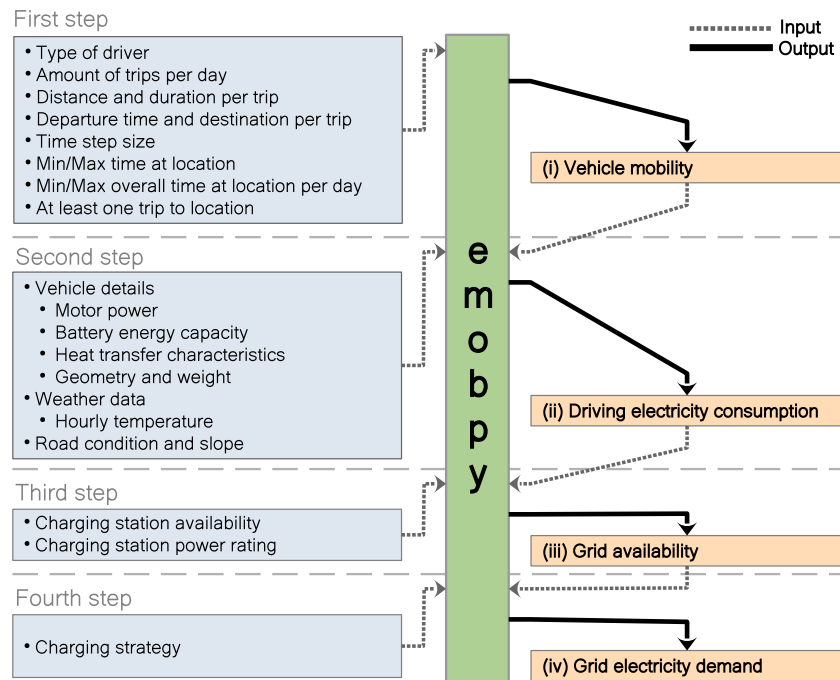
Such profiles are core input data for a wide range of model applications in energy, environmental, and economic studies on BEV. Technology developments as well as energy and climate policy measures drive the deployment of BEV in many countries<sup>1</sup>. Growing BEV fleets can have substantial impacts on the power sector. They increase the electric load, but may also provide temporal flexibility for integrating variable renewable energy sources and contribute to decarbonizing transportation<sup>2</sup>. Many model-based analyses investigate potential power sector interactions of future BEV fleets<sup>3–6</sup> and thus depend on a meaningful representation of electric vehicles' mobility patterns.

Yet such data are often not publicly available. In general, empirical data are scarce because BEV fleets are still small in most countries. And if respective time series are available, they are often specific to the conditions in which the data was collected and subject to data protection provisions. Past approaches make either stylized coarse assumptions<sup>7</sup>, derive data from mobility statistics, but lack documentation, transparency or reproducibility<sup>8–14</sup>, or are idiosyncratic with respect to geographic characteristics or assumed driver behavior<sup>11,15,16</sup>.

Following<sup>3</sup>, we argue that new models are needed to derive relevant time series in a transparent and flexible way. As a first step in this direction, the tools *Vencopy*<sup>17</sup> and *RAMP-mobility*<sup>18</sup> recently emerged. To further fill this gap, we developed *emobpy*. Our tool takes empirical mobility statistics, physical properties of vehicles, and customizable assumptions as inputs and delivers BEV profiles as output. Figure 1 gives a stylized account. We first discuss the outputs, then the required inputs.

Four output time series constitute one BEV profile. These profiles have a customizable length and resolution. A handy format for many applications is all hours of one year. But other formats are also possible by discretion of the researcher. Likewise, the researcher can choose how many profiles she wants to create.

<sup>1</sup>German Institute for Economic Research (DIW Berlin), Mohrenstr. 58, D-10117, Berlin, Germany. <sup>2</sup>Workgroup for Infrastructure Policy (WIP), Technische Universität Berlin, D-10623, Berlin, Germany. ✉e-mail: [wschill@diw.de](mailto:wschill@diw.de)



**Fig. 1** Inputs and outputs of *emobpy* and sequence of generating four types of time series. The boxes on the left-hand side show customizable input assumptions, the boxes on the right-hand side indicate the four types of time series.

The time series of vehicle mobility (i) contains the location of the vehicle at each time step and the time steps during which the vehicle is driving with information of the distance traveled. The driving electricity consumption time series (ii) provides information on how much electricity the vehicle consumes for driving in each time step. It also includes energy needs for accessories and heating/cooling, as well as the charging and discharging losses of the battery. The time series of (iii) grid availability provides information whether a vehicle is connected to the electricity grid in a time step and if so, with what power rating for charging or discharging. The time series of grid electricity demand (iv) provides information on how much electricity a vehicle demands from the electricity grid in a time step. Time series (i), (ii) and (iii) are core inputs for models that endogenously determine the timing of charging (and, potentially, discharging to the grid); the time series (iv) are core inputs for models that do not endogenously determine the grid interactions of BEV, but use exogenous input data for vehicle charging.

The required input data for the time series of vehicle mobility (i) are the relative frequencies of different driver types, e.g., commuters, of the number of trips per day, of the destination, distance and duration of trips, and of the departure hours. Such information can often be derived from national mobility statistics. If required or desired, a researcher can also make own assumptions or resort to the pre-set values from German mobility statistics. *emobpy* makes sure that the resulting time series are feasible and consistent. To this end, a minimum and maximum number of hours at specific locations can be specified, and it is assumed that the last trip of a day heads home. With a Monte Carlo approach, *emobpy* ensures variability across profiles.

Based on the vehicle mobility time series, the driving electricity consumption (ii) time series is derived. This requires further input data, such as information on nominal motor power, curb weight, drag coefficient, and dimensions, which the tool includes for several current BEV models. Ambient temperature is also a significant parameter that affects the consumption of BEV<sup>8,19</sup>. For that reason, *emobpy* is endowed with a database of hourly temperature for European countries with a registry of the last 17 years. Additionally, the vehicle cabin insulation characteristics are required; this data is not widely available and thus assumed independently of the BEV models database. Driving cycles are also important input parameters that are used to simulate every individual trip. The model includes two driving cycles, *Worldwide Harmonized Light Vehicles Test Cycle* (WLTC) and *Environmental Protection Agency* (EPA). This input data is already provided within the tool, and the user can select a particular BEV model, country weather, and driving cycle. Alternatively, *emobpy* also allows providing user-defined custom data.

The required input data for the grid availability time series (iii) is the driving electricity consumption time series (ii). Further, data or assumptions on the power rating of charging stations at different generic locations as well as their availability probabilities are needed. Variability across profiles is, again, introduced through a Monte Carlo approach, while *emobpy* makes sure that the time series (iii) is consistent within each profile.

The required input data for the grid electricity demand time series (iv) includes the created time series on driving electricity consumption (ii) and grid availability (iii). Additionally, users can choose a charging strategy, such as immediate full charging or night-time charging, or make customary assumptions.

Number of trips	Working days	Weekend days
0	35.4	50.7
1	0.0	0.0
2	29.9	27.5
3	8.3	4.4
4	12.5	10.2
5	13.9	7.2

**Table 1.** Probability distributions (given in %) for the amount of trips per day by days of the week. *Note:* Data adapted from<sup>20</sup>. Commuters have the same distribution of daily trips as non-commuters. Data corresponds to the group of respondents that have a yearly mileage in the range of 10,000–15,000 km.

Distance	Trip duration (minutes)						
	10	10–15	15–20	20–30	30–45	45–60	60–185
1 km	2.9	0.3	0	0	0	0	0
1–2 km	3.5	4.8	0.8	0	0	0	0
2–5 km	8.4	10.2	5.7	0	1.2	0.4	0
5–10 km	1.3	12.2	14.4	0	2.4	0.6	0.7
10–20 km	0	0.9	6.3	0	4.7	1.3	0.5
20–50 km	0	0	0	0	8.6	2.1	1.6
50–100 km	0	0	0	0	0	0.6	2.1
100–400 km	0	0	0	0	0	0	1.5

**Table 2.** Joint probability distributions (given in %) for the distance travelled by trip and trip duration. *Note:* Data adapted from<sup>20</sup>. Numbers rounded to one decimal. Data corresponds to the group of respondents that have a yearly mileage in the range of 10,000–15,000 km. All values add up to 100%.

## Results

**Application to Germany: parameterization and setup.** For a first application of *emobpy*, we draw on the comprehensive German mobility survey *Mobilität in Deutschland* (Mobility in Germany)<sup>20</sup>. The survey features mobility data relating to different types of households, vehicles, individuals, and trips. In this application, we make three general assumptions: first, we assume that individuals with access to a vehicle carry out all their trips with the same vehicle; second, we assume that future BEV drivers have similar mobility patterns as current drivers of conventional vehicles covered by the underlying mobility statistics; and third, for simplicity and tractability, we assume that there are only four BEV models: Hyundai Kona, Renault Zoe, Tesla Model 3 and Volkswagen ID.3. These models had the largest market shares in Germany by the time of writing. Again, all of these pre-set assumptions can easily be modified in *emobpy*.

We generate 50 profiles for each BEV model, i.e., 200 BEV profiles overall, each consisting of four time series. We focus on two types of drivers: commuters (62% of all drivers) and non-commuters (38% of all drivers). For commuters, we further differentiate between full-time and part-time employees, with a split of 78 to 22%<sup>21</sup>. We exclude commuting students, apprentices, and trainees, who represent only a small share of all commuters in the initial dataset. The amount of trips per day varies between 0–5 with different probabilities for weekdays and weekend days (Table 1).

The trip distance and duration follows a probability distribution derived from the input data (Table 2). As the underlying mobility statistics features a category that includes any trips with more than 100 km distance and more than 60 minutes duration, we cap the maximum distance travelled per trip at 400 km and the trip duration at 185 minutes. We also ensure that the average velocity resulting from every possible combination of distance and duration cannot exceed 130 km/h.

The probability of departure times is specific to the trip destination, type of driver, and day of the week (Table 3). It is distributed according to the underlying mobility statistics. Following the input data, we consider six trip destinations: *workplace*, *shopping*, *errands*, *escort*, *leisure*, and *home*. An example for errands is a visit to the doctor or to the authorities. In the case of escort destinations, the driver transports other persons, for example children. A set of rules is implemented in this case study to select only consistent day trips. The rules are applied depending on the day of the week and the type of driver (compare Methods section).

Table 4 contains information on the four BEV models used for this case study. Most of the parameters serve to calculate driving electricity consumption, with the exception of nominal battery capacity that is used to generate the grid availability time series (iii) and grid demand time series (iv). Many other parameters are also provided by *emobpy* to calculate the driving electricity consumption, such as efficiencies, auxiliary power and heat transfer data (compare additional Tables in the Methods section). These are default values in the tool; however, they can be modified if desired by the user.

We model time steps of 15 minutes. In each time step, a vehicle is either *driving* in case a trip takes place, or is in one of the locations *workplace*, *shopping*, and so on. Depending on the vehicle location, a charging station to connect the vehicle to the grid may be available with a location-specific power rating. For this application, we assume four generic types of charging stations with different probability distributions for each vehicle location.

Commuter	Workplace		Shopping		Errands		Escort		Leisure		Home	
	yes		yes	no	yes	no	yes	no	yes	no	yes	no
<b>Departure</b>	<b>Working days</b>											
05:00–08:00	11.1		0.5	0.7	0.5	0.7	1.1	0.7	0.5	0.7	0.8	0.7
08:00–10:00	3.1		1.8	4.5	1.4	4.1	0.8	0.9	1.4	3.2	1.8	3.6
10:00–13:00	1.3		2.7	6.7	2.3	5.4	0.7	1.3	3.2	4.7	5.5	11.7
13:00–16:00	1.1		2.5	3.7	2.2	4.0	1.8	1.5	3.8	5.9	8.9	8.2
16:00–19:00	0.3		3.0	1.9	2.2	2.2	1.4	1.0	4.9	4.5	14.0	9.3
19:00–22:00	0.3		0.4	0.1	0.6	0.4	0.4	0.3	2.4	1.5	6.1	4.0
22:00–05:00	0.6		0.0	0.0	0.1	0.1	0.1	0.1	0.4	0.2	2.4	1.3
	<b>Saturday</b>											
05:00–08:00	0.9		1.2	1.2	0.3	0.3	0.2	0.2	0.8	0.8	0.8	0.8
08:00–10:00	0.5		4.8	4.9	1.9	2.0	0.7	0.7	2.7	2.8	3.0	3.1
10:00–13:00	0.4		7.1	7.3	3.5	3.6	1.4	1.5	5.2	5.4	9.1	9.3
13:00–16:00	0.2		3.4	3.5	2.5	2.6	1.2	1.2	7.0	7.1	7.6	7.8
16:00–19:00	0.1		2.3	2.4	1.7	1.7	1.1	1.1	6.0	6.1	9.5	9.7
19:00–22:00	0.1		0.4	0.4	0.5	0.5	0.4	0.4	2.5	2.6	4.9	5.0
22:00–05:00	0.2		0.0	0.0	0.1	0.1	0.2	0.2	0.8	0.8	3.0	3.1
	<b>Sunday</b>											
05:00–08:00	0.8		0.3	0.3	0.2	0.2	0.1	0.1	0.8	0.8	0.4	0.4
08:00–10:00	0.4		1.5	1.5	1.4	1.5	0.6	0.6	4.8	4.9	2.0	2.1
10:00–13:00	0.3		0.7	0.7	2.8	2.8	1.3	1.3	11.7	11.9	7.2	7.4
13:00–16:00	0.3		0.5	0.5	2.6	2.6	1.4	1.4	13.7	14.0	8.8	9.0
16:00–19:00	0.2		0.2	0.2	1.8	1.9	1.0	1.0	6.8	7.0	13.3	13.6
19:00–22:00	0.2		0.1	0.1	0.5	0.5	0.4	0.5	2.0	2.1	6.4	6.6
22:00–05:00	0.1		0.0	0.0	0.1	0.1	0.1	0.1	0.4	0.4	2.0	2.1

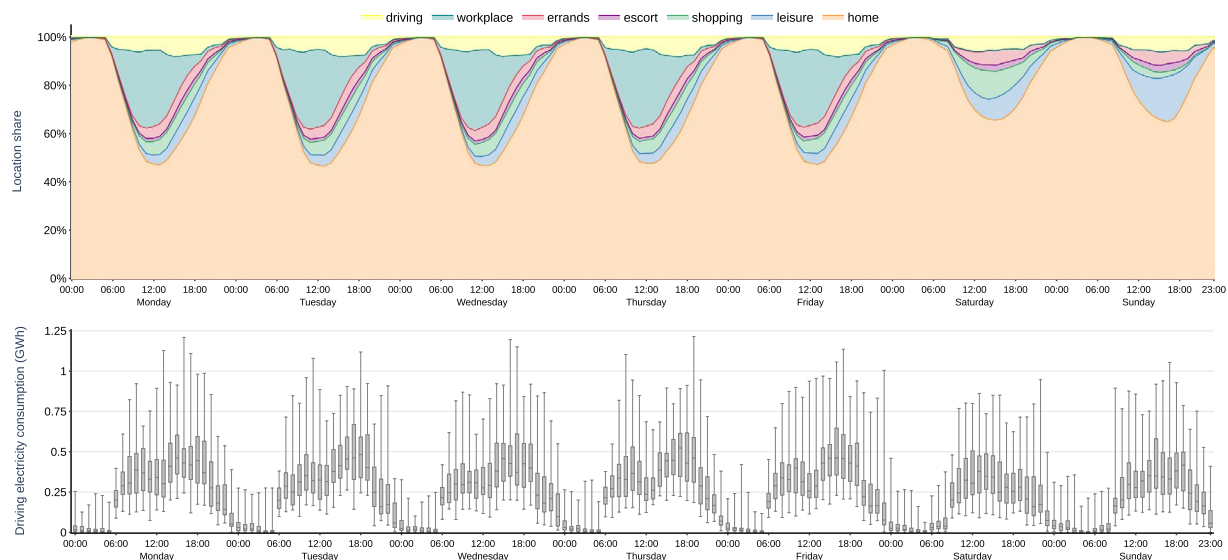
**Table 3.** Joint probability distributions (given in %) for trip destinations and departure times, differentiated for commuters and non-commuters and days of the week. *Note:* Data adapted from<sup>20</sup>. Numbers rounded to one decimal.

Parameter	Unit	BEV models				Description
		Model 3 (Tesla)	ID.3 (VW)	Kona (Hyundai)	Zoe (Renault)	
$N_{motor}$	kW	358	93	150	65	Nominal motor power
$N_{battery}$	kWh	79.5	45.0	64.0	45.6	Nominal battery capacity
$m_c$	kg	1860	1600	1685	1480	Curb weight
$C_d$	—	0.23	0.27	0.29	0.29	Drag coefficient
$h$	m	1.44	1.55	1.57	1.56	Height
$w$	m	1.85	1.81	1.80	1.73	Width
$r_{gear}$	—	9.0	10.0	8.0	9.3	Gear ratio
$PMR$	W/kg	192	58	89	44	Power to mass ratio

**Table 4.** BEV models' parameters derived from manufacturer data<sup>26</sup>.

The charging stations are at *home*, in the *public* area, or at the *workplace*, or *none* is available. Respective power ratings are 3.6, 22, 11, and 0 kW, based on<sup>22</sup>. The tool also considers fast charging; this feature is available for long-distance trips that are larger than the vehicle maximum range. The charging capacities selected for this application are 75 and 150 kW. This can be interpreted as the vehicle making a short stop during a longer trip. Charging efficiency is set to 90%<sup>6,9,15,23</sup>.

When at *home*, 81% of all drivers park their vehicles in a carport or garage and 19% on public streets according to<sup>20</sup>. For the group of vehicle profiles that have a carport at home, we assume a 100% charging availability. For those without a private charging station, we set a probability of 50% to find a *public* charging station and 50% of finding *none*. For commuters, we consider three charging groups with different grid connection opportunities during work hours: charging at the *workplace*, charging in the *public* area, or *none*. When commuters park their BEV at the *workplace*, we assume that 50% of them can charge their vehicles there, with a 100% probability of finding a charging station; 25% of commuters charge in a *public* area, with a 50% probability of finding a charging station; and the remaining 25% of commuters are assumed to have a 100% probability of not having a charging station available during work hours (*none*). For the vehicle locations *shopping*, *errands*, *escort*, and *leisure*, we assume a probability of 50% to find a *public* charging station and 50% to find *none*. When driving, grid connection is not available, with the exception of fast-charging for very long trips.



**Fig. 2** Simulated time series of vehicle locations (top panel) and driving electricity consumption (bottom panel) of one million BEV, given as averages and box plots for each hour of the week.

To derive time series of BEV grid electricity demand, we apply four exemplary charging strategies. Note that these charging strategies do not take into account any power sector or electricity market price information:

- *immediate - full capacity*: BEV charge their batteries at full power rating as soon as they arrive at charging stations. Charging stops when the battery is full, or when the next trip starts. This mimics a setting where drivers have no incentives and/or no technical possibility to charge their vehicle batteries in a more balanced way, which is likely to be sub-optimal with respect to the electricity market or network situation.
- *immediate - balanced*: BEV start charging their batteries as soon as they arrive at charging stations, however with constant power rating (usually below the power rating of the charging station), such that a 100% state of charge is reached just before starting the next trip, assuming perfect foresight of the next departure time. This approximates a smoother and potentially more system-oriented charging behavior.
- *at home - balanced*: similar to the previous charging strategy, but BEV only charge *at home*, even when additional charging options are available at other locations. This reflects a preference or economic incentives for home charging.
- *at home night-time - balanced*: similar to the previous charging strategy, but with charging time restricted to the time window between 23:00 and 8:00. This mimics the effect of potential tariff incentives for night-time (off-peak) charging.

**Vehicle mobility.** Figure 2 summarizes all 200 simulated vehicle mobility time series. For visualization, we aggregate the quarter-hourly time series to full hours. For each hour, vehicle locations are averaged over all profiles and weeks of the year. Hourly driving electricity consumption is summarized in box plots, rendering the dispersion over the simulated profiles through the weeks of the year. All numbers are linearly scaled up to represent one million BEV, so the setting may be interpreted as a German near-term future scenario.

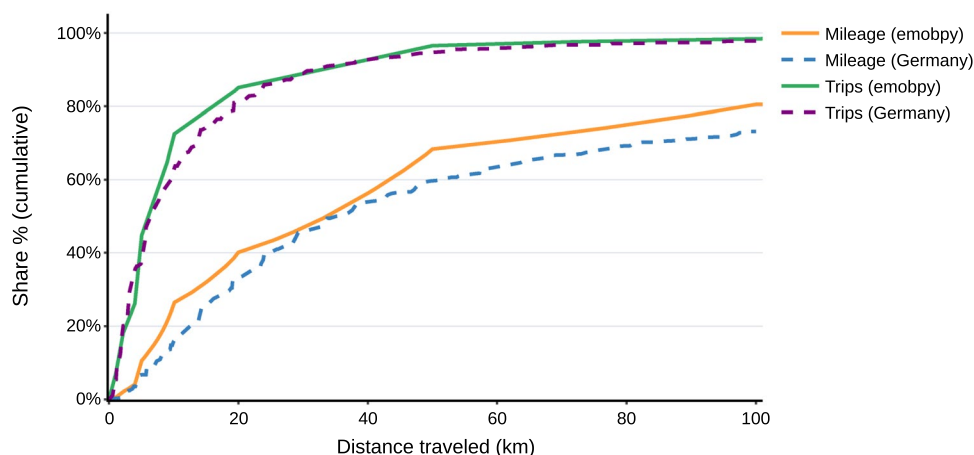
Most of the time, vehicles are parking (top panel). At night, between 23:00 and 5:00, more than 96% of the fleet are, on average, at *home*. During daytime, still the majority of vehicles are at *home*, but also a large proportion of vehicles is at the *workplace*, peaking at 32% at 11:00 on working days. During weekends, more vehicles stay at *home*, and the shares of *shopping*, *errands*, *escort*, and *leisure* increase. Commuters have a positive but very small probability of going to the workplace on weekends (Tables 3 and 5), so it is hardly visible in Fig. 2. Every day between 6:00 and 22:00, at least 3% of the fleet are *driving*, with a peak between around 15:00 and 17:00 with about 9% of the fleet *driving*.

To validate *emobpy* results, we compare the cumulative distributions of trips and mileage to the underlying German mobility statistics<sup>20</sup>. For the two metrics, the cumulative distributions follow a similar pattern (Fig. 3). Both our *emobpy* application and the official German statistics indicate that about 90% of all trips have a distance travelled of 50 km or below. The cumulative mileage – the overall distance travelled by all vehicles in a year – also has a similar shape in *emobpy* and in the official statistics up to 40 km. The Figure also allows inferring that long-distance trips above 100 km represent 25% of the yearly mileage, while those trips only account for 3% of all trips<sup>20</sup>.

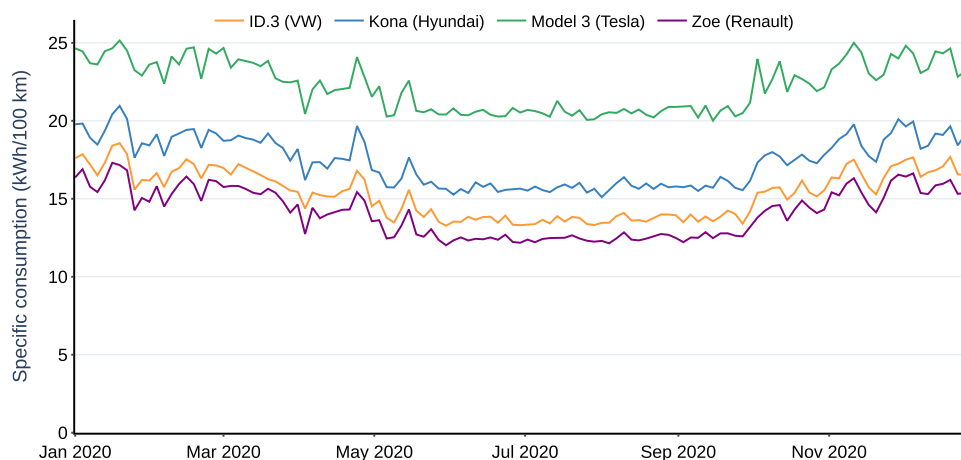
**Driving electricity consumption.** The overall hourly driving electricity consumption of one million BEV (Fig. 2, bottom panel) peaks between 15:00 and 17:00 on working days with an annual median around 450 MWh (i.e., 0.45 kWh per vehicle), and an absolute maximum at 19:00 of 1250 MWh (1.25 kWh per vehicle). During the weekend, overall consumption is lower and has less distinctive evening peaks. Note that the values also include energy needs for accessories and heating/cooling, as well as battery losses.

Rule		Non-commuter		Full-time commuter		Part-time commuter	
		Working day	Weekend	Working day	Weekend	Working day	Weekend
Minimum time at	home	0.5 hrs	0.5 hrs	0.5 hrs	0.5 hrs	0.5 hrs	0.5 hrs
	workplace	—	—	3.5 hrs	3.0 hrs	3.5 hrs	3.0 hrs
	other destinations	0.5 hrs	0.5 hrs	0.5 hrs	0.5 hrs	0.5 hrs	0.5 hrs
Minimum time per day at	home	9 hrs	6 hrs	9 hrs	6 hrs	9 hrs	6 hrs
	workplace	—	—	7 hrs	3 hrs	3.5 hrs	3 hrs
	other destinations	—	—	—	—	—	—
Maximum time per day at	home	—	—	—	—	—	—
	workplace	—	—	8 hrs	4 hrs	4 hrs	4 hrs
	other destinations	—	—	—	—	—	—
At least one trip to	home	yes	yes	yes	yes	yes	yes
	workplace	—	—	yes	no	yes	no

**Table 5.** Rules implemented to select consistent day trips.

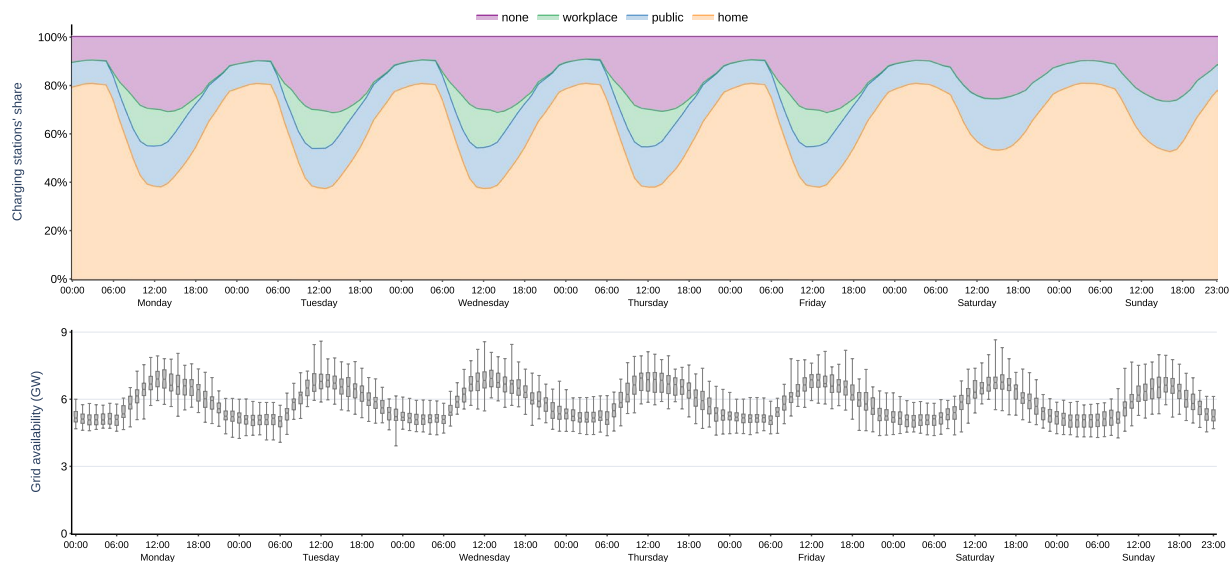


**Fig. 3** Comparison of cumulative shares of trips and mileage per distance travelled. “Germany” represents German mobility statistics<sup>20</sup>, which reports these aggregate shares up to a distance of 100 km.



**Fig. 4** Specific consumption of four selected BEV models throughout a year. The values are calculated as the medians of all trips taken for every three days.

The average specific consumption for all trips over the 200 profiles and four models is 22.2 kWh/100 km. The median specific consumption values for the different BEV types Model 3, Kona, ID.3 and Zoe are 22.7, 17.9, 15.7, and 14.5 kWh/100 km, respectively (Fig. 4). The ambient temperature variation has a clear impact on the specific consumption. On average, specific consumption is lowest in summer with 20.7 kWh/100 km, and highest in winter with 23.6 kWh/100 km.



**Fig. 5** Simulated time series summarized for different types of charging stations (top panel) and grid-connected power rating (bottom panel) of one million BEV, given as averages and box plots for each hour of the week.

**Grid availability.** The cumulative simulated grid availability time series is shown in Fig. 5. On working days, the time series on the types of charging stations (top panel) has a recurring pattern that corresponds to the pattern of vehicle locations. The share of vehicles with a charging station available reaches a 90% peak between 3:00 and 5:00 at night. Here, around 80% of vehicles are connected at *home* charging stations, and 10% at charging stations on *public* streets. Nearly all vehicles park at home during these hours, and we assume 80% of these to be in garages which have a connection to the grid (“home” category in the Figure). For the remaining 20% of vehicles that are parked on the street, we assume only 50% to be able to find a public charging station, and the other 50% have no grid connection. Between 11:00 and 12:00, average grid availability is at a minimum level of 70%. During daytime, a relevant proportion of available charging stations is at the *workplace*. On weekends, the charging station time series is less peaky, with higher proportions at *home* and on *public* streets during daytime.

The grid-connected power rating is lowest between 19:00 and 8:00, with a median between 5.0 and 5.6 GW for a fleet of one million BEV (bottom panel). This is due to the high share of home charging stations with a low power rating of 3.6 kW. During daytime, the median grid-connected power rating is greater than 7 GW because charging stations available either at the *workplace* or in *public* areas have a power rating of 11 and 22 kW, respectively.

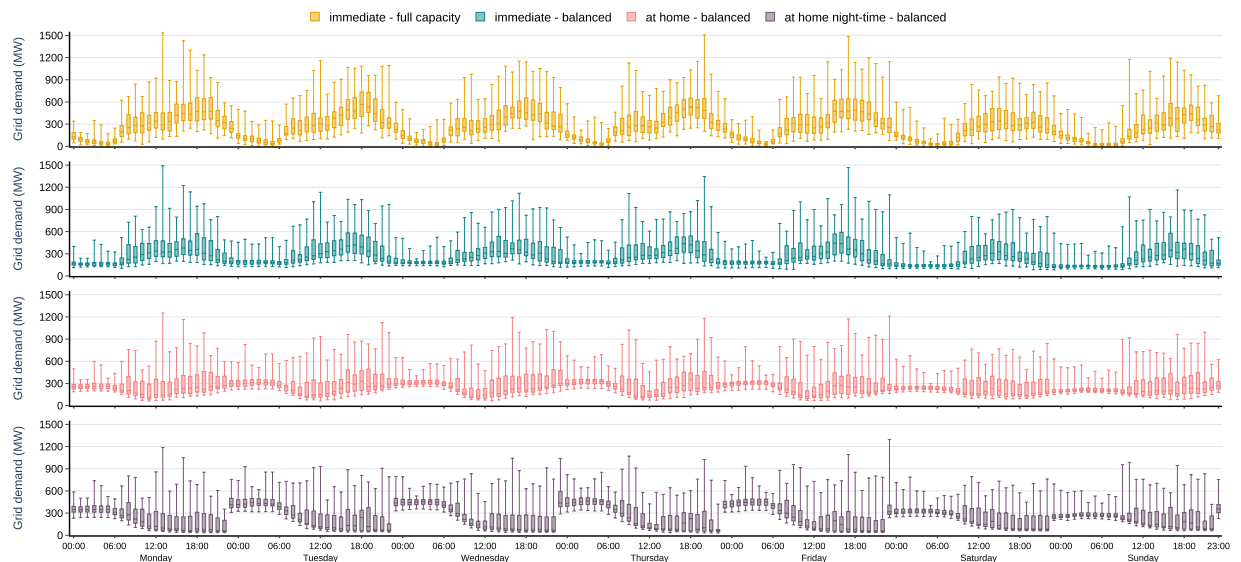
**Grid electricity demand.** The grid electricity demand time series of the four exemplary charging strategies are summarized in Fig. 6. The *immediate - full capacity* charging strategy leads to a volatile cumulative BEV grid electricity demand both over the week and over the year, with a pronounced diurnal pattern. A distinctive peak of hourly electricity demand from the grid, with median values around 460 MW for a fleet of one million BEV (0.46 kW per vehicle), occurs on working day afternoons between 17:00 and 20:00, when many vehicles arrive at *home* and charge immediately at full power rating. As the entire BEV fleet is assumed to charge similarly in this scenario, such a charging strategy would substantially add to the evening peak of electric load. It could thus have substantial repercussions on the power sector and other electricity consumers. Load peaks could increase even further if higher power rating for charging at home was considered.

The *immediate - balanced* and *at home - balanced* charging strategies have smoother temporal grid electricity demand patterns with lower peaks, because vehicles do not get charged at full power rating once they reach a charging station. Both the variance of medians and (interquartile) ranges are lower. Likewise, the median hourly consumption of the one million BEV fleet rarely exceeds 400 MW for *immediate - balanced*, and 300 MW for *at home - balanced* (0.4 and 0.3 kW per vehicle, respectively). During weekdays, fluctuations are more pronounced for *at home - balanced*, as most vehicles are at *home* every night. Compared to *immediate - full capacity*, such smoother charging may be more compatible with the power sector.

The *at home night-time - balanced* charging strategy shows a distinct load peak at working day nights, with median hourly grid electricity demand of one million BEV around 420 MW (0.42 kW per vehicle). Between Friday evening and Monday morning, median demand at night-time is lower than 300 MW because the vehicles are less used on weekends than on working days. Accordingly, any regulatory measures that shift BEV charging to night-time periods would lead to substantially less smooth patterns compared to all-day charging. Yet the power sector implications of these charging strategies are less clear and should be investigated in detail with dedicated energy models.

## Discussion

The open-source tool *emobpy* allows to derive electric mobility time series from empirical mobility data in a transparent and customizable way. The central outputs are profiles for individual BEV, consisting of four basic types of time series covering a full year: vehicle mobility, driving electricity consumption, grid availability, and grid electricity demand. The number of vehicle profiles can be freely chosen. A greater number of profiles represents a



**Fig. 6** Simulated grid electricity demand time series for a fleet of one million BEV for four charging strategies, summarized in box plots for each hour of the week.

large and diverse BEV fleet more realistically, yet may lead to greater computational burden when using the time series in energy model applications. Users may customize the tool and alter both the German mobility data used here and the various assumptions we made, such as the shares of driver types or the availability and power rating of charging stations.

The generated vehicle profiles can be used as inputs for a wide range of model analyses of electrified and decarbonized mobility futures. Research questions in energy, environmental, and economic studies requiring temporally detailed data of BEV are abundant. These comprise the role of BEV as a flexibility resource to make efficient use of renewable electricity, emission effects of electric mobility, the impact of new loads from BEV on electricity prices, or electricity market repercussions of optimized versus user-driven charging schedules.

Several limitations offer scope for future research. First, *emobpy* focuses on vehicles only. Addressing the individual choice of the modal split between different modes of transport would provide an interesting complementary approach. This would also allow to relax the assumption that all trips are made with the same vehicle. Second, *emobpy* draws on past mobility behavior data that does not necessarily reflect future behavior. While this is a generic issue in *ex-ante* analyses, the model is generally flexible to accommodate alternative input data assumptions for future or counterfactual scenarios. However, to represent structural changes of future mobility demand, such as a potential shift toward car sharing or ride hailing concepts, possibly combined with autonomous driving, some further development of the model code may be required. Third, using input data on the distance and destination of trips, *emobpy* determines vehicle locations as background information for creating a BEV profile. While this is a convenient approach to simulate temporal variation, it has no explicit spatial resolution. We argue that this is a minor drawback because many energy, environmental, and economic models rather address a macro perspective without zooming into fine spatial detail. Further, we exclude a group of drivers that have a *service* trip destination according to<sup>20</sup>. This refers to profiles with numerous work-related trips per day, e.g., taxi drivers, which are conceptually challenging to model in our current framework. As we publish the code open-source under a permissive license, we expect that future and potentially collaborative development could address these and other options for improvement.

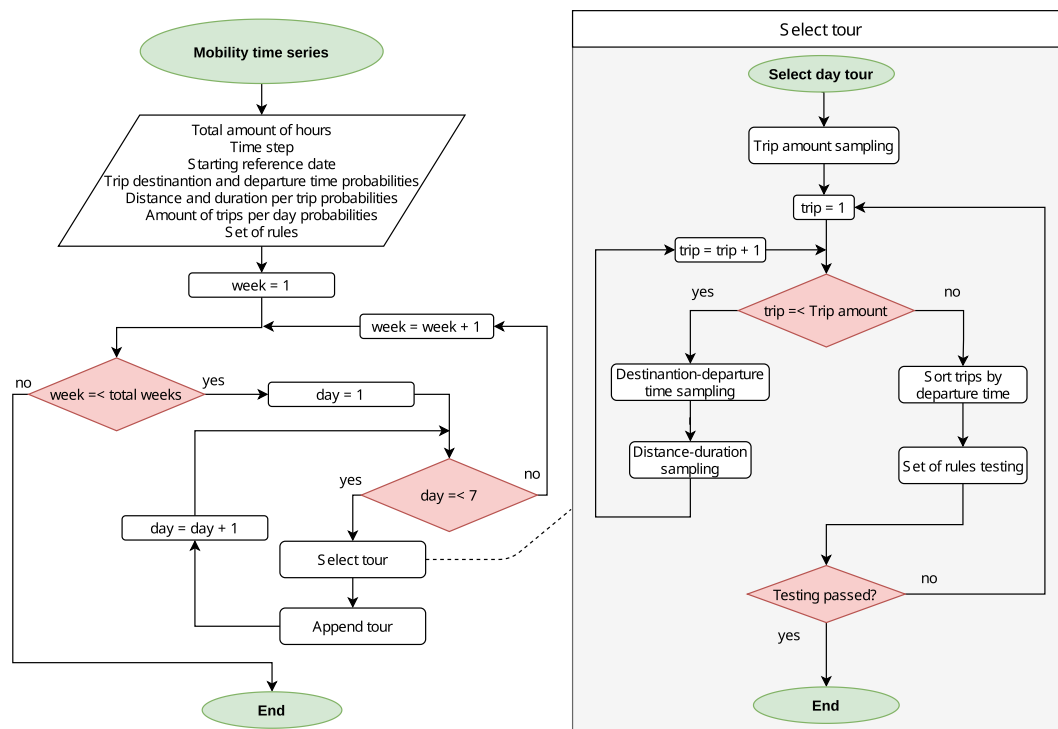
## Methods

One BEV profile consists of four time series: (i) vehicle mobility, (ii) driving electricity consumption, (iii) grid availability, and (iv) grid electricity demand. Time series (i) is created first. All the following time series will build up from this time series as it has locations at every time step and distance travelled while driving. Then, time series (ii) is calculated, taking time series (i) as an input. Time series (iii) is created, based on time series (ii); and time series (iv) is generated taking into account (ii) and (iii). For this Methods section, we introduce the following definitions:

- Edge: link or vertex that connects two nodes, where each node comprises an origin or a destination of a trip.
- Trip: edge with departure time, distance travelled and duration of the travel as attributes.
- Tour: also referred to as a day tour, it consists of a list of chronologically sorted trips by departure time. A tour contains all trips carried out by a BEV in a day.

The sampling approach consists of a sampling procedure of discrete choices. Input parameters are discrete choices with given corresponding probabilities<sup>24</sup>. Additionally, and only for the sampling of distance-duration-relations of trips, a second sampling is carried out if the probability distribution contains





**Fig. 7** Vehicle mobility time series flow diagram.

discrete distance ranges and duration ranges. In this case, a uniform distribution of integers is assumed to obtain a distance value which is within the distance range. The duration of the trip is subsequently obtained by interpolating the sampled distance with the respective distance range and duration range (see Table 2).

**Vehicle mobility time series.** The flow diagram shown in Fig. 7 illustrates how *emobpy* creates the time series of vehicle mobility. The input data are shown in the parallelogram in the left panel. The proportion of commuters and non-commuters is based on empirical data or assumptions. Additionally, the total time frame as a number of total weeks must be specified. A reference date can be used to map the day of the week. This is not only useful when the input statistics differentiates between weekdays, but also for allocating the temperature, which is a step required in the creation of time series (ii). Further inputs are three probability distributions that contain the number of day trips, the destinations and departure times, and the distance-duration-relation of the trips (compare Tables 1, 2, 3). Finally, a set of rules ensures that the tours are plausible (compare Table 5).

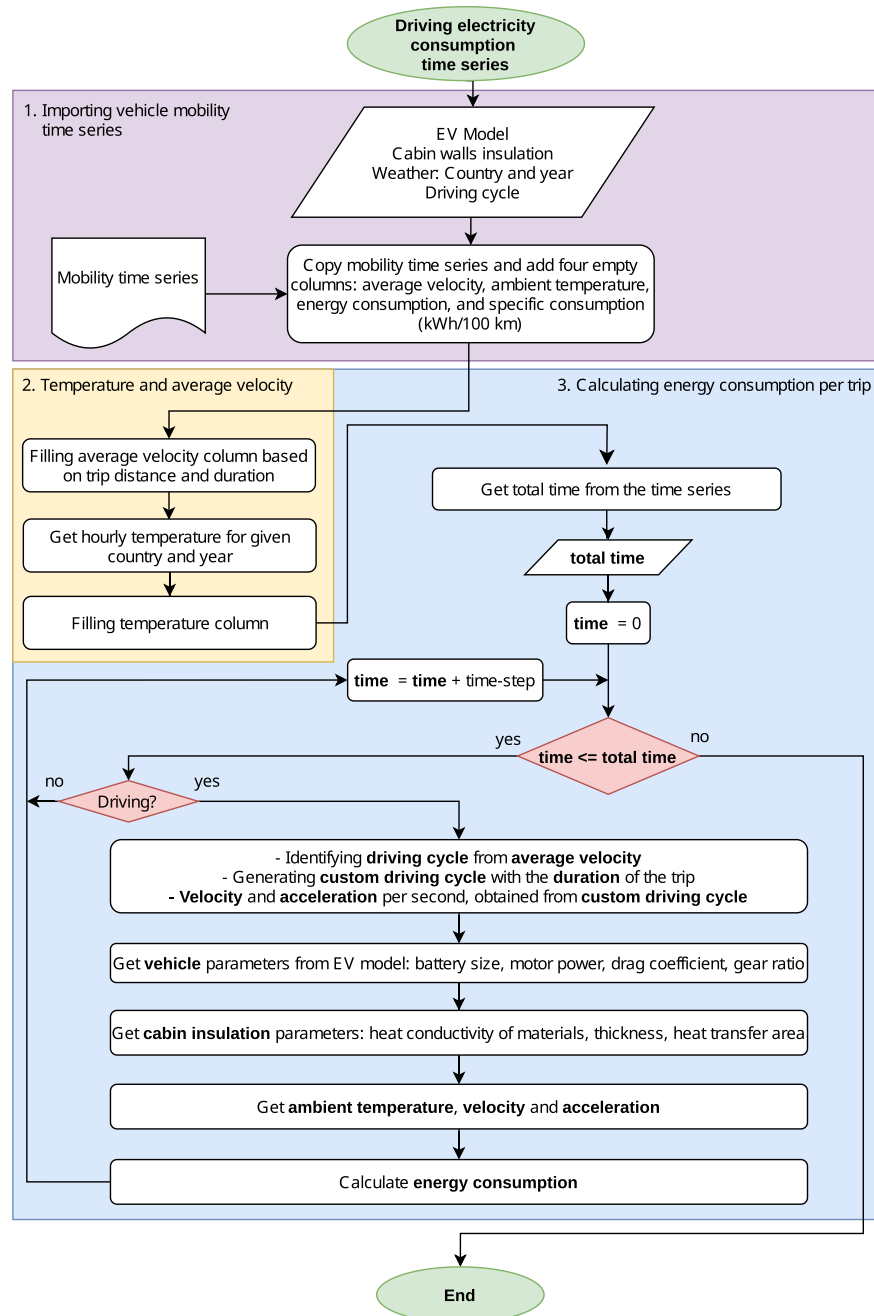
The function *Select Tour* creates a plausible day tour. Its output is a chronologically sorted list of trips, where each trip is represented by an edge of two locations (origin and destination) with departure time, distance travelled, and trip duration. This function is used twice as displayed in the left-hand side of Fig. 7.

For every day of the calculation period, the function *Select Tour* is called. Initially, a number of trips for the current day is obtained by sampling from the probability distribution that matches the type of driver. Trips are sampled according to the joint probability distribution of destinations and departure times. The sampled trips are stored in a sequential order. For each new sampled trip, *emobpy* disregards all tuples that contain the departure time of the already selected tuples, and the probability of the remaining tuples is normalized to add up to 100%. This avoids selecting a destination-departure time tuple with the same departure time as the one already selected. Once the total amount of tuples matches the number of daily trips, the sampling is finished and the tuples are ordered chronologically.

From the chronologically ordered tuples of destination and departure time, the eventual trips are created by establishing an origin-destination edge with its departure time as an attribute. The distance travelled and trip duration for each trip is sampled from the probability distribution provided by distance-duration statistics, such as shown in Table 2. Distance and duration of each trip are also attached to the origin-destination edge as attributes.

The duration time at each location is calculated from the arrival time and departure time. The arrival time is estimated from the previous trip departure time and trip duration. The next step evaluates the feasibility of the tour by checking the set of rules (Table 5), such as the minimum time at the workplace or whether the last trip heads home. All rules must be satisfied, or the current tour is discarded and the process is repeated until feasible results are obtained.

**Driving electricity consumption time series.** The flow diagram displayed in Fig. 8 illustrates how *emobpy* creates the driving electricity consumption time series. The first block describes the input data, including the vehicle mobility time series. Different types of input parameters are required. Parameters associated with



**Fig. 8** Driving electricity consumption flow diagram.

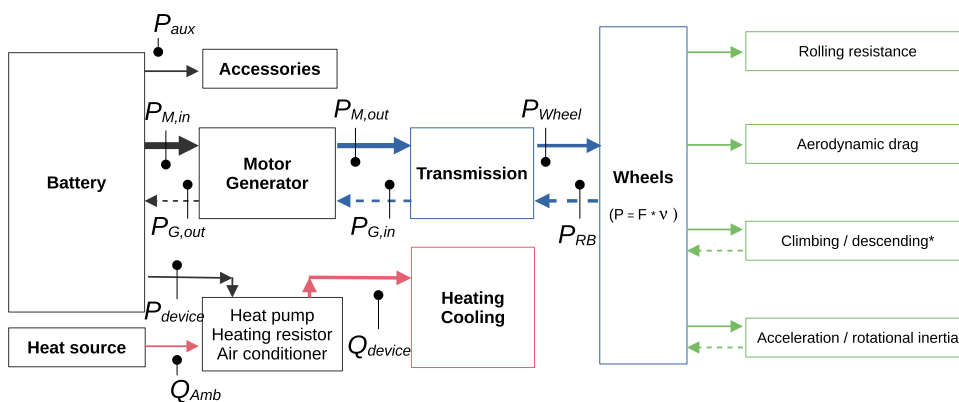
the vehicle can be obtained by selecting a BEV model. This includes nominal motor power, the battery energy capacity, the curb weight, the drag coefficient, height and width to calculate the frontal area, the gear ratio, and power-to-mass ratio (compare Table 4). Also, we make additional parameter assumptions associated with vehicles, such as battery charging and discharging efficiency, transmission system efficiency, cabin air volume, coefficient of performance of heat pumps and accessories' average power. The tool also requires passenger-related parameters, such as average weight and sensible heat, and the average number of passengers. Ambient temperature as well as driving cycle assumptions are also required. *emobpy* has access to three types of datasets: a) Hourly temperature time series can be obtained for 39 European countries<sup>25</sup>; b) Parameters of 25 BEV models that can be retrieved from<sup>26</sup>; and c) cabin thermal insulation based on<sup>27</sup>. The default values used in our case study are defined in Tables 6 and 7. The second block consists of incorporating the temperature time series. Trip distance and duration are used to calculate the average velocity for each trip. The third block shows the steps for calculating the energy consumption for each trip. The respective trip average velocity and trip duration are used to generate a custom driving cycle from a standard driving cycle sub-class. In doing so, velocity and acceleration are simulated at high-resolution (per seconds), which enables us to calculate power flow and energy consumption as described in the following sections.

		Layers [j]						Area (m <sup>2</sup> ) [A <sub>k</sub> ]
		Laminated glass	Tempered glass	Metal	PU foam	Polyester	Fiberglass	
Zones [k]	Windshield	✓						1.7
	Side windows		✓					1.5
	Rear window		✓					1.4
	Rest			✓	✓	✓	✓	9.9
Thermal conductivity ( $\frac{W}{mK}$ ) [ $\lambda_j$ ]		0.6	1.38	60	0.02	0.64	2	
Layer thickness (mm) [ $x_j$ ]		4.5	3.5	0.9	58	2	1	

**Table 6.** Configuration of the vehicle cabin insulation<sup>27,32,36,37</sup>. Note: PU: Polyurethane.

Parameter	Unit	Value	Description	Reference
$\eta_{tr}$	%	95	Transmission efficiency	23,28,30
$\eta_{charge}$	%	90	Efficiency for battery charging	6,9,15,23
$\eta_{discharge}$	%	95	Efficiency for battery discharging	38
$h_{cabin}$	$\frac{W}{m^2K}$	10	Cabin air convective heat transfer coefficient	32,37
$m_p$	kg	75	Person mass	Own assumption
$q_{sensible}$	W	70	Person sensible heat of a driver or passengers	36,39
$n_p$	quantity	1.5	Number of passengers in the vehicle	40
$T_{amb}$	Celsius	Germany (2016)	Time series hourly temperature	25
$T_{cabin}$	Celsius	17 (20)	Target cabin temperature for heating (cooling)	Own assumption
$V_{cabin}$	m <sup>3</sup>	3.5	Air volume of vehicle's cabin	38
$\dot{V}_{in}, \dot{V}_{out}$	$\frac{m^3}{s}$	0.02	Input (output) air flow at ambient (cabin) temperature for ventilation	41,42
$P_{aux}$	W	300	Auxiliary power for electronic accessories and battery heating	23,28
Driving cycle	—	WLTC	Driving cycles	43
COP	—	2	Coefficient of performance. Values >1 imply the use of a heat pump; similar COP for heating and cooling assumed	44,45

**Table 7.** Parameters used for all BEV models to determine driving electricity consumption.



**Fig. 9** Block diagram of the power flows at the components of the electric vehicle while driving.  $P$ : power,  $F$ : forces,  $v$ : velocity,  $P_{aux}$ : auxiliary power,  $P_{M,in}$ : motor input power,  $P_{M,out}$ : motor output power,  $P_{device}$ : electrical power for heating/cooling devices,  $Q_{Amb}$ : heat transfer rate from ambient by heat pump,  $Q_{device}$ : heat transfer rate for heating/cooling,  $P_{G,in}$ : generator input power,  $P_{G,out}$ : generator output power,  $P_{Wheel}$ : power at wheels,  $P_{RB}$ : regenerative braking power. [Black lines: electrical power, blue: mechanical power, red: heat transfer rate, green: acting forces. Dashed lines represent flows related to regenerative braking. Line thickness indicates typical flow magnitudes].

To calculate a trip's energy consumption, we calculate the power requirements for vehicle traction, heating and cooling. We further include (customizable) assumptions on auxiliary power. Figure 9 shows the power flows between the battery and the wheels, the heating/cooling devices and the accessories.

**Custom driving cycles.** A custom driving cycle is required to simulate a vehicle's driving pattern based on the trip average velocity and trip duration. This is necessary to calculate the power flow of a vehicle journey. The *Worldwide Harmonized Light Vehicles Test Cycle* (WLTC) is the tool's default driving cycle. A driving cycle emulates driving velocity patterns in cities, suburban areas, or highways, represented by driving cycle sub-classes. Every driving cycle sub-class has an average velocity which is calculated, including stops. The tool first selects the sub-class, whose average velocity is closest to the current trip's average velocity. The driving cycle sub-class selected is divided by the average velocity of the sub-class and multiplied by the trip's average velocity to create a custom driving cycle. This approach modifies the original driving cycle only to a small extent. Finally, as driving cycles have a finite duration, the custom driving cycle is replicated sequentially until the total driving cycle length reaches the trip duration. Acceleration is calculated from the variation of the velocity.

**Vehicle tractive effort.** Tractive effort  $F_{te}$  is the force required to surpass the opposing forces to the movement of a vehicle, expressed in Eq. 1, where  $F_{rr}$  is the rolling resistance force,  $F_{ad}$  is the aerodynamic drag,  $F_g$  is the climbing force and  $F_{acc}$  is the linear acceleration and inertia force<sup>28</sup>.

$$F_{te} = F_{ad} + F_{rr} + F_g + F_{acc} \quad (1)$$

$$F_{ad} = \frac{1}{2} \rho \cdot A_{frontal} \cdot C_d \cdot \nu^2 \quad (2)$$

$$F_{rr} = f_{rr} \cdot m \cdot g \cdot \cos\theta \quad (3)$$

$$F_g = m \cdot g \cdot \sin\theta \quad (4)$$

$$F_{acc} = (m + m_i) \alpha \quad (5)$$

The aerodynamic drag force, as defined in Eq. 2, depends on  $\rho$  moist air density,  $A_{frontal}$  frontal area of the vehicle,  $C_d$  drag coefficient, and  $\nu$  vehicle's velocity. The rolling resistance force is displayed in Eq. 3, where  $f_{rr}$  is the rolling resistance coefficient,  $m$  is the vehicle mass,  $g$  is the gravitational acceleration, and  $\theta$  is the slope in radians. Climbing force is shown in Eq. 4, and linear acceleration and inertia force of rotating parts is presented in Eq. 5 where  $\alpha$  is linear acceleration, and  $m_i$  is the inertial mass, a mass that represents the inertia of moving parts<sup>23</sup>. The inertial mass is defined in Eq. 7 that depends on the curb mass of the vehicle  $m_c$  and the gear ratio  $r_{gear}$ , while the mass of the vehicle  $m$  is the sum of the curb mass  $m_c$  and the passengers mass  $m_p$  as shown in Eq. 6.  $f_{rr}$  is a parameter that depends on the ambient temperature  $T_{amb}$  and velocity according to Eq. 8. This equation is derived from empirical data<sup>29</sup>.

$$m = m_c + m_p \quad (6)$$

$$m_i = m_c (0.04 + 0.0025 r_{gear}) \quad (7)$$

$$f_{rr} = 1.9 \times 10^{-6} T_{amb}^2 - 2.1 \times 10^{-4} T_{amb} + 0.013 + 5.4 \times 10^{-5} \nu \quad (8)$$

**Motor power.** Power at wheels  $P_{Wheel}$  is estimated at each time step, as shown in Eq. 9 where  $F_{te}$  is non-negative. Otherwise,  $P_{Wheel}$  is zero and regenerative braking power takes the absolute value of  $F_{te}$  as shown in the following section. The output power of the motor  $P_{M,out}$  is defined in Eq. 10 where  $\eta_{tr}$  is the transmission system efficiency. The input power of the motor  $P_{M,in}$  depends on its output power and the motor efficiency  $\eta_m$ , as shown in Eq. 11.

$$P_{Wheel} = F_{te} \cdot \nu \quad \text{if } F_{te} > 0 \quad (9)$$

$$P_{M,out} = \frac{P_{Wheel}}{\eta_{tr}} \quad (10)$$

$$P_{M,in} = \frac{P_{M,out}}{\eta_m} \quad (11)$$

The motor efficiency  $\eta_m$  depends on the motor's angular speed and torque. This value can be determined experimentally for each vehicle model or can be provided by the manufacturer. We have implemented a more general approach described in<sup>23,28</sup> (Eq. 13). The efficiency function depends on the motor load fraction  $Load_m$  as defined in Eq. 12 where  $N_{motor}$  is the nominal power capacity of the motor.

$$Load_m = \frac{P_{M,out}}{N_{motor}} \quad (12)$$

$$\eta_m = f(\text{Load}_m) \quad (13)$$

*Regenerative braking.* Regenerative braking power  $P_{RB}$  occur when  $F_{te}$  is negative as defined in Eq. 14 where the absolute values of  $F_{te}$  is used.

$$P_{RB} = |F_{te}| \cdot \nu \quad \text{if } F_{te} < 0 \quad (14)$$

The input power of the generator  $P_{G,in}$  is described in Eq. 15 where  $\eta_{rb}$  is the efficiency of the regenerative braking. The regenerative braking efficiency represents the fraction of the regenerative braking power that can be effectively recovered. The Eq. 16 shows the regenerative braking efficiency is a function of the acceleration  $\alpha$ <sup>30</sup>.

$$P_{G,in} = P_{RB} \cdot \eta_{tr} \cdot \eta_{rb} \quad (15)$$

$$\eta_{rb} = \left[ e^{\frac{0.0411}{|\alpha|}} \right]^{-1} \quad (16)$$

$$P_{G,out} = P_{G,in} \cdot \eta_g \quad (17)$$

Assuming a generation efficiency  $\eta_g$ , we can estimate the output power of the generator  $P_{G,out}$  as indicated in Eq. 17. The load fraction of the generator  $\text{Load}_g$  is required to calculate the  $\eta_g$  as shown in Eq. 18, where  $N_g$  is the nominal power capacity of the generator that is in fact also the nominal power capacity of the motor. The dataset with corresponding  $\eta_g$  by  $\text{Load}_g$  is obtained from<sup>23,28</sup> (see Eq. 19).

$$\text{Load}_g = \frac{P_{G,in}}{N_{gen}} \quad (18)$$

$$\eta_g = f(\text{Load}_g) \quad (19)$$

*Heating, cooling and accessories.* We aim to estimate the power that an electric device has to provide for heating or cooling a vehicle cabin to keep the temperature on a level of comfort for the passengers. To do so, we use a heat balance model<sup>31,32</sup>. The heat balance equation is shown in Eq. 20. The left-hand side expression represents the amount of heat accumulated in the cabin air, where  $V_{cabin}$  is the cabin volume,  $\rho_{air, T_{cabin}}$  is moist air density at cabin temperature,  $C_p$  is the specific heat of air,  $T_{cabin}$  is the cabin temperature, and  $\frac{dT_{cabin}}{dt}$  is the temperature change in the cabin over time. The right-hand side expression of the heat balance considers the following mechanisms: a) enthalpy of outside air  $\dot{Q}_{inflow}$ , b) enthalpy of discharged air to outside  $\dot{Q}_{outflow}$ , c) heat transfer through the cabin walls  $\dot{Q}_{wall}$ , d) sensible heat of passengers  $\dot{Q}_{person}$ , and e) the heat provided by a device to keep the target temperature in the cabin  $\dot{Q}_{device}$ . The device may be either a resistor or a heat pump. Radiation heat transfer and latent heat by condensation/evaporation are features not considered in this model.

$$V_{cabin} \cdot \rho_{air, T_{cabin}} \cdot C_p \frac{dT_{cabin}}{dt} = \dot{Q}_{device} + \dot{Q}_{person} + \dot{Q}_{inflow} - \dot{Q}_{outflow} - \dot{Q}_{wall} \quad (20)$$

$$\dot{Q}_{inflow} = \rho_{air, T_{amb}} \cdot \dot{V}_{in} \cdot C_p \cdot T_{amb} \quad (21)$$

$$\dot{Q}_{outflow} = \rho_{air, T_{cabin}} \cdot \dot{V}_{out} \cdot C_p \cdot T_{cabin} \quad (22)$$

$$\dot{Q}_{wall} = (T_{cabin} - T_{amb}) \sum_{k=1}^n \frac{1}{R_k} \quad (23)$$

$$R_k = \frac{1}{A_k} \left( \frac{1}{h_{cabin}} + \sum_{j=1}^m \frac{x_j}{\lambda_j} + \frac{1}{h_{amb}} \right) \quad (24)$$

$$h_{cabin} = \text{constant} \quad (25)$$

$$h_{amb} = \begin{cases} 6.14\nu^{0.78}, & \text{if } \nu > 5 \text{ m/s} \\ 6.14 \cdot 5^{0.78}, & \text{otherwise} \end{cases} \quad (26)$$

$$\dot{Q}_{person} = q_{sensible} \cdot n_p \quad (27)$$

$$P_{device} = \frac{|\dot{Q}_{device}|}{COP} \quad (28)$$

$$P_{aux} = constant \quad (29)$$

The enthalpy of outside air  $\dot{Q}_{inflow}$  is described in Eq. 21, where  $\rho_{air, T_{amb}}$  is the moist air density in the ambient,  $\dot{V}_{in}$  is the volume inflow of air for ventilation, and  $T_{amb}$  is the ambient temperature. The enthalpy of discharged air to outside  $\dot{Q}_{outflow}$  is defined in Eq. 22, where  $\dot{V}_{out}$  is the output volume flow of air. The heat transfer through the cabin walls  $\dot{Q}_{wall}$  is shown in Eq. 23, where  $R_k$  is the heat transfer resistance and  $k$  is the set of cabin zones. The heat transfer resistances  $R_k$  is defined in Eq. 24, where  $A_k$  is the area of every cabin zone,  $h_{cabin}$  is the convection heat transfer coefficient between the cabin air and the vehicle wall,  $h_{amb}$  is the convection heat transfer coefficient between the wall and ambient air,  $x_j$  is the thickness of thermal insulation material of the wall,  $\lambda_j$  is the thermal conductivity, and  $j$  is the set of insulation materials. The cabin convection heat transfer coefficient  $h_{cabin}$  is defined in Eq. 25 where it has been assumed a constant value. Typical values are 10–20  $\frac{W}{m^2K}$ <sup>32</sup>.

The ambient convection heat transfer coefficient  $h_{amb}$  is defined in Eq. 26, where  $v$  is the outside wind speed, which we consider to be equal to the vehicle's velocity<sup>33</sup>. The sensible heat of passengers  $\dot{Q}_{person}$  is presented in Eq. 27, where  $q_{sensible}$  is the sensible heat per person and  $n_p$  is the number of passengers. The heat balance equation is solved for  $\dot{Q}_{device}$  to get the heat requirement. The electric power for the heating/cooling  $P_{device}$  is defined in Eq. 28, where  $COP$  is the coefficient of performance of the heater/cooler or heat pump. A constant power for accessories  $P_{aux}$  is assumed as shown in Eq. 29. To estimate the heat transfer that occurs by heat conduction, Table 6 displays the default insulation configuration used in *emobpy*.

**Energy consumption.** Positive or negative values can be expected at the battery  $P_{battery}$ . Suppose the sum of motor input power, generator output power, auxiliary power and power for heating/cooling  $P_{all}$  is positive (see Eq. 30). The battery then provides energy to the vehicle as it discharges (see Eq. 31) and the discharging efficiency is used  $\eta_{discharge}$ . If  $P_{all}$  is negative, then the battery is charged via regenerative braking. In such a case, the battery load  $P_{battery}$  is negative, hence the charging efficiency  $\eta_{charge}$  is utilized.

$$P_{all} = P_{M,in} - P_{G,out} + P_{aux} + P_{device} \quad (30)$$

$$P_{battery} = \begin{cases} \frac{P_{all}}{\eta_{discharge}}, & \text{if } P_{all} > 0 \\ P_{all} \cdot \eta_{charge}, & \text{otherwise} \end{cases} \quad (31)$$

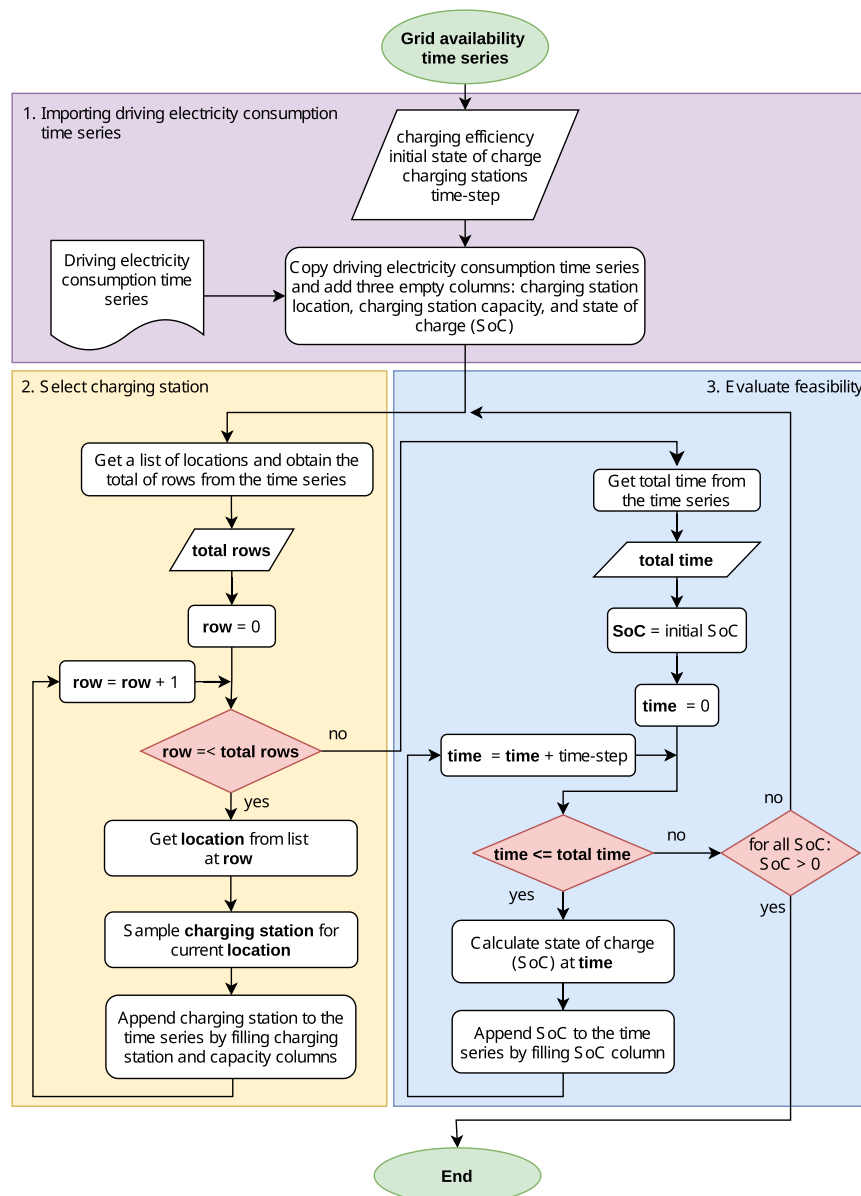
$$E_{total} = \sum_{t=1}^{\tau} P_{battery,t} \quad (32)$$

The total energy consumption per trip  $E_{total}$  is defined in Eq. 32, where battery load is aggregated through the set  $t$  that consists of the duration of the trip at every second. Table 7 provides parameters required for estimating the driving electricity consumption. For reasons of simplicity and data availability, we assume that these parameter do not differ between BEV models.

**Grid availability time series.** The flow diagram shown in Fig. 10 illustrates how *emobpy* creates the grid availability time series. Inputs are the time series of driving electricity consumption as well as locations and distances created in step 1. Further, *emobpy* requires data or assumptions on the battery size, charging efficiency, the initial state of charge (SoC), and the probability distributions of charging stations at different locations including their respective power rating, as indicated in the parallelogram in the first box of Fig. 10.

Initially, a time series containing the time step, location, distance, and consumption columns is imported from the driving electricity consumption time series. Next, different types of charging stations are selected for each time step. For each parking location (arrival time plus subsequent parking time steps until next trip), the types and respective power ratings are sampled from the corresponding probability distributions.

After a candidate grid availability time series is created, *emobpy* evaluates its feasibility. This check takes into account the driving electricity consumption time series of the profile as well as the charging station power rating available in each time step. To this end, the SoC of the battery is calculated for each time step by adding the energy taken from the grid for charging if connected to the grid, or subtracting the energy consumed from the battery if driving. For the first time step, we use an exogenous value. To simulate the SoC of the battery, we assume a charging strategy called *immediate - full capacity* as introduced in the section "Grid electricity demand". It draws electricity from the grid at full rating of the charging station as soon as the BEV is connected and until the battery is full. The following section provides more detailed information. After calculating the SoC for all time steps, *emobpy* verifies if each SoC lies within 0–100%. If this is the case, the allocation of charging stations throughout



**Fig. 10** Grid availability time series flow diagram.

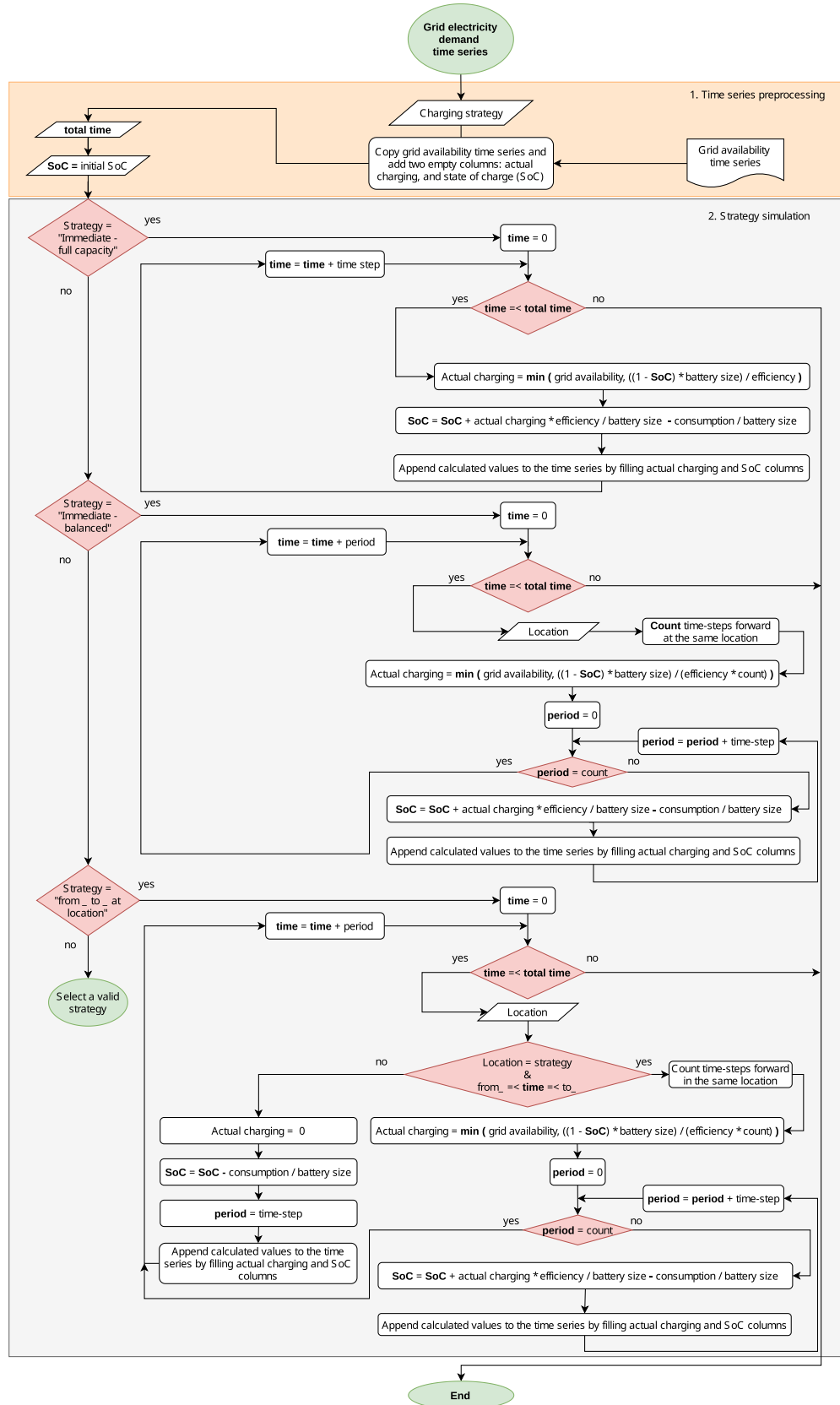
the time series horizon allows to create a grid availability time series. If this is not the case, a new allocation is carried out. In case of many unsuccessful allocations, *emobpy* returns a warning. Reasons comprise a low availability of charging stations and/or low power ratings compared to trip lengths or a low battery capacity.

**Grid electricity demand time series.** Figure 11 shows a flow diagram of how *emobpy* creates the grid electricity demand time series. The inputs are the grid availability time series, including the charging power rating, and the driving electricity consumption time series, including vehicle locations. Further inputs are data or assumptions on the battery size, initial SoC, and charging efficiency. Based on the inputs, *emobpy* calculates, for each time step, the SoC and, as output, the actual charging that represents the electricity drawn from the grid to charge the battery. To this end, two pre-set charging strategies (*immediate - full capacity*, *immediate - balanced*) or a customized charging strategy can be applied.

In the first pre-processing stage, *emobpy* imports the input data. In the second stage, charging – depending on the pre-set strategy – and the according SoC of the battery are determined.

For the strategy *immediate - full capacity*, *emobpy* iterates over all time steps without any foresight. It aims at reaching 100% SoC as fast as possible. If the current time step indicates grid availability, this strategy charges the BEV at the full power rating, except when less than the full rate is required to obtain 100% SoC. If the current time step corresponds to driving, actual charging is zero and the electricity consumed by the motor is subtracted from the SoC.

For the strategy *immediate - balanced*, *emobpy* also charges the BEV as soon as a grid connection is available. Yet, based on perfect foresight, the model executes its iteration over all consecutive time steps a vehicle is parking at the same location. To this end, the energy required to fill up the battery completely (100% SoC) is determined,



**Fig. 11** Grid electricity demand time series flow diagram for the charging strategies *immediate - full capacity*, *immediate - balanced*, and a customized charging strategy.



and the resulting value is divided by the number of time steps that the vehicle remains parked. The actual charging equals the maximum station power rating only in case a 100% SoC cannot be reached before the next trip. Otherwise, the actual charging rating is lower than the charging station power rating.

A customized charging strategy allows to derive alternative grid electricity demand time series. Such a strategy is passed to the model as text, e.g., *From\_23\_to\_06\_at\_home*. In this example, the actual charging occurs in the time window defined in hours of the day (23–06) and when the vehicle is parked at a predetermined location (*home*). The charging is performed in *balanced* configuration as described above. If a negative SoC is identified in the time series, the model may charge the battery outside the boundary defined by the customized charging strategy.

### Data availability

The dataset generated for the current study is available in the Zenodo repository <https://doi.org/10.5281/zenodo.3931663><sup>34</sup>.

### Code availability

The tool can be installed from the Python Package Index (PyPI) at <https://pypi.org/project/emobpy/>. The code is provided under a permissive license in Zenodo<sup>35</sup>. We also provide the script created to generate the 200 BEV profiles for the current case study at [https://gitlab.com/diw-evu/emobpy/emobpy\\_examples](https://gitlab.com/diw-evu/emobpy/emobpy_examples).

Received: 4 May 2020; Accepted: 21 May 2021;

Published online: 11 June 2021

### References

1. International Energy Agency. Global EV Outlook 2019. <https://doi.org/10.1787/35fb60bd-en> (2019).
2. de Coninck, H. *et al.* Strengthening and implementing the global response. In *Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty* (IPCC, 2018).
3. Daina, N., Sivakumar, A. & Polak, J. W. Modelling electric vehicles use: a survey on the methods. *Renewable and Sustainable Energy Reviews* **68**, 447–460, <https://doi.org/10.1016/j.rser.2016.10.005> (2017).
4. Mwasilu, F., Justo, J. J., Kim, E.-K., Do, T. D. & Jung, J.-W. Electric vehicles and smart grid interaction: A review on vehicle to grid and renewable energy sources integration. *Renewable and Sustainable Energy Reviews* **34**, 501–516, <https://doi.org/10.1016/j.rser.2014.03.031> (2014).
5. Richardson, D. B. Electric vehicles and the electric grid: A review of modeling approaches, impacts, and renewable energy integration. *Renewable and Sustainable Energy Reviews* **19**, 247–254, <https://doi.org/10.1016/j.rser.2012.11.042> (2013).
6. Taljegard, M., Walter, V., Göransson, L., Odenberger, M. & Johnsson, F. Impact of electric vehicles on the cost-competitiveness of generation and storage technologies in the electricity system. *Environmental Research Letters* **14**, 124087, <https://doi.org/10.1088/1748-9326/ab5e6b> (2019).
7. Kempton, W. & Tomić, J. Vehicle-to-grid power fundamentals: Calculating capacity and net revenue. *Journal of Power Sources* **144**, 268–279, <https://doi.org/10.1016/j.jpowsour.2004.12.025> (2005).
8. Fischer, D., Harbrecht, A., Surmann, A. & McKenna, R. Electric vehicles' impacts on residential electric local profiles – a stochastic modelling approach considering socio-economic, behavioural and spatial factors. *Applied Energy* **233–234**, 644–658, <https://doi.org/10.1016/j.apenergy.2018.10.010> (2019).
9. Kiviluoma, J. & Meibom, P. Methodology for modelling plug-in electric vehicles in the power system and cost estimates for a system with either smart or dumb electric vehicles. *Energy* **36**, 1758–1767, <https://doi.org/10.1016/j.energy.2010.12.053> (2011).
10. Muratori, M. Impact of uncoordinated plug-in electric vehicle charging on residential power demand. *Nature Energy* **3**, 193–201, <https://doi.org/10.1038/s41560-017-0074-z> (2018).
11. Robinson, A., Blythe, P., Bell, M., Hübner, Y. & Hill, G. Analysis of electric vehicle driver recharging demand profiles and subsequent impacts on the carbon content of electric vehicle trips. *Energy Policy* **61**, 337–348, <https://doi.org/10.1016/j.enpol.2013.05.074> (2013).
12. Schäuble, J., Kaschub, T., Ensslen, A., Jochem, P. & Fichtner, W. Generating electric vehicle load profiles from empirical data of three EV fleets in Southwest Germany. *Journal of Cleaner Production* **150**, 253–266, <https://doi.org/10.1016/j.jclepro.2017.02.150> (2017).
13. Schill, W.-P. & Gerbaulet, C. Power system impacts of electric vehicles in Germany: Charging with coal or renewables? *Applied Energy* **156**, 185–196, <https://doi.org/10.1016/j.apenergy.2015.07.012> (2015).
14. Heinz, D. Erstellung und Auswertung repräsentativer Mobilitäts- und Ladeprofile für Elektrofahrzeuge in Deutschland. Karlsruhe Institut für Technologie (KIT), Working Paper Series in Production and Energy, 30, <https://doi.org/10.5445/IR/1000086372> (2018).
15. Chen, X. *et al.* Impacts of fleet types and charging modes for electric vehicles on emissions under different penetrations of wind power. *Nature Energy* **3**, 413–421, <https://doi.org/10.1038/s41560-018-0133-0> (2018).
16. Wolinetz, M., Axsen, J., Peters, J. & Crawford, C. Simulating the value of electric-vehicles' grid integration using a behaviourally realistic model. *Nature Energy* **3**, 132–139, <https://doi.org/10.1038/s41560-017-0077-9> (2018).
17. Wulff, N. *et al.* Comparing power-system and user-oriented battery electric vehicle charging representation and its implications on energy system modeling. *Energies* **13**, <https://doi.org/10.3390/en13051093> (2020).
18. Lombardi, F., Balderrama, S., Quoilin, S. & Colombo, E. Generating high-resolution multi-energy load profiles for remote areas with an open-source stochastic model. *Energy* **177**, 433–444, <https://doi.org/10.1016/j.energy.2019.04.097> (2019).
19. Brown, T., Schlachtberger, D., Kies, A., Schramm, S. & Greiner, M. Synergies of sector coupling and transmission reinforcement in a cost-optimised, highly renewable European energy system. *Energy* **160**, 720–739, <https://doi.org/10.1016/j.energy.2018.06.222> (2018).
20. Kuhnimhof, T. & Nobis, C. Mobilität in Deutschland – MiD: Ergebnisbericht. <https://elib.dlr.de/125879/> (2018).
21. OECD. OECD Labour Force Statistics 2018. [https://doi.org/10.1787/oecd\\_lfs-2018-en](https://doi.org/10.1787/oecd_lfs-2018-en) (2018).
22. BMVI. Förderrichtlinie Ladeinfrastruktur für Elektrofahrzeuge in Deutschland. <https://www.bmvi.de/SharedDocs/DE/Anlage/G/foerderrichtlinie-ladeinfrastruktur-fuer-fahrzeuge-in-deutschland.pdf> (2017).
23. Iora, P. & Tribioli, L. Effect of ambient temperature on electric vehicles' energy consumption and range: Model definition and sensitivity analysis based on Nissan leaf data. *World Electric Vehicle Journal* **10**, <https://doi.org/10.3390/wevj10010002> (2019).
24. van der Walt, S., Colbert, S. C. & Varoquaux, G. The numpy array: A structure for efficient numerical computation. *Computing in Science Engineering* **13**, 22–30, <https://doi.org/10.1109/MCSE.2011.37> (2011).
25. De Felice, M. Country averages of Copernicus ERA5 hourly meteorological variables. *Zenodo* <https://doi.org/10.5281/zenodo.1489915> (2018).
26. EV Database. Electric Vehicle Database. <https://ev-database.org> (2021).
27. Wirth, S., Eimler, M. & Niebling, F. Thermal insulation of the passenger cabin of electric vehicles. *ATZ worldwide* **115**, 46–51, <https://doi.org/10.1007/s38311-013-0129-0> (2013).

28. Genikomsakis, K. N. & Mitrentsis, G. A computationally efficient simulation model for estimating energy consumption of electric vehicles in the context of route planning applications. *Transportation Research Part D: Transport and Environment* **50**, 98–118, <https://doi.org/10.1016/j.trd.2016.10.014> (2017).
29. Wang, J., Besselink, I. & Nijmeijer, H. Electric vehicle energy consumption modelling and prediction based on road information. *World Electric Vehicle Journal* **7**, 447–458, <https://doi.org/10.3390/wevj7030447> (2015).
30. Fiori, C., Ahn, K. & Rakha, H. A. Power-based electric vehicle energy consumption model: Model development and validation. *Applied Energy* **168**, 257–268, <https://doi.org/10.1016/j.apenergy.2016.01.097> (2016).
31. Shibata, S. & Nakagawa, T. Mathematical model of electric vehicle power consumption for traveling and air-conditioning. *Journal of Energy and Power Engineering* **9**, 269–275, <https://doi.org/10.17265/1934-8975/2015.03.006> (2015).
32. Klemm, D., Roessner, W., Widdecke, N. & Wiedemann, J. Reduced model of a vehicle cabin for transient thermal simulation. SAE Technical Paper <https://doi.org/10.4271/2018-37-0022> (2018).
33. Fujita, A., Ichi Kanemaru, J., Nakagawa, H. & Ozeki, Y. Numerical simulation method to predict the thermal environment inside a car cabin. *JSAE Review* **22**, 39–47, [https://doi.org/10.1016/S0389-4304\(00\)00101-6](https://doi.org/10.1016/S0389-4304(00)00101-6) (2001).
34. Gaete-Morales, C. emobpy: application for the German case. *Zenodo* <https://doi.org/10.5281/zenodo.3931663> (2021).
35. Gaete-Morales, C. An open tool for creating battery-electric vehicle time series from empirical data - emobpy. *Zenodo* <https://doi.org/10.5281/zenodo.3675456> (2021).
36. Brèque, F. & Nemer, M. Cabin thermal needs: Modeling and assumption analysis. *Proceedings of the 12th International Modelica Conference, Prague, Czech Republic, May 15–17* **132**, 771–781, <https://doi.org/10.3384/ecp17132771> (2017).
37. Rashid, R. M. Thermal management of vehicle interior temperature for improvement of fuel economy. PhD thesis <https://scholar.uwindsor.ca/etd/7564> (2018).
38. Zhang, C., Shen, K., Yang, F. & Yuan, C. Multiphysics modeling of energy intensity and energy efficiency of electric vehicle operation. *Procedia CIRP* **80**, 322–327, <https://doi.org/10.1016/j.procir.2019.01.058> (2019).
39. Torregrosa-Jaime, B., Bjurling, F., Corberán, J. M., Di Sciuillo, F. & Payá, J. Transient thermal model of a vehicle's cabin validated under variable ambient conditions. *Applied Thermal Engineering* **75**, 45–53, <https://doi.org/10.1016/j.applthermaleng.2014.05.074> (2015).
40. Deutscher Bundestag. Antwort der Bundesregierung. Drucksache 19/1162 <https://dip21.bundestag.de/dip21/btd/19/011/1901162.pdf> (2018).
41. Leong, J., Tseng, C.-Y., Tsai, B.-D. & Hsiao, Y.-F. Cabin heat removal from an electric car. *World Electric Vehicle Journal* **4**, 760–766, <https://doi.org/10.3390/wevj4040760> (2010).
42. Ott, W., Klepeis, N. & Switzer, P. Air change rates of motor vehicles and in-vehicle pollutant concentrations from secondhand smoke. *Journal of Exposure Science & Environmental Epidemiology* **18**, 312–325, <https://doi.org/10.1038/sj.jes.7500601> (2008).
43. Ciuffo, B. *et al.* Development of the worldwide harmonized test procedure for light-duty vehicles: Pathway for implementation in European Union legislation. *Transportation Research Record* **2503**, 110–118, <https://doi.org/10.3141/2503-12> (2019).
44. Bellocchi, S., Leo Guizzi, G., Manno, M., Salvatori, M. & Zaccagnini, A. Reversible heat pump hvac system with regenerative heat exchanger for electric vehicles: Analysis of its impact on driving range. *Applied Thermal Engineering* **129**, 290–305, <https://doi.org/10.1016/j.applthermaleng.2017.10.020> (2018).
45. Cuevas, C., Declaye, S. & Lemort, V. Experimental characterization of a reversible heat pump for hybrid and electric vehicles. *Advances in Mechanical Engineering* **11**, 1687814019845806, <https://doi.org/10.1177/1687814019845806> (2019).

## Acknowledgements

We thank three anonymous reviewers for valuable comments. We acknowledge a research grant of the German Federal Ministry of Education and Research (BMBF) via the START project (FKZ 03EK3046). Wolf-Peter Schill carried out parts of his work on this paper during a research stay at the Climate and Energy College at the University of Melbourne. The authors also thank all colleagues of the Open Energy Modelling Initiative for providing feedback during the Open Energy Modelling Workshop - Berlin 2020.

## Author contributions

Conceptualization, W.P.S. and A.Z.; Methodology, C.G., H.K., W.P.S. and A.Z.; Software, C.G.; Investigation, C.G.; Writing - original draft, C.G., H.K., W.P.S. and A.Z.; Writing - review & editing, C.G. and W.P.S.; Visualization, C.G.; Project administration and funding acquisition, W.P.S.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to W.-P.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021