



HAL
open science

A review of electric vehicle load open data and models

Yvenn Amara-Ouali, Yannig Goude, Pascal Massart, Jean-Michel Poggi, Hui Yan

► **To cite this version:**

Yvenn Amara-Ouali, Yannig Goude, Pascal Massart, Jean-Michel Poggi, Hui Yan. A review of electric vehicle load open data and models. 2020. hal-03028375v1

HAL Id: hal-03028375

<https://inria.hal.science/hal-03028375v1>

Preprint submitted on 27 Nov 2020 (v1), last revised 2 Apr 2021 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A review of electric vehicle load open data and models

Yvenn Amara-Ouali^{a,*}, Yannig Goude^{b,*}, Pascal Massart^a, Jean-Michel Poggi^{a,c}, Hui Yan^b

^a*Laboratoire de Mathématiques d'Orsay (LMO), CNRS, Université Paris-Saclay, Faculté des Sciences d'Orsay, bat 307, 91405 Orsay, France*

^b*EDF Lab, 7 bd Gaspard Monge, 91120 Palaiseau, France*

^c*Université de Paris, IUT STID, 143 avenue de Versailles, 75016 Paris*

Abstract

The field of electric vehicle charging load modelling has been growing rapidly in the last decade. Various models have been applied to this new industrial problem. In light of the Paris Agreement, it is crucial to keep encouraging better modelling techniques for successful electric vehicle adoption. Additionally, numerous papers highlight the lack of charging station data available in order to build models that are consistent with reality. In this context, the purpose of this article is threefold. First, to provide the reader with an overview of the open datasets available and ready to be used in order to foster reproducible research in the field. Second, to review electric vehicle charging load models with their strengths and weaknesses. Third, to provide suggestions on matching the models reviewed to six datasets found in this research that have not previously been explored in the literature. The open data search covered more than 860 repositories and yielded around 60 datasets that are relevant for modelling electric vehicle charging load. These datasets include information on charging point locations, historical and real-time charging sessions, traffic counts, travel surveys and registered vehicles. The models reviewed range from statistical characterisation to stochastic processes and machine learning and the context of their application is assessed.

Keywords: Electric Vehicles, Charging Point, Load Modelling, Smart

*Corresponding authors

1. Introduction

Assuming a low-carbon energy mix, Electric Vehicles (EVs) are a credible alternative to internal combustion engine vehicles (ICEVs) supporting the transportation sector in its low-carbon transition. A substantial number of governments are heavily investing in electric mobility with more than 5.1 million electric passenger cars on the roads globally in 2018, according to the International Energy Agency (IEA) [1]. Several countries are achieving high rates of EV adoption such as Norway which approached an EV market share of almost 47% in 2019 [1]. This is due in large part to major incentives implemented by governments to foster EV uptake [2]. The EV30@30 Campaign [3] sets a target of 30% EV market share by 2030 for the member countries of the Electric Vehicle Initiative (EVI) [4]. This enthusiasm for EVs comes hand in hand with great concern about how to manage the surge in electricity demand which could greatly disrupt the current schedule [5].

In order to overcome potential pitfalls, businesses and researchers are proposing solutions including pricing strategies [6] and smart charging [7]. The goal of these solutions is to avoid dramatically shifting EV users' behaviours and power plants production schedules. However, their implementation requires a precise understanding of charging behaviours. Thus, EV load models are necessary in order to better understand the impacts of EVs on the grid. With this information, the merit of EV charging strategies can be realistically assessed.

In this article, the term "EVs" refers to small vehicles (e.g., light motorcycles), passenger vehicles (e.g., cars) and goods-carrying vehicles (e.g., trucks) as per the classification from the European Commissions' official "Mobility & Transport: Vehicle Categories" [8] [9]. Passenger vehicles constitute the majority of EVs. Additionally, all energy system management are considered: Battery EV, Fuel-Cell EV or Plug-In Hybrid EV [10]. Furthermore, Electric Vehicle Supply Equipment (EVSE) will be referred to as any type of charging point, be it

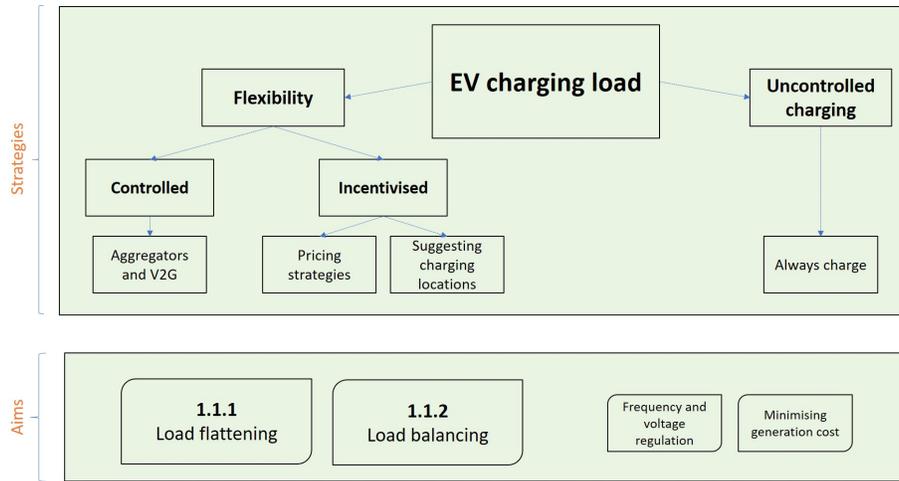


Fig. 1. The variety of strategies envisaged for handling EV load and their associated aims. Incentivised flexibility and controlled flexibility are used to specifically achieve energetic and financial aims while uncontrolled charging lets the market decide the prioritisation of these aims. Load flattening and load balancing are the most common aims found in the literature and they are the focus of paragraphs 1.1.1 and 1.1.2.

public or private. Finally, an EV charging session (or transaction) refers to the period of time an EV has spent charging at an EVSE.

1.1. Aims and strategies for EV charging schemes

Electricity distribution occurs such that at any point in time and space, the consumption has to be equal to the production in order to avoid severe consequences such as blackouts [11]. A significant rise in the number of EVs in circulation leads to an increase in electricity demand which could cause such a blackout if the balance in the grid is not effectively maintained. Therefore, EVs have an important role to play in maintaining this balance [5]. The purpose of this section is to explore the different aims and strategies required to overcome the potential difficulties caused by increased EV penetration. Figure 1 summarises these aims and strategies.

1.1.1. Load flattening

While some studies show minimal impact of EVs on peak load [12, 13], the consensus in the field is that by keeping up with business as usual, the grid will not be able to sustain its operations with such demand from EVs [6, 9, 14, 15, 16, 17, 18, 19].

One of the first articles dealing with the impact of EVs on load management was written in 1983 [20]. In this article, EVs were suggested as a way to minimise the overall grid load factor f . This factor is defined as the ratio of the average load (L) over the maximum load in a given period of time: $f = avg(L)/max(L)$. The maximisation of this quantity results in a more efficient distribution of resources over time. The article proposed that using off-peak recharging of EVs will significantly increase the load factor. This means shifting the EV demand to times when the rest of the demand is low (e.g., night time) in order to flatten the load curve. The flexibility analysis produced in [21] suggests that it is possible to shift the EV charging to the afternoon and night valleys for different clusters of users without changing their behaviours. This could lead to peak reduction and load factor maximisation with little change to users' requirements and lifestyles.

Articles such as [22] strived to estimate the benefits of this kind of controlled or incentivised EV charging. However, these articles do not always account for potential mistakes in load forecasting, therefore the benefits calculated could be inaccurate. Hence, it is critical to improve EV load forecasting models in order to alleviate the risk of unrealistic optimisation schedules for maximising the load factor.

1.1.2. Load balancing

An early article from 1997 [23] considered using EVs as a source of electricity for the grid when demand is high. In other words, using EVs plugged-in to the grid as an ancillary service or as way to bring flexibility to the overall shape of the load. According to a study focused on 400,000 EV charging transactions from 2012 to 2016 in the Netherlands, 75% of EVs connected at public EVSEs are already fully charged [24]. This study therefore supports the strategy of

using fully recharged EVs which are still connected, as a source of energy in order to supply the grid. This paradigm shift, using what could be a major constraint and treating it as an opportunity, is called “Vehicle-to-grid” (V2G).

Additionally, integrating renewable energy sources is also the focus of numerous studies [7]. Many countries with climate related commitments are aiming to increase the share of renewables in their energy mix. However, the main drawback of renewable energies is their intermittent delivery of supply. Indeed, solar panels and wind farms are highly weather-dependent. In this context, EVs can adequately complement the energy coming from renewable power plants and operate as an ancillary service to the grid. This strategy consists in considering multiple EVs acting as a large battery or electricity storage system which can be discharged back into the grid when weather conditions do not allow renewable power plants to produce enough energy [19].

Although V2G has many advantages, one drawback is that it reduces battery lifetime by adding unnecessary cycles of charge and discharge to the vehicle [25]. Furthermore, this strategy requires the existence of global and local communication and monitoring channels which do not exist yet. These channels are necessary for the development of EVs in general and particularly for V2G and load balancing [26, 27]. Finally, in order to ensure effective communication, EV load models are critical as they can reduce uncertainty and minimise contradicting signals from what is expected and what is observed by operations management.

1.2. Paper Structure and Contributions

The purpose of this article is to enable a better understanding of EV load data available and models produced in the literature. The main contributions of this article are as follows:

- The results of an in-depth open data search with a structured list of datasets available for use

- A comprehensive review of EV load models including their strengths, weaknesses and their application in the literature
- A preliminary study on matching EV load models to six open datasets found in this research and not previously explored in the literature

The rest of the article is structured as follows. Section 2 defines EV load and its most common drivers. Section 3 presents the open data found which can be used to model EV load. Section 4 reviews EV load models comparing the different approaches taken. Section 5 explores charging session data not previously explored in the literature and suggestions are provided on the models reviewed that could be applied to these datasets. Finally, Section 6 highlights the current knowledge gaps and discusses the different options in order to pave the way for future work.

2. EV load and its main drivers in the literature

2.1. EV load as a model output

EV load corresponds to the power or energy consumed at all EVSEs in scope over time. This can be referred to as the aggregated output of EV load models. The same model output can be envisaged in a disaggregated fashion. Two setups are widely used in practice. The first is vehicle-centric which considers the contribution of each EV member of a fleet to the aggregated load. The second is EVSE-centric which considers the perspective of one or multiple EVSEs. Neither approach is mutually exclusive and the two setups can be combined to model EV load.

2.1.1. Aggregated

The aggregated approach is shown in various articles such as [24] and [28] where the total EV load across multiple EVSEs is modelled. In [24] 1,750 charging stations (2,900 charging points) are used while [28] uses a single station with many charging piles. This kind of approach usually performs well due to

the smoothness of the aggregated load curve assuming there are enough EVs or charging stations in scope. While they bring a holistic view of the charging load, they can lack detail with regards to the temporal and spatial distribution of the load which is one of the key concerns raised in the literature [5].

2.1.2. Vehicle-centric

In order to explore the finer details of EV load, a vehicle-centric approach can be adopted. In [29] all individual EV loads are calculated and summed up in order to recover the aggregated load. This approach can be qualified as a vehicle-centric approach as it uses individual outputs of EVs. However, it is only significant in terms of aggregated load and does not hold any particular veracity at a finer grain. The same comment can be made for [30] and [31]. A similar study can be found in [32] where the EV load outputs are separated into urban and rural behaviours while [13] looks at public and residential charging. This can give a better understanding of the spatial and temporal properties of EV load.

Additionally, more elaborate models which consider EV charge at a finer scale are detailed in [9] and [18]. For instance, in [18] four schedules for EVs are identified which enables one to better distinguish and evaluate their temporal impact on the grid. Furthermore, the spatial dimension is addressed by modelling EV charging locations. Both outputs are brought together in order to better model the aggregated EV load.

2.1.3. EVSE-centric

The EVSE-centric approach is rare in the literature as it usually is superseded by the vehicle-centric approach. However, there are some occurrences of such work for instance in [12] where residential charge is envisaged from each household perspective. The authors used a bottom-up approach to forecast the aggregated EV load using each household individual load. The debate of using EVSE-centric over vehicle-centric approaches is illustrated in [33]. In this article, it was found that both approaches yield comparable prediction errors even

though the EVSE-centric approach was slower to compute.

2.2. Input data used for EV load modelling

2.2.1. Battery

Battery inputs are variables which closely relate to the charging demand of EV load from a “physical/chemical” perspective. The most common ones used across the literature are the State of Charge (SoC), Energy Consumption (E) and battery capacity (C). Generally speaking, the SoC is the rate at which the battery is charged whether the EV is plugged-in, idle or travelling [34]. The SoC when the EV arrives at an EVSE is a critical influential factor of EV demand. This is referred to as the initial SoC (SoC_{init}) in the literature. By incorporating the distance travelled (D) by the EV, the following formula can be written:

$$SoC_{init} = \frac{C - E \times D}{C} \quad (1)$$

with C in kWh, E in kWh/km, D in km and SoC_{init} in %.

On one hand, battery capacity and other engine specifications are usually assumed to be known constants. Based on EU MERGE data, probability functions were derived to characterise EV specifications in [9]. On the other hand, the SoC and energy consumption evolve over time and with vehicle usage. Both are highly correlated and they can be deduced from each other from the formula above or by a set of assumptions. For instance, the initial SoC of EVs is assumed to be equal to 0%, 30%, 60% in [35] to match different scenarios. Similarly, the initial SoC is used as an input in [36] along with D , C and the charging rate of the EV charging model in [37]. Furthermore, in [30] and [38] battery specifications and stochastic characteristics are also part of model inputs. Finally, the EV load itself can be used as an input when considering time series approaches [33].

2.2.2. Travel

From this literature review, it appears that travel behaviours are the most widely used exogenous factor driving EV load models. It is important to distin-

guish between travel inputs extracted from travel surveys [14, 39] or estimated pattern data [30, 40, 9] (which usually require further statistical treatment to be part of a model), and real-world traffic patterns (which are deduced either from pilot experiments [41] or direct GPS driving data [42]).

The input variables used in most papers (whether they are estimated or recorded) are the daily distance travelled and travel time. In [42] daily travel distance and individual trip distances distributions were extracted from a survey conducted between 2012 and 2013 in Beijing with real-world GPS data collected on 112 volunteer vehicle owners. Likewise, a pilot experiment was put in place for a week in Germany [41] in order to record the evolution of daily trips through GPS data.

When such exact data is not available, researchers use travel surveys instead [43]. These datasets hold valuable general information on drivers behaviours and can be used in order to estimate parameters of statistical distributions for daily distance travelled or travel time. However, they can lack accuracy as the information is usually collected through questionnaires. For instance, in [39] the authors used the 2009 National Household Travel Survey (NHTS) as well as the New-York State Transportation Federation Traffic Data Viewer in order to extract traffic statistics such as EV speed travelling from one charging station to another. In [14] daily trips from a single real-world vehicle from the NHTS is randomly assigned to a fictional EV used in the model. This procedure is applied to the desired number of EVs to obtain a fictional EV traffic. Similarly, [30] used Barcelona’s mobility patterns while [44] used the 2008 transportation data from the Dutch Ministry of Transportation in order to extract traffic statistics.

2.2.3. Weather

EV load models have stemmed from electrical load models. They have been developed over 100 plus years [45] and are comprised of some strongly established characteristics. One such characteristic is the thermosensitivity of electrical load [46]. In short, this means that some obvious patterns can be derived from analysing both load demand and temperature. Thus, it is natural that the most

frequently used input for EV load models is temperature and its traditionally associated statistics (e.g., average, maximum, minimum) [12]. Even though temperature is used in most electrical load models it is rarely used in EV load models. Nevertheless, there exists reasonable arguments to include weather data in EV load models.

[47] explores the influence of different weather variables on daily EV charging demand. This includes, minimum, maximum and mean daily temperature as well as mean wind speed, maximum gust, rainfall, global radiation and sunny hours. The results of this study showed that temperature and specifically mean air temperature is the most correlated weather input to daily EV load relative to the others reaching a 27% correlation relationship in one of the regions considered.

Similarly, in [6] the authors argue that temperature can be used to model EV load as it is correlated to electricity prices and demand. However, there is no mention of other potential weather factors which could be included.

A relational analysis is used in [48] to assess the impact of weather factors on traffic volume in South Korea. It was found in this case study that maximum and average temperature as well as average humidity are the most influential weather factors on traffic volume. Average wind speed on the other hand is less influential and was discarded in their model.

Finally, it is also argued in [49] that temperature has a great impact on EV charging station load while wind and humidity are discarded.

2.2.4. Economy

Amongst the articles covered only a few include economical factors such as electricity prices [28, 6, 39], GDP [30] or trends [47, 40]. While some locations still provide free charging as an incentive to foster EV adoption, most public EVSEs have a charging price based off a subscription or peak/off-peak tariffs. However, China is one of the countries where real-time electricity pricing affects the price consumers pay at EVSEs. Thus, for [28] it is natural to include time of use tariffs as this study was made on Chinese EVs. In [39], the authors

also include electricity prices as it can have an impact on the decision making undertaken by an EV driver when choosing which station to charge their vehicle.

Interestingly, GDP is included as a model input in [30] as it was shown in previous work [50, 51, 52] that GDP and other socio-economical variables such as place of residence and household characteristics have an impact on EV load and can be leveraged using an vehicle-centric approach. This is something worth exploring as these variables are easily accessible in travel surveys and general country statistics. They can be used to better anticipate charging behaviours in various locations of the grid. Global EV trend usage with uptake scenarios [40] or calculated trends [47] can also be used as model inputs.

2.2.5. Calendar

Temporal inputs are used in most model set-ups. They are easy to integrate and bring consistency as well as performance with the strong explanatory power they hold. They require no heavy statistical treatment as opposed to other variables (e.g., travel and battery) which makes them easy to use. For instance, in [31] and [8] day of the week and time of day are used in EV load models and more generally, EV load is derived in most research papers from day of the week, time of day and seasonal variation.

3. Open Data Search

3.1. Overview

The purpose of this section is to provide an understanding of the type of open data available globally which can be used in order to model EV load. Few review articles that deal with related topics to EV load modelling have included information regarding open data with associated references [53]. To the best of the authors' knowledge, there exists no article at the time of writing which has attempted this type of endeavour for EV load models. Indeed, a great majority of articles produced in the EV load modelling domain are based off simulated

data or information owned by private entities which are very rarely made available [28, 49]. This prevents reproducible work and slows down research in the field.

Therefore, the objective is to fill this gap by providing the community with a structured and carefully selected list of open datasets ready to be used in order to foster data-driven research in the field. This open data search was possible in great part thanks to the open data inception initiative which gathers links to more than 3,500 open data repositories on their website all across the world [54].

Links to the datasets are provided throughout this section and are up to date at the time of writing. However, the links may change over time and new datasets may emerge. Supplementary materials are available at this github repository [55] with the exhaustive list of resources gathered as an outcome of this research.

3.2. Research Criteria

This study focuses on datasets which give information on transactions between EVs and EVSEs. In other words, charging sessions.

Additionally, datasets holding information on exogenous variables such as traffic, travel surveys and air quality have also been considered. These variables are widely used in the domain in order to simulate travel behaviours especially when considering spatiotemporal models. Weather data is also used for EV load modelling and electrical load modelling in general [46]. However this type of information was excluded from this data research as global resources which provide high quality weather data already exist. For example, the *riem* package [56] written in R retrieves data from airport weather stations all over the world via the Iowa Environment Mesonet website. Alternatively, the National Oceanic and Atmospheric Administration (NOAA) also provides extensive weather data [57].

In terms of the perimeter of this research, the top 14 countries active in the EVI during the period covering 2018 to 2019 have been targeted. They

are ranked by market share of electric cars according to the IEA [1]. This list includes, Norway, Iceland ¹, Sweden, Netherlands, Finland, China, Portugal ², USA ³, Canada, France, New Zealand, United Kingdom, Germany and Japan [59].

Most of the repositories covered used native language, therefore, the use of direct query search was minimised as it can be approximate, especially in a foreign language. Thus the following standardised process was used for each repository covered: every time a categorical hierarchy was available, datasets under the following categories were searched for: “Environment”, “Natural Resources”, “Infrastructure”, “Transportation”, “Traffic”, “Climate & Weather”, “Urban Development”, “Planning”. If a category search was not enabled, then the following key words were used with their translated variants: “Travel (Survey)”, “Electric Vehicle (or Car)”, “Charge-Charging”, “Traffic”, “Station”, “Air Quality”, “Mobility”.

3.3. Open datasets

Overall, more than 860 repositories have been explored and more than 60 relevant datasets have been found that are directly (endogenous) or indirectly (exogenous) useful for modelling EV load. Table 1. summarises the results found across all countries covered with the most relevant datasets in each category. Regarding EVSE data, a distinction is made between real-time and historical charging session data. Historical data gives information on charging sessions which occurred in the past. This is the essential type of data sought to model EV load. Real-time data refers to EVSE occupation information which is updated on short time frames (every few minutes) and not stored. It requires regular

¹Iceland is not officially part of the EVI but has the same ICEV ban by 2030 target than other countries in the EVI and is often mentioned alongside them in IEA reports and charts [58]. The magnitude of its EV market share makes it highly relevant to this analysis.

²As an observer

³Participation was being assessed at the time and at time of writing the USA are no longer part of the EVI

scraping to be transformed into a historical charging session dataset and only then can it be leveraged for EV load modelling.

Countries	EVSE data			Exogenous data		
	Location	Charging Sessions		Traffic Counts	Travel Survey	Registered Vehicles
		Historical	Real-time			
Norway	[60]		[60]	[61] [62]		[63]
Iceland				[64]		[65]
Sweden	[66]		[60]			[67]
Netherlands	[68] [69]	[70]	[68]	[71]	[72]	[73]
Finland	[60]		[60]	[74]		[75]
China	[76]			[77] [78]	[79]	
Portugal	[80]					[81]
USA	[82]	[83] [84] [85] [86]		[87]	[43] [88]	[89]
Canada	[90]	[91]		[92]	[93]	[94]
France	[95]	[96] [97]	[98]	[99] [100]	[101]	[102]
New Zealand				[103]		[104]
UK	[105] [106]	[107] [108]		[109] [110]	[111]	[112]
Germany	[113]		[114]	[115] [116]	[117] [118]	[119]
Japan					[120]	[121]

Table 1. The most relevant open dataset available found in this research with the associated references.

For each country the corresponding EV market share from the IEA [1] is provided as well as an estimated value of the number of EVs to which this market corresponds. The estimated number of EVs sold is calculated by using the number of passenger sales in 2019 given on [122] multiplied by the EV market share from the IEA [1]. In Figure 2, the national EV market share and estimated number of EVs sold are shown, coloured by the type of data available for each country. It is interesting to note that countries with the highest market share and number of EVs sold are not the ones for which historical charging session data was found. First of all, countries for which historical charging session data was found will be discussed as it is the most relevant and rarest information to find. Then, the information available from countries without historical charging

session data but with real-time charging session data will be outlined. Finally, the countries where only traffic information is available will be presented.

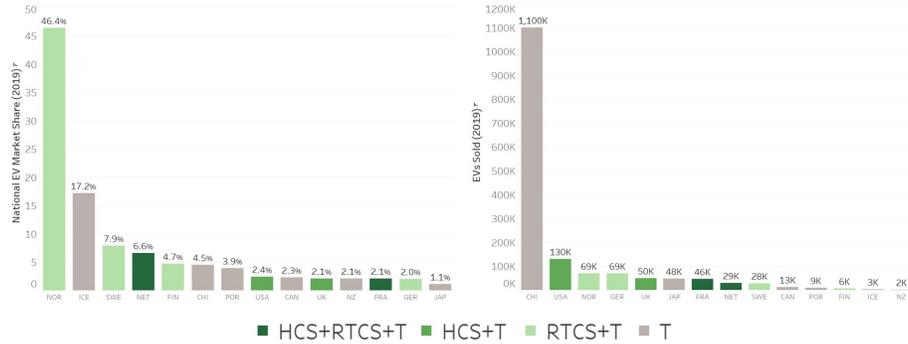


Fig. 2. For each of the 14 countries in scope, the national EV market share [1] and the estimated number of EVs sold [122] is shown. HCS refers to historical charging session data, RTCS refers to real-time charging session data and T refers to traffic counts and/or travel survey data. This demonstrates the existing gap between EV penetration in each country and the availability of open charging session data.

3.3.1. Countries with historical charging session data

Netherlands. 6.6% national EV market share [1] equating to approximately 29,000 EVs sold in 2019 [122]

23 repositories were covered in the Netherlands with every type of relevant data found. First of all, ElaadNL [70] holds historical charging sessions which were studied in multiple papers [21, 24]. With regards to traffic data, the *Onderweg in Netherlands* (ODiN) is the national travel survey published on a yearly basis [72]. While its tables are quite hard to study as-is for non-native speakers, they are summarised in another website in English [71]. Real-time data on utilisation and consumption at public EVSEs installed in Rotterdam can be found on the EV-BOX website which is one of the EVSE providers [68]. Registered vehicles [73] and public EVSE locations are also available (e.g., in Eindhoven [69]). Additionally, historical traffic data from 2010 extracted from 24,000 measure points which stores information on vehicles such as speed and travel time [123] was found.

USA. 2.4% national EV market share [1] equating to approximately 130,000 EVs sold in 2019 [122]

The open data search for the USA was extensive. Around 370 repositories were covered in the analysis. Among them three relevant charging session datasets were found [83, 84, 85]. The first provides a continuous dump of session data from 2018 on EV sessions recorded at city-owned EVSEs in Boulder (Colorado) [83]. The second gives the same information for charging sessions of EVs in the city of Palo Alto (California) from 2011 to 2017 [84]. Finally, the third provides us with an aggregated monthly view of transactions in the city of Evanston (Illinois) between 2016 and 2017 [85]. Furthermore, a charging session open dataset from Caltech, which is continuously updated in collaboration with Power Flex, is available at [86] and an exploration of this dataset was produced in [124]. On top of these charging session datasets, EVSE locations are also available from the Alternative Fuels Data Center [82], as well as many travel surveys including the National Household Travel Survey (NHTS) [43], which are frequently used to simulate EV behaviours from conventional vehicles. In particular, a mobility survey was performed in April 2019 for the City of Boulder on 203 residents. Information extracted from [88] brought together with the EV charging session dataset of the city of Boulder [83] could lead to more consistent and accurate representation of EV load than by using the more general NHTS. Finally, a large proportion of states share traffic volumes in various municipalities across the country (e.g., the city of Houston [87]).

France. 2.1% national EV market share [1] equating to approximately 46,000 EVs sold in 2019 [122]

France (mainland) also has a large number of open data repositories. In total, 151 repositories were explored. Among them all kinds of relevant data were found. Firstly, charging sessions were recorded from April to May 2017 on Belib' stations in Paris [96]. Furthermore, the Paris Data website provides the Belib' real time availability of public EVSEs in Paris [98] which can be scraped on a regular basis via an API in order to reconstruct a historical dataset. Regarding

private EVSEs, the charging sessions of a fleet of EVs owned by SAP Labs France have been recorded from June 2017 [97]. This dataset is updated every three months. On top of charging session data, registered vehicles across the territory [102], traffic counts in numerous cities [99], real-time traffic [100], and a national travel survey [101] are available in order to perform a spatiotemporal analysis of EV load. Different road traffic open data repositories are gathered on the Cerema website [125].

United Kingdom. 2.1% national EV market share [1] equating to approximately 50,000 EVs sold in 2019 [122]

72 repositories were covered for the UK mainland which yielded multiple charging session datasets. Two of them are situated in Scotland: Dundee City [107] and Perth & Kinross City Councils [108]. The former gathers two years of charging session data from 2017 to 2018 while the latter covers four years from 2016 to 2019 to the granularity of each session. Additionally, the UK government has led an EVSE analysis over the year 2017, with domestics [126], and public [127], [128] chargers. The raw datasets available include charging session data for each type of EVSE. There were also some initial trials led by the UK power networks in 2013-2014 which can yield useful information [129]. Public EVSE locations are available in numerous municipalities of the UK [106] as well as a national charging point registry [105] with real-time [109] and historical traffic counts [110]. Moreover, yearly national surveys are also available [111].

3.3.2. Countries with real-time but no historical charging session data

Norway. 46.4% national EV market share [1] equating to approximately 69,000 EVs sold in 2019 [122]

Norway is by far the country which has the highest penetration rate of EVs to date. Thus, it is no surprise that some highly relevant data for EV load modelling was found regardless of a relatively small number of repositories explored (13). Norway was an early-mover in fostering EV adoption. In 2009, the first large investments were made by cities and the government with Oslo

being the major contributor [2]. The most relevant data feed comes from the NOBIL database API [60]. This service provides (after benefiting of an API key from NOBIL) real-time information on EVSEs all across Norway, Sweden, Finland and Denmark (e.g., location, usage, details). Historical dumps do not seem to be available through the API, however a regular scraping may be put in place in agreement with NOBIL in order to reconstruct historical data. Other data sources which describes exogenous variables are available such as traffic volumes [61], [62] and vehicle registrations by fuel types [63] which gives an overview of the trend in EV adoption.

Sweden. 7.9% national EV market share [1] equating to approximately 28,000 EVs sold in 2019 [122]

Sweden, with 18 repositories covered, also benefits from the NOBIL API which gathers real-time information on public EVSEs activity across the territory [60]. Some of NOBIL's information is gathered on an external Swedish website which provides historical statistics on EV public charge use [130]. On top of that data source, the map of public EVSEs [66] and statistics on newly registered vehicles per county, town and fuel type on a monthly basis are also available [67]. This latter dataset can be used in load forecasting models as a variable explaining the trend in EV usage particularly thanks to its monthly granularity.

Finland. 4.7% national EV market share [1] equating to approximately 5,700 EVs sold in 2019 [122]

Finland is also one of the countries which has adopted the NOBIL database API [60]. Amongst the 20 repositories covered, exogenous information with traffic in real time in a few municipalities (e.g., the city of Tampere was found [131]) as well as registered vehicles between 1922 and 2019 [75] and average distance travelled by vehicles between 1980 and 2015 [74]. Even though these sources, provide us with extensive historical data, the most recent years are the most relevant for EV load models. These datasets can give an overall understanding of the overall traffic trends in Finland.

Germany. 2% national EV market share [1] equating to approximately 69,000 EVs sold in 2019 [122]

With regards to Germany, the most relevant datasets found among the 52 repositories covered were real-time public EVSE usage [114] and real-time traffic data [115] in the city of Bonn. Scraping both sources and associating these can lead to precise EV load models. In addition, travel surveys at fine levels of details are available from the German Mobility Panel [118] as well as the Rheinisch-Westfälisches Institute (RWI) [117]. The RWI dataset was used for a study on mobility patterns in [132]. Furthermore, the number of vehicles registered [119] and traffic counts in several municipalities [116] can give an understanding of the trend in EV usage across the country. Finally, as for most other countries, public EVSE locations are also available [113].

3.3.3. Countries with traffic data and no charging session data

Iceland. 17.2% national EV market share [1] equating to approximately 3,100 EVs sold in 2019 [122]

As for Iceland, 4 repositories were covered and the most relevant datasets found do not include any charging sessions but descriptive statistics on transports in Reykjavik [64] as well as vehicles distance and fuel consumption between 1995 and 2019 [65]. This can enable an understanding of the trends in EV adoption and high-level travel behaviours. However, limited analysis can be conducted as real charging session data is unavailable and would have to be simulated from other markets. Additionally, no real-world traffic data or travel survey was found which also limits spatial studies.

China. 4.5% national EV market share [1] equating to approximately 1,100,000 EVs sold in 2019 [122]

Being the country with the largest volume of EVs, China is at the forefront of EV deployment worldwide. However, this research did not result in finding any charging session data for China. One explanation for this is that more than 90% of EVSEs are owned by private firms [133]. Most of the articles which use

data from charging stations on Chinese territory do not make it available as it is usually part of an agreement between the researcher and the entity owning the data. Nevertheless some relevant traffic data [77], [78] for the whole territory was found and travel surveys [79] as well as EVSE locations [76] specifically in Hong-Kong.

Portugal. 3.9% national EV market share [1] equating to approximately 8,900 EVs sold in 2019 [122]

Regarding Portugal's open data, traffic statistics with the number of vehicles registered by type and fuel was found [81]. Additionally, EVSE locations in Lisbon were also available [80]. No charging session data was found.

Canada. 2.3% national EV market share [1] equating to approximately 13,000 EVs sold in 2019 [122]

Being the co-lead of the EVI activities along with China [1], Canada is a major player in the field of EV deployment. Around 76 repositories were explored with numerous travel surveys which describe various aspects of drivers' behaviours [93]. Traffic volumes [92] and EV registrations [94] are also available with details on EVSEs available for public use in some municipalities (e.g., the city of Edmonton [90]). Even though no historical nor real-time charging session data was found, there exists an EV Home Charging Program [91] which gathers residential charging session data. However, this dataset is not open at the time of writing but might be accessed with the relevant access grants.

New Zealand. 2.1% national EV market share [1] equating to approximately 2,300 EVs sold in 2019 [122]

For New Zealand, 22 repositories were covered with successful findings in traffic statistics and vehicle registrations. Several locations in New Plymouth record traffic count [103] and the number of vehicles registered by type across the country is also available [104]. This data as-is is difficult to exploit for EV load modelling as it lacks EVSE locations and charging sessions.

Japan. 1.1% national EV market share [1] equating to approximately 48,000 EVs sold in 2019 [122]

Finally, with Japan, 14 repositories which did not contain any charging session or station location information were covered. Nevertheless, exogenous data can be extracted with numerous travel surveys [120] and some statistics on registered vehicles [121].

4. EV load models

4.1. Statistical Characterisation

The goal of statistical characterisation models is to produce a distributional analysis for the outputs shall it be data-driven [24] or entirely deduced from exogenous variables such as travel data and statistical assumptions [14]. The different characterisations of EV load and proxy variables such as charging duration or inter-arrival time are summarised in Table 2.

Model	Strengths	Weaknesses	Ref
Gaussian	Particularly suited for large simulations	Unrealistic as negative values have a non-zero probability	[14]
Weibull, Lognormal, Exponential	Rapid to implement while providing an approximation consistent with reality	Fail to capture significantly diverse behaviours in the data	[13]
Mixtures (e.g., Beta, Gaussian)	Captures significantly different users' behaviours in the data and respects real-world constraints	Unsuitable for medium or large dimension problems with numerous covariates	[24] [134] [124] [135]
KDE	Highly versatile model as no explicit prior on the distribution is required	Weak interpretability power in addition to a sensitivity to outliers	[136] [137] [138] [139] [140]

Table 2. Statistical characterisation models for EV load

In [14], the authors did not benefit from any EVSE data. Nevertheless, they used the NHTS [43] ICEV behaviours from 2009 to derive EV travel patterns in order to simulate an EV fleet and characterise their behaviours. In their

work, the simulation showed that the power consumption can be seen as a normal distribution without any loss of accuracy. This can be true in practice, however, it is usually more consistent to assign distributions which are defined on \mathbb{R}^+ as it is unrealistic to observe negative power demand in that context. It is however convenient for model conciseness and computational speed.

In [13] a statistical analysis is conducted on data extracted from an EV trial conducted in Victoria (Australia) on 33 EVs on a 3-month period. This article showed that the Weibull distribution was the best fit for charging duration compared to the exponential and lognormal laws. They have also characterised the time to the next charging event as a mixture of two lognormal distributions. This is a vehicle-centric approach which considers the time to next charge from the EV perspective. These characterisations were used on a Monte-Carlo simulation which created 4,000 EVs by random sampling and assessed their overall impact on the grid. While these distributions are more consistent than a Gaussian distribution, they still fail to capture the irregularity of EV drivers' behaviours hidden in the data.

In [24], a dataset provided by Elaad NL [70] has been studied. This paper characterises EV load through a mixture of beta distributions. Its parameters are optimised by minimising the Root Mean Squared Error (RMSE) of the point-wise difference with the empirical distribution. Additionally, Kolmogorov-Smirnov testing was used to assess the goodness-of-fit. From the observations that weekly charging sessions present two peaks (namely a morning and a late afternoon peak) it was reasonable to consider a mixture of distributions to account for the different modes. In [134], 13 different charging session profiles were identified using Gaussian mixture clustering based on data provided by the G4 cities of the Netherlands. Other recent studies complement this work by using Gaussian mixtures to model the triplet (Arrival Time, Charging Duration, Energy Consumed) in order to characterise EV load. In [124] the triplet is modelled by a multivariate Gaussian mixture while in [135] only the couple (Charging Duration, Energy Consumed) is modelled by a Gaussian mixture with the Arrival Time modelled by an exponential distribution. The results

produced are more accurate than for elementary distributions. However, they are structurally limited to the joint use of few covariates which keeps from fully integrating exogenous information.

A few articles also modelled EV load with a kernel density estimator (KDE). Two main types have been used in the literature: the Gaussian kernel density estimator (GKDE) and the diffusion kernel density estimator (DKDE). These methods are highly versatile because no prior knowledge over the distribution is hypothesised. Thus, they can reach high accuracy when fitting empirical data at the cost of weak interpretability. Looking at [136], a GKDE is used to estimate daily trip distance and end time of the last trip. Both variables are critical for EV charging schedules and this method improves the accuracy of the distributions compared to parametric methods. A similar conclusion is drawn in [137] from a GKDE estimating the triplet (Arrival time, Charging duration, Charging capacity). In [138] and [139] the authors have compared both the GKDE and DKDE when estimating EV load. Thanks to its optimal bandwidth selection process, DKDE was found to produce better load estimations. Finally, in order to make the best of both GKDE (which is less sensitive to outliers) and DKDE (which has a higher overall accuracy), [140] has proposed a hybrid density estimator (HKDE). This HKDE reached significantly better root-mean square performance in estimating the EV load than the DKDE and GKDE on their own on the dataset used for this study.

4.2. Stochastic Processes

In the context of EV load models, three main types of stochastic processes have been detailed in the literature: purely temporal, spatiotemporal and queuing theory viewpoints. The various stochastic processes presented are summarised in Table 3.

Model	Strengths	Weaknesses	Ref
Temporal	Adequate for modelling one EVSE or one EV	Generally assumes independence between EVSEs	[15] [141] [37] [142] [44]
Spatiotemporal	Suited for modelling clusters of charging stations simultaneously	Large increase in complexity with scale	[9] [18] [32]
Queues	Easily scalable with strong theoretical grounds	Restricts reality with simplifying assumptions	[29] [143] [16] [39]

Table 3. Stochastic Processes for EV load

4.2.1. Temporal

One of the early works on EV temporal load models was completed in [15] where the authors explored the stochastic nature of EV load by using probabilistic travel patterns to determine initial SoC and starting time of battery charge. In particular, assuming battery type is known, recharge starting time is then assumed to be a random variable with a probability density function (pdf) determined by the tariff structure (scenarios) and patterns of EV usage. Initial SoC is also considered as a random variable dependent on the total distance travelled since last charge. Introducing a lognormal pdf for the daily distance driven, the initial SoC can be derived assuming a linear discharge (also assuming that it was fully recharged originally). Finally, they obtain a discretised version of the stochastic process of the load on half hourly intervals for a single EV which is then extended to an arbitrary number of EVs.

In [141], the authors defined a temporal stochastic process modelling charging patterns at a public EVSE with a Markov Chain comprising three states: unoccupied, charging and plugged-in but not charging. Essentially, the Markov Chains setup assumes that the current state of the process, conditionally to all past states, only depends on the previous state. It simplifies the calculation and has been extensively studied in the literature through many applications [144]. In [141], after initialising the transition probability matrix which drives

the path of the process they let the system evolve and assess the revenue made by the charging station.

Auto-regressive integrated moving averages (ARIMA) are a particular type of temporal process. Box and Jenkins [145] formalised a precise methodology to estimate the different orders of ARIMA processes. In [37] the ARIMA process is quantised on hours of the day. In other words, 24 sub-processes are estimated in their model. The final process obtained is thus a day-ahead hourly forecaster of EV load. In a following paper [142], they improved the performance of their model by forecasting separately conventional load and EV load. The results obtained in this paper reinforces the argument that EV load is structurally different from conventional load and requires specific load forecasting models.

Similarly, in [44] the authors modelled household EV load demand by using stochastic behaviours of three random variables: start-time of trip, end-time of trip and travelled distance. With a vehicle-centric approach, they present a Monte Carlo simulation method to derive overall system load. A particularity of this model is that it used a copula to characterise the multivariate distribution function of model variables. Then, using typical EV charging profiles, they derived the electricity demand at different EV uptake levels while observing the grid impacts.

Purely temporal models are particularly suited for one EVSE or one EV. They are not consistent for modelling cluster of EVSEs which require spatial considerations.

4.2.2. Spatiotemporal

Spatiotemporal models are usually designed for disaggregated approaches. The EV load at different stations is modelled separately using temporal features as well as travel patterns. They are rare in the literature as they require the combination or simulation of both the charging sessions and EV trips. Furthermore, they are limited as they cannot scale to large geographical scopes. Nevertheless, they can explore in fine details the intricacy of the relationship between EVs and EVSEs in specific regions.

In [9], the authors introduced a spatiotemporal model using Monte-Carlo simulation to specifically assess EV load demand in urban areas. The core of this method lies in the origin-destination analysis used to determine daily travel patterns of EVs. Additionally, probability functions to describe EV technologies were identified and using both information they ran a Monte-Carlo estimation of EV charging load for each busbar. By construction, this model can also be used for probabilistic assessment which indicates the branches most vulnerable to potential overloading.

In [18], the authors modelled both temporal and spatial stochastic aspects of PHEV owners behaviours to then derive their pdf. They modelled the temporal dimension with a uniform distribution for the start and end of charging time. As for the spatial dimension, they described the number of PHEVs arriving at an EVSE by a Poisson process according to driving behavior and traffic state. Assuming that both dimensions are independent, they derived the joint spatiotemporal pdf by multiplying both individual pdfs for charging times and arrival at EVSE. Ultimately, they expressed the effect on the daily load curve under various number of PHEVs for 150 PHEVs dispersed in the test system.

Finally, [32] proposed another probabilistic approach to characterise the spatiotemporal diversity of EV charging demand specifically on peak load demand. A Monte-Carlo simulation was used to evaluate the impacts of charging demand on the grid in urban and rural environments. It showed that this diversity of location helped the grid handle the demand better.

4.2.3. Queuing theory

Queuing theory models often use Kendall's shorthand notation which describes the arrival (A), the serving time (B) and the number of servers (C) in a compact form: A/B/C. EV load models are a suitable context for this theory as it was detailed in numerous articles [29, 39, 16, 38, 143].

One of the early works on EV load modelling was performed in [29]. This simple theoretical approach proposed to use an $M/M/n_{max}$ queue where the two first components characterises the Poisson processes for the number of EVs

arriving at a public EVSE and the number of EVs served while n_{max} refers to the number of maximum parallel charging EVs at charging points. A case study was proposed on the first car produced by Tesla, the Roadster Model, in order to assess the stochastic power demand output from the model.

The same queuing model was also used in [38] and was compared to a Monte-Carlo simulation in order to ultimately fit a distribution for the entire load demand of PHEVs. Additionally, in [143] the authors also used this queuing model and complemented it with a fluid traffic model in order to look at EV charging load on highway charging stations.

In a more general fashion, the authors of [16] have opted for an $M_t/GI_t/\infty$ queue where the number of arrivals follows an inhomogeneous Poisson process (indicating that the intensity function varies over time), the serving time is a general time-dependent distribution with an infinite amount of servers or EVSEs in the EV load context. Using some established results of queuing theory and previous work on estimating non-homogenous Poisson process rates, the authors managed to forecast each disaggregated intensity function for day-ahead forecasting. This paper is the only one found for stochastic processes applied to EV load which uses both travel patterns from the NHTS and real charging session data.

Thus, one important advantage of queuing network analysis applied in a spatiotemporal context of EV load is that it can capture interactions among multiple charging stations. In that sense, BCMP networks (named after their inventors: Baskett, Chandy, Muntz and Palacios) introduced in [146] were applied in [39] to produce an EV load model. BCMP networks are a type of queuing networks which yield a product-form stationary distribution. This kind of network is commonly used to study interconnected queues. In the EV load context, it means that it enables the model to take into account the potential shift of users from one station to another and control it to envisage different scenarios.

It is clear that queuing models are to be reserved for theoretical considerations rather than for operational implementation. Nevertheless, thanks to their solid mathematical foundations, they bring great insights for understanding EV

load behaviours especially when EVSE data is scarce.

4.3. Machine Learning

Four machine learning branches have mainly been explored for modelling EV load: Linear Model (LM), Support Vector Machine (SVM), Random Forest (RF) and Artificial Neural Network (ANN). In [31] the authors compare decision trees/tables, SVM and ANN. SVM demonstrated the best performance while the ANN and decision trees are ten times quicker to test on new data. A limitation of this work is that the dataset does not come from real charging session data but an aggregated distributional analysis produced by ECOtotality [147]. In [148], SVM, RF, k-Nearest Neighbours (k-NN) and a method called Modified Pattern-based Sequence Forecasting (MPSF) which uses k-means are compared. They found that SVMs and RF reach the best performance with regards to the Mean Absolute Error while k-NN and MPSF achieve better performance with regards to the Symmetric Mean Absolute Percentage Error. Since, k-NN and MPSF are much faster to compute predictions, they concluded that MPSF and k-NN were better suited for operational use. The different machine learning branches studied for EV load models are gathered in Table 4.

Model	Strengths	Weaknesses	Ref
Linear Model	Easily interpretable with fast implementation	Structurally limited for capturing complex and irregular patterns	[149] [150] [151]
Support Vector Machines	Easy implementation and effective in high-dimensional spaces	Not suitable for large and complex datasets which are not linearly separable even in high-dimensional spaces	[31] [33] [8] [148]
Random Forest	Versatile model with no prior assumptions on the shape of the data	Weak interpretability with no ability to extrapolate from training data	[33] [12] [152] [148]
Neural Networks	Can reach the highest level of performance	Architecture selection process can be laborious with long training time	[31] [33] [41] [49] [28] [153]

Table 4. Machine Learning models for EV load

4.4. Linear Model

It is common practice to start addressing a machine learning problem with simple models such as LM. In [149], and [150] LM was chosen as a first step to implement a smart charging strategy. This gives a more realistic operational context as opposed to other articles which skip predictive models before implementing an optimal charging strategy. Furthermore, in [151] an assessment of model inputs is presented using LM. They found that the voltage level of each EV had a critical influence over their model. However, these models are limited as they cannot capture irregular patterns in the data which is expected across EV drivers.

4.4.1. Support Vector Machines

SVM were originally defined by Vapnik [154]. In a nutshell, the idea behind this algorithm is to find the hyperplane which maximises the margin between different sets of populations. It is easy to implement but yields relatively long training times when working with large datasets.

In the context of EV load, SVM were compared to a Monte-Carlo forecasting technique in [8] and showed a better performance on a theoretical charging session dataset. Additionally, in [33] SVMs are used alongside other machine learning algorithms in order to model EV load from a vehicle-centric as well as from an EVSE-centric perspective. Because the EVSE-centric approach requires more data, it demonstrates a significantly longer running time as expected for SVMs. This study used a dataset extracted from UCLA campus parking lots. Thus, it is unlikely that these kinds of models will scale adequately for a larger scope of charging stations. Furthermore, articles using SVMs are now becoming rare as other alternatives with similar or better performances can be found.

4.4.2. Random Forests

RF is a learning algorithm which was popularised by Leo Breiman [155]. In short, it is an ensemble method which uses decision trees as elementary components for its construction.

On top of SVMs, [33] also used RF to model EV load. The few hyperparameters required to be tuned (e.g., number of trees, sampling rates) enables a fast and easy implementation with the possibility to iterate rapidly. In [12], RF demonstrated their ability to forecast day-ahead EV load charging blocks for households in an EVSE-centric fashion. As mentioned in previous paragraphs, the EVSE-centric approach can be difficult to implement as it requires large amounts of data and complex modelling. Thus, the use of RF for this kind of disaggregated approaches is adequate. [152] precisely illustrates the ability of RF to handle the EV load problem from both time and spatial dimensions. This article shows that RF can model both a single station as well as a group of stations considering spatial and temporal inputs. Single station models are more accurate as they have more consistent behaviours while the group of station models is slightly less accurate in terms of the mean absolute percentage error but brings a more holistic view to the problem. Other ensemble methods which stemmed from the same area of machine learning such as gradient boosting could also be considered [150], [156].

4.4.3. *Neural Networks*

ANNs were initially presented by Frank Rosenblatt in 1958 [157] in their most elementary form in the name of the perceptron. They were extended shortly afterwards to the Multi-Layer Perceptron (MLP). After being forgotten for a few decades, ANNs have experienced a rebound in interest from the end of the 80s in particular with the formulation of the backpropagation algorithm [158] and the breakthroughs in computer vision with Convolutional Neural Networks (CNN) [159].

In [41] an MLP with tilted loss function is used for probabilistic forecasting of EV load. It is compared with a kernel density estimator as well as quantile regression and it showed the best performance using the same inputs and outputs. It is quite common across various scientific fields that ANNs reach the highest performance on many problems compared to other machine learning or statistical methods. The main drawback is the lack of interpretability of such models which are highly complex [160], [161]. Elaborate ANNs architectures have been explored for modelling EV load. CNNs in [49] and Recurrent Neural Networks (RNNs) in [28]. [153] compares 12 different architectures including CNNs, RNNs. In this last article, the LSTM architecture showed the best performance on the dataset studied. From all these articles it is challenging to decide which ANN architecture is the overall best for EV load modelling. However, some clear conclusions can be made. Recurrent Neural Networks are particularly performant as they take into account historical EV load values. In the current operational context, this is information that is hard to obtain at fine time steps. Thus, until real-time communication channels are available, it is likely that the most useful ANN models will be CNNs or RNNs with larger timesteps.

5. Matching EV load models to open datasets: a preliminary study

Six datasets dealing with historical charging session information have been selected for their completeness and accessibility: Boulder [83], Palo Alto [84],

Dundee, [107], Perth, [108], Paris [96] and Domestic UK [126]. According to this research, none of these datasets were used in the EV load modelling literature so far. The purpose of this section is to identify the variables available and to enable a high-level understanding of charging behaviours. In addition, an association of the 6 datasets selected with the models reviewed in Section 4 is proposed.

5.1. Variables and data quality

The fields available in each of the six datasets selected are summarised in Figure 4. These six datasets provide us with session start and end times as well as the energy consumed. With the exception of the Domestic UK dataset, the station address (location), and the power level of the charging port are available in these datasets. In addition, the Palo Alto and Boulder datasets contain gasoline and greenhouse gases (GHG) savings as well as the charge duration which represents the amount of time the vehicle was plugged-in and actively charging. This is different to the park duration which also captures the time a vehicle was plugged-in and no longer charging which is a variable only given in the Palo Alto dataset. However, this park duration can be deduced from the session start and end times in the remaining datasets. Finally, for customer specific information, the Paris data provides a unique identifier per customer badge and the Palo Alto dataset gives the post code registered by the driver. Information regarding postcodes is interesting for models that include travel inputs such as the distance between the driver’s home and stations surrounding.

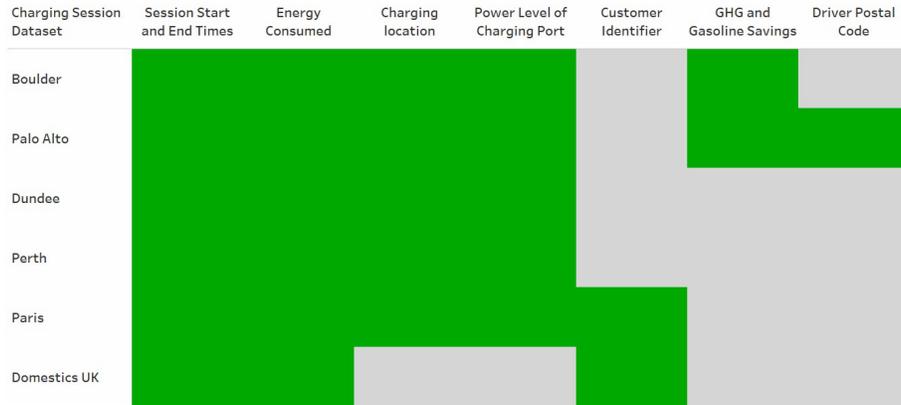


Fig. 3. Fields available in the six datasets selected. Fields available are in green and fields not available are in grey for each dataset considered.

Additionally, a data quality analysis was conducted on the six datasets. In [153], outliers were identified by using a set threshold from the variability between current and previous values. Instead, in this analysis, fixed boundaries were chosen and the following set thresholds were observed:

- Charge and/or Park Duration has to be positive and less than 24 hours
- Energy Consumption needs to be positive and less than 100 kWh

The first criterion is important as some datasets have some obvious errors in the end times column which are set in 1970. This might indicate a manipulation error from the customer which led to a computational mistake along the process of data collection. Also, as most charging sessions last for a few hours, charging sessions that lasted for more than a day were discarded.

Similarly, recorded energy consumption values for Perth [108] and Dundee [107] are highly variable, reaching anomalously negative and highly positive values indicative of potential errors. The 100 kWh upper bound was chosen as it is close to the highest capacity of the Tesla Model S which is the EV with the largest battery capacity amongst the most widespread models [162].

If a transaction does not fit these criteria, it is discarded for the following analysis. This preparation had very little impact on the Palo Alto dataset with

only 0.17% of transactions discarded, while the Boulder, Dundee, Perth, Paris and Domestic UK datasets have seen 8%, 11%, 4% , 14% and 7% of their data discarded respectively.

5.2. Exploratory analysis

Figure 4 shows the trend in the total number of transactions per day over each dataset specific time frame. Due to increasing EV uptake [1], an increase in EV charging sessions is expected as illustrated by Palo Alto and Perth. However, this is not the case for Boulder and Dundee. Instead, a decreasing number of charging sessions at the end of each time series can be observed. This could be due to external factors such as an increase in charging session prices. As for the Domestic UK, only one year of data is available which indicates that the plot shown describes the yearly cycle rather than the long-term trend. Similarly, the Paris data cannot be extrapolated as it only represents two months of data.

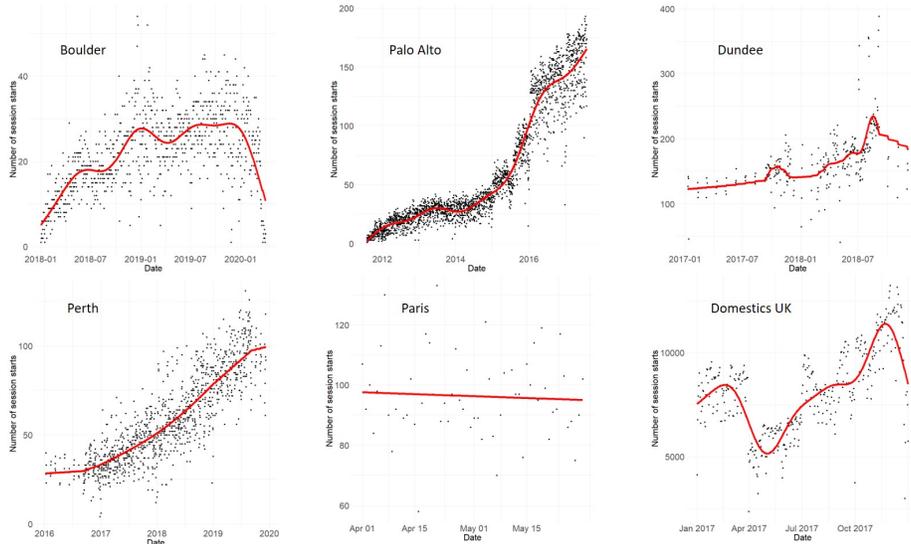


Fig. 4. The black dots represent the daily number of sessions while the red curve is the output of a generalised additive model used to smooth the values. Each model was produced using the *mgcv* package in R and its default settings [163].

Overall statistics of the six datasets in scope are provided in Table 5. The

dataset which covers the largest time frame is from the city of Palo Alto with 6 years, then comes Perth with 4 years, Boulder with a little over 2 years, Dundee with 2 years, Domestic UK with 1 years and Paris with 2 months. In terms of the transactions (or sessions), Domestic UK records the largest number of transactions. Moreover, Dundee and Boulder both cover 2 years of data but Dundee has close to three times more transactions. Naturally, there are consistently more transactions on weekdays than on weekends in total and on average across all datasets. Furthermore, the average park duration and charge duration varies significantly across the datasets. Indeed, while for Palo Alto the average park duration is around 2 hours and 40 minutes, in Perth it is closer to 1 hour and 15 minutes, so less than half of the time. Additionally, the average Park Duration for Domestic UK is greater than 9 hours which is expected as this dataset describes residential charge instead of public charge for the others. The Charge duration on the other hand is relatively close for Boulder and Palo Alto which are both American cities (located in Colorado and California). Finally, the average energy demand is consistently between 8 to 11 kWh across all datasets.

dataset		Boulder	Palo Alto	Dundee	Perth	Paris	Domestics UK
First Transaction Date		2018-01-01	2011-07-29	2017-01-09	2016-01-09	2017-04-01	2017-01-01
Last Transaction Date		2020-03-31	2017-07-31	2018-12-05	2019-12-08	2017-05-30	2017-12-31
Total Transactions	All	18,052	133,329	47,051	63,936	4,225	2,956,198
	Weekdays	13,487	101,486	34,434	46,607	4,907	2,208,695
	Weekends	4,565	31,843	12,617	17,329	1,555	747,503
Average Transactions	Weekdays	23	65	173	60	101	8,495
	Weekends	20	51	160	55	86	7,119
Park Duration (h)	Mean	-	2.7	2.29	1.24	1.7	9.16
	Standard Deviation	-	2.41	4.56	2.13	2.93	6.49
Charge Duration (h)	Mean	1.81	2.05	-	-	-	-
	Standard Deviation	1.34	1.39	-	-	-	-
Energy Demand (kWh)	Mean	8.42	8.18	9.16	11.01	8.51	8.88
	Standard Deviation	7.08	6.76	6.53	8.49	6.71	7.55

Table 5. Overall statistics and high-level information for the six datasets in scope.

5.3. Suggested matching of EV load models with the datasets considered

From this exploratory analysis, some suggestions can be given on how to match the EV load models reviewed in Section 4 with the six datasets presented above. The first criterion identified for this selection is whether the dataset describes public charging or residential charging. Only two of the models reviewed deal with residential charging and will thus be assigned to the Domestics UK dataset. They would benefit from the large number of records and the customer identifier provided in this dataset. Thus, it is a good setup for vehicle-centric approaches [44] and machine learning models [12].

Looking now at public charging sessions, Boulder and Palo Alto are the only two datasets which gathered greenhouse gases and gasoline savings. These fields are rather uncommon across charging session datasets and they can enable an environmental impact analysis of EVs. However, it would be limited to EV usage rather than a holistic environmental impact with lifecycle assessments [164].

Thus, no mention of this kind of analysis was found in the articles reviewed.

Nevertheless, Palo Alto also possesses the driver’s post code. With this knowledge, fine spatiotemporal processes can be derived as proposed in [32]. Additionally, the large amount of records available is suited to test the scalability of queuing models [16, 39] and spatiotemporal processes [9, 18] which require travel information. It also provides an ideal setup for deep learning models which require large training sets [33, 41, 49, 28, 153].

As discussed in Section 3, Boulder not only holds a charging session dataset but also a travel survey led in 2018 with a focus on EVs [88]. It is a rather qualitative survey and can be used in combination with the NHTS [43] to address the more specific behaviours inherent to the city of Boulder. With both travel and charging session data, this is also a favourable setup to apply spatiotemporal models [9, 18, 32]. Considering that this dataset is continuously updated and holds recent data, it would also be interesting to apply models which were precisely built for operational use such as [148] and [152] in the hope of taking consistent conclusions with real-world applications.

The Paris dataset holds customer identifier information which encourages vehicle-centric approaches. However, the small amount of data deters the use of models which leverage numerous parameters. Instead, statistical characterisation techniques with unimodal distributions could yield a sufficient approximation of the phenomenon as proposed in [13] and [14] along with LM [149, 150, 151]. The remaining statistical characterisation models (mixtures [24, 13, 124, 135, 134] and KDEs [136, 137, 138, 139, 140]) can capture diverse patterns and thus could be applied to medium-sized datasets. The Paris dataset could also be used to verify the consistency of simple queuing models as they usually struggle to find concrete applications [29, 143].

As Perth and Dundee are two neighbouring cities of the UK, it would be interesting to compare the difference in charging behaviours between them. Considering that they have the same fields, it would be interesting to independently compare their behaviours as despite their closeness, it is unlikely that there is a significant spatial impact between these for public charging. Thus, tempo-

ral processes produced in [15, 141, 37, 142] would be well suited. Additionally, thanks to the medium-size of both datasets, SVM models described in [33, 8] could also be a good option here.

Charging session dataset	Specificities	EV load models	Models references
Boulder [83]	Public Charging Sessions Medium-sized Dataset GHG and Gasoline Savings 2 years	Spatiotemporal RF Mixtures KDE	[32] [9] [18] [148] [152] [24] [135] [124] [134] [136] [137] [138] [139] [140]
Palo Alto [84]	Public Charging Sessions Large Dataset Driver's Post Code 6 years	Spatiotemporal Queues ANN	[32] [9] [18] [16] [39] [33] [41] [49] [28] [153]
Dundee [107] & Perth [108]	Public Charging Sessions Medium-sized Datasets 2 years (Dundee) and 4 years (Perth)	Temporal SVM Mixtures KDE	[15] [141] [37] [142] [33] [8] [24] [135] [124] [134] [136] [137], [138] [139] [140]
Paris [96]	Public Charging Sessions Small dataset Customer Identifier 2 months	Unimodal distributions LM Queues	[13] [14] [149] [150] [151] [29] [143]
Domestics UK [126]	Residential charging sessions Large dataset Customer Identifier 1 year	Temporal RF	[44] [12]

Table 6. Charging session datasets with associated EV load models suggested.

6. Discussion and future work

The purpose of this section is to highlight and discuss the current gaps and limitations from both open data and EV load models perspectives (Table 7).

	Section	Keywords	References
Data	6.1 Usage and privacy issues	Variables Accessibility Regulations	[57] [56] [41] [165]
	6.2 Other types of relevant data	General Electrical Load Non-Intrusive Load Monitoring Synthetic data	[53] [166] [167] [168] [135]
Models	6.3 Composite Approaches	Stochastic Processes Machine Learning	[169] [170] [171]
	6.4 Link with optimisation	Smart Charging Probabilistic Approach Reinforcement Learning	[6] [172] [173]

Table 7. Gaps and future work for EV load data and models. Sections 6.1 and 6.2 deal with data prospects and limitations while Sections 6.3 and 6.4 describe new ways of modelling and tie EV load models to optimisation of charging schedules.

6.1. Data usage and privacy issues

With this article, the community has clear visibility on a carefully selected list of open datasets useful for modelling EV load [55]. In most research papers, datasets obtained from system operators or other entities were not explicitly made available and no clear indication was given on how to retrieve them if it was possible. Our hope is to foster the practice of sharing supplementary materials with both the data used and code produced in order to encourage reproducible work in the field.

While the open data search provides visibility on charging session and traffic data, no repositories merging both was found. Thus, it is likely that the standard will remain to manipulate separate datasets for charging and traffic data as it is already the case in the literature [39, 48, 41] at least in the near future. And as such, different locations and different grains of data will still need to be leveraged in order to perform a complete data-driven spatiotemporal description of EV load.

Battery inputs are intrinsically complex to obtain. It would involve estab-

lishing an Internet of Things (IOT) between EVs, charging stations and controllers when considering a smart charging scenario [26, 27]. This type of work is currently in progress [41] and the community could benefit from new types of information for EV load models in the near future. However, so far, the articles which include these variables simulate them from prior statistical distributions.

Some of the datasets presented in Section 3 provide unique identifiers for vehicles and even driver’s registered post codes [84]. However, data regulations (e.g., GDPR in Europe [165]) may prevent spreading battery and travel inputs openly. Thus, more elaborate and complex models will be required in order to capture hidden information for disaggregated approaches.

The other variables of interest pinpointed in this review are easier to retrieve. For example, weather information can be obtained from the R package *riem* for a wide range of locations [56] or on the NOAA website [57]. If finer information is required, meteorological grid models can be used for that purpose. Economical and calendar variables can be tailored for each analysis depending on the grain chosen.

6.2. Other types of relevant data

The open data search presented in Section 3 mainly focuses on charging session and traffic data. However, it is also possible to consider general electrical load open data [53]. Indeed, if a region switches to EVs in a given time period, this change can reflect on the regional load curve. In this context, EV load would be a latent or hidden variable contributing to the general electrical load.

In addition, grid networks data [174] and big cities electrical load data [166] are becoming more and more available. Combining them with charging session and traffic data may lead to models which have a holistic and data-driven understanding of the reality.

To model the load of specific appliances with general electrical load data, Non-Intrusive Load Monitoring (NILM) methodologies have received a lot of interest in the related literature [175, 176]. The question addressed is whether it is possible to identify and characterise EV load within a general electrical

load curve [167], [168].

Synthetic data can also be used in order to produce EV load models. In most research papers, simulators rely heavily on assumptions derived from travel surveys and not so much on real charging session data [141, 15, 14]. Nevertheless, a data-driven simulator has been recently proposed in [135] which was trained on real-world charging sessions and thus can represent more accurately real world charging behaviours.

Finally, there are semi-open or closed data. Most of these closed datasets are related to residential load [91, 177] as it is less feasible to retrieve them without raising data privacy concerns.

6.3. Composite Approaches

From a methodology perspective, it is interesting to note that very few stochastic processes approaches used real data [16]. These models are usually theoretical and can be useful for mid-term or long-term scenarios but less relevant for short-term forecasting. Alternatively, the machine learning and statistical characterisation approaches presented were highly data-driven.

In the corpus of articles considered in this review, there exists no article that deals both with stochastic processes and machine learning algorithms in the context of EV load models. Thus, it would be interesting to compare them in terms of performance but also to assess what they can bring to each other in a composite model [171, 169, 170].

Furthermore, it was shown in many articles reviewed in Section 4 that the influx of vehicles at EVSEs is highly time dependent. Consequently, homogeneous poisson processes used in articles from the corpus will not be enough to capture the reality of drivers' behaviours [18]. More elaborate processes such as inhomogeneous poisson or self-exciting point processes [178] have to be considered to account for this time dependence. Using these stochastic processes hand in hand with machine learning algorithms will foster consistency, conciseness and performance of EV load models.

Finally, another gap brought to light in this review is the lack of work on

stacking models or bottom-up approaches [12] which are indeed more costly from a computational perspective but can bring a deeper understanding of EV load.

6.4. Link with optimisation

As mentioned in the introduction, EV load models are part of a two-step process. Firstly, behaviours relating to EV load demand must be understood and then current schedules optimised depending on the aim (e.g., load flattening or load balancing). The articles introducing methodologies for optimising charging schedules usually assume a clear knowledge of the future short-term demand. It is less common to see articles which account for the potential uncertainty of EV load models. This is also due to the fact that there has been less focus given to probabilistic EV load models which could yield confidence intervals for evaluating risks of surpassing the energy supply at a given time. Additionally, probabilistic forecasting proposes a more exhaustive representation of the demand as it does not solely focus on the mean demand.

Solutions which include both forecasting and optimisation aspects in the same model or process are required [6], [179]. Again, using the same data for this purpose is essential, as it enables the development of solutions by researchers specialised in different fields such as forecasting and optimisation. To unify both, methodologies can be also developed using reinforcement learning [173]. In addition, specific losses related to the exploitation of probabilistic forecasts in smart charging strategies could be relevant [172].

7. Conclusion

In this paper, the reader is provided with a comprehensive list of open data that can be used to model EV load. Additionally, an organised review of EV load models is presented. Finally, six datasets are explored to provide recommendations on how they can be matched to the EV load models reviewed.

The open data search focused on the top 14 countries of the EVI ranked by national EV market share. A total of 860+ open data repositories was covered which yielded more than 60 open datasets relevant for modelling EV load.

Across the literature, a wide spectrum of EV load models were reviewed. This includes statistical characterisation models from parametric (unimodal distributions and mixtures) to non-parametrical estimation (KDE). Furthermore, stochastic processes with purely temporal models, spatiotemporal models and queues were also included. Finally, machine learning models including LM, SVM, RF and ANN were reviewed.

From the open data research, six datasets which have not been previously studied in the literature were considered. Recommendations were provided on how the models reviewed could be matched to each dataset.

We hope that this article will encourage the use of the open datasets and models reviewed in order to foster reproducible work and breakthroughs in the field of EV load modelling.

Acknowledgements

The authors would like to acknowledge the contribution of Benjamin Mousseau and Emma Chieusse-G erard from EDF Energy who brought to our knowledge the existence of the Domestic UK dataset as well as Emma Yule from Edinburgh University who kindly proofread the article.

Appendix

EV load models

Study	Model(s)	Output variable(s)	Approach	Input dataset(s)
[14]	Gaussian Distribution	Power Demand	Vehicle-centric	NHTS daily trips [43]
[13]	Weibull and Lognormal Distributions	Charge Duration, Daily Charge Frequency, Energy Demand, Time to next charge	Vehicle-centric	Charging sessions from the Victorian EV Trial [180]

[24]	Beta Mixture	Number of plugged-in and plugged-out EVs, Charge Duration	Aggregated and vehicle-centric	Charging sessions from ElaadNL [70]
[134]	Gaussian Mixture	Clusters of EV user profiles	-	Charging sessions provided by the G4 cities of the Netherlands
[124]	Gaussian Mixture	Arrival Time, Charge Duration, Energy Demand	Vehicle-centric	Charging sessions from the ACN data [86]
[135]	Poisson Distribution and Gaussian Mixture	Arrival Time, Charge Duration, Energy Demand	Vehicle-centric	Charging sessions from ElaadNL [70]
[136]	GKDE	Daily Trip Distance and end time of last trip	Vehicle-centric	NHTS daily trips [43]
[137]	GKDE	Arrival Time, Charge Duration, Charge Capacity	Vehicle-centric	Charging sessions from an EV charging service company platform in Nanjing
[138]	GKDE and DKDE	Charge Duration and Energy Demand	Vehicle-centric	Charging sessions from UCLA campus
[139]	GKDE and DKDE	Charge Duration, Energy Demand	Vehicle-centric	Charging sessions from UCLA campus and from My Electric Avenue Project trial [181]

[140]	HKDE	Charge Duration and Energy Demand	Vehicle-centric	Charging sessions from UCLA campus and from My Electric Avenue Project trial [181]
[15]	Temporal stochastic process with scenarios	Power Demand	Vehicle-centric	EV charging load for lead-acid and lithium-ion batteries and UK National Travel Survey [182]
[141]	3-state temporal Markov Chain	Energy Demand	Vehicle-centric	-
[37]	ARIMA and rule-based probabilistic model	Power Demand (via the SoC after charge)	Vehicle-centric	PJM historical load data [183], EV drivers simulated data from the Power Systems Engineering Research Center [184]
[142]	ARIMA and rule-based probabilistic model	Power Demand (via the SoC after charge)	Vehicle-centric	PJM historical load data [183]
[44]	Monte-Carlo Simulation with copulas and scenarios	Energy Demand (via initial SoC)	Vehicle-centric	Transportation data from the Dutch Ministry of Transportation [185]
[9]	Monte-Carlo Simulation with origin destination analysis	Power Demand	Vehicle-centric	EU merge EV database ⁴

⁴The link provided in the article is no longer available online

[18]	Spatiotemporal stochastic process with uniform and poisson distributions on different schedules and state transitions	Power Demand	Vehicle-centric and aggregated	BMW Mini E (PHEV) battery characteristics from the smart grid integration project (Shanghai expo garden)
[32]	Spatiotemporal monte-carlo simulation based on EV trial data	Power Demand	Vehicle-centric	Charging sessions and EV journeys from the Switch EV trial
[29]	$M/M/n_{max}$ queue	Power Demand	Vehicle-centric and EVSE-centric	Tesla Roadster EV characteristics [186]
[38]	$M/M/n_{max}$ queue	Power Demand	Vehicle-centric and EVSE-centric	EV drivers simulated data from the Power Systems Engineering Research Center [184]
[143]	$M/M/n_{max}$ queue	Power Demand	Vehicle-centric and EVSE-centric	Traffic fluid model [187]
[16]	$M_t/GI_t/\infty$ queue	Power Demand	Vehicle-centric and EVSE-centric	NHTS [43], charging sessions from the UC Davis PH&EV center [188]

[39]	BCMP network	Power Demand	Vehicle-centric and EVSE-centric	NHTS [43], New York State Transportation Federation Traffic Data Viewer [189] and Ontario Electricity prices ⁵
[31]	Decision trees/tables, SVM, ANN	Power Demand	Aggregated	Residential charging sessions provided by ECOtality [147], public EVSE data pilot EV project in France
[148]	MPSF, kNN, SVM and RF	Energy Demand	EVSE-centric	Charging sessions from UCLA campus data
[149]	Mean estimation and Linear Regression	Energy Demand (Charging Profile)	Vehicle-centric	Charging sessions from UCLA campus data
[150]	Linear Regression, Gradient Boosting, ANN	Departure time	Vehicle-centric	EV charging records from SAP Walldorf
[151]	Linear Regression	Power Demand (Charging Speed)	Aggregated	Charging sessions from EVNET and NUON energy providers in Amsterdam
[33]	MPSF, SVM and RF	Energy Demand (Charging Profile)	Vehicle-centric and EVSE-centric	Charging sessions from UCLA campus data
[8]	SVM and Monte-Carlo	Power Demand	Vehicle-centric	Beijing Olympic Games EV charging station

⁵The link provided in the article is no longer available online

[12]	RF	Power Demand	Vehicle-centric and EVSE-centric	Charging sessions from Pecan Street [177]
[152]	RF, SVM and Decision tree	Energy Demand (Charging capacity)	EVSE-centric	Charging sessions in Shenzhen
[41]	Quantile regression, MLP and KDE	Trip Distance and Parking Duration	Vehicle-centric	Parking events derived from trip data from MiD2008 [190]
[49]	Lion algorithm with Niche Immunity for a CNN	Power Demand	Aggregated	Charging sessions in one charging station in Beijing
[28]	ANN, RNN, Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU)	Energy Demand	Aggregated	Charging sessions in Shenzhen
[153]	ANN, RNN, GRU, Stacked auto-encoders and LSTM	Power Demand	Aggregated	Charging sessions in Shenzhen

Table 8. Exhaustive list of models presented in Section 4 with the input dataset used (if applicable), the approach taken (Aggregated, Vehicle-centric or EVSE-centric) and the output variable(s) modelled. For the input datasets, the data repositories or data reports are provided as references when they were clearly made available by the authors of the paper.

References

- [1] International Energy Agency, Global EV Outlook 2019: Scaling-up the transition to electric mobility, OECD, 2019. doi:10.1787/35fb60bd-en.
- [2] H. Kvisle, The Norwegian Charging Station Database for Electromobility (NOBIL), World Electric Vehicle Journal 5 (3) (2012) 702–707. doi:10.3390/wevj5030702.

- [3] International Energy Agency, Electric vehicle stock in the EV30@30 scenario, library Catalog: www.iea.org.
URL <https://www.iea.org/data-and-statistics/charts/electric-vehicle-stock-in-the-ev3030-scenario-2018-2030>
[accessed 2020-05-03]
- [4] Clean Energy Ministerial, Electric Vehicles Initiative.
URL <http://www.cleanenergyministerial.org/initiative-clean-energy-ministerial/electric-vehicles-initiative> [accessed 2020-05-04]
- [5] M. D. Galus, M. G. Vayá, T. Krause, G. Andersson, The role of electric vehicles in smart grids: The role of electric vehicles in smart grids, *Wiley Interdisciplinary Reviews: Energy and Environment* 2 (4) (2013) 384–400. doi:10.1002/wene.56.
- [6] K. L. Lopez, C. Gagne, M.-A. Gardner, Demand-Side Management Using Deep Learning for Smart Charging of Electric Vehicles, *IEEE Transactions on Smart Grid* 10 (3) (2019) 2683–2691. doi:10.1109/TSG.2018.2808247.
- [7] R. Fachrizal, M. Shepero, D. van der Meer, J. Munkhammar, J. Widén, Smart charging of electric vehicles considering photovoltaic power production and electricity consumption: A review, *eTransportation* 4 (2020) 100056. doi:10.1016/j.etrans.2020.100056.
- [8] E. S. Xydas, C. E. Marmaras, L. M. Cipcigan, A. S. Hassan, N. Jenkins, Forecasting Electric Vehicle charging demand using Support Vector Machines, in: 2013 48th International Universities' Power Engineering Conference (UPEC), IEEE, Dublin, 2013, pp. 1–6. doi:10.1109/UPEC.2013.6714942.
- [9] Y. Mu, J. Wu, N. Jenkins, H. Jia, C. Wang, A Spatial–Temporal model for grid impact analysis of plug-in electric vehicles, *Applied Energy* 114 (2014) 456–465. doi:10.1016/j.apenergy.2013.10.006.

- [10] S. F. Tie, C. W. Tan, A review of energy sources and energy management system in electric vehicles, *Renewable and Sustainable Energy Reviews* 20 (2013) 82–102. doi:10.1016/j.rser.2012.11.077.
- [11] K. Yamashita, J. Li, P. Zhang, C.-C. Liu, Analysis and control of major blackout events, in: *2009 IEEE/PES Power Systems Conference and Exposition, 2009*, pp. 1–4. doi:10.1109/PSCE.2009.4840091.
- [12] A. Gerossier, R. Girard, G. Kariniotakis, Modeling and Forecasting Electric Vehicle Consumption Profiles, *Energies* 12 (7) (2019) 1341. doi:10.3390/en12071341.
- [13] Y. B. Khoo, C.-H. Wang, P. Paevere, A. Higgins, Statistical modeling of Electric Vehicle electricity consumption in the Victorian EV Trial, Australia, *Transportation Research Part D: Transport and Environment* 32 (2014) 263–277. doi:10.1016/j.trd.2014.08.017.
- [14] K. Sun, M. R. Sarker, M. A. Ortega-Vazquez, Statistical characterization of electric vehicle charging in different locations of the grid, in: *2015 IEEE Power & Energy Society General Meeting, IEEE, Denver, CO, USA, 2015*, pp. 1–5. doi:10.1109/PESGM.2015.7285794.
- [15] K. Qian, C. Zhou, M. Allan, Y. Yuan, Modeling of Load Demand Due to EV Battery Charging in Distribution Systems, *IEEE Transactions on Power Systems* 26 (2) (2011) 802–810. doi:10.1109/TPWRS.2010.2057456.
- [16] M. Alizadeh, A. Scaglione, J. Davies, K. S. Kurani, A Scalable Stochastic Model for the Electricity Demand of Electric and Plug-In Hybrid Vehicles, *IEEE Transactions on Smart Grid* 5 (2) (2014) 848–860. doi:10.1109/TSG.2013.2275988.
- [17] N. Z. Xu, C. Y. Chung, Challenges in Future Competition of Electric Vehicle Charging Management and Solutions, *IEEE Transactions on Smart Grid* 6 (3) (2015) 1323–1331. doi:10.1109/TSG.2014.2373401.

- [18] H. Jiang, H. Ren, C. Sun, D. Watts, The temporal-spatial stochastic model of plug-in hybrid electric vehicles, in: 2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), IEEE, Torino, Italy, 2017, pp. 1–6. doi:10.1109/ISGTEurope.2017.8260233.
- [19] D. B. Richardson, Electric vehicles and the electric grid: A review of modeling approaches, Impacts, and renewable energy integration, *Renewable and Sustainable Energy Reviews* 19 (2013) 247–254. doi:10.1016/j.rser.2012.11.042.
- [20] G. T. Heydt, The impact of electric vehicle deployment on load management strategies, *IEEE transactions on power apparatus and systems* (1983) 1253–1259.
- [21] N. Sadeghianpourhamami, N. Refa, M. Strobbe, C. Develder, Quantitative analysis of electric vehicle flexibility: A data-driven approach, *International Journal of Electrical Power & Energy Systems* 95 (2018) 451–462. doi:10.1016/j.ijepes.2017.09.007.
- [22] E. C. Kara, J. S. Macdonald, D. Black, M. Bérge, G. Hug, S. Kiliccote, Estimating the benefits of electric vehicle smart charging at non-residential locations: A data-driven approach, *Applied Energy* 155 (2015) 515–525. doi:10.1016/j.apenergy.2015.05.072.
- [23] W. Kempton, S. E. Letendre, Electric vehicles as a new power source for electric utilities, *Transportation Research Part D: Transport and Environment* 2 (3) (1997) 157–175. doi:10.1016/S1361-9209(97)00001-1.
- [24] M. G. Flammini, G. Pretticco, A. Julea, G. Fulli, A. Mazza, G. Chicco, Statistical characterisation of the real transaction data gathered from electric vehicle charging stations, *Electric Power Systems Research* (2019) 15.
- [25] G. Saldaña, J. I. San Martin, I. Zamora, F. J. Asensio, O. Oñederra, Electric Vehicle into the Grid: Charging Methodologies Aimed at Provid-

- ing Ancillary Services Considering Battery Degradation, *Energies* 12 (12) (2019) 2443. doi:10.3390/en12122443.
- [26] B. Chokkalingam, S. Padmanaban, P. Siano, R. Krishnamoorthy, R. Selvaraj, Real-Time Forecasting of EV Charging Station Scheduling for Smart Energy Systems, *Energies* 10 (3) (2017) 377. doi:10.3390/en10030377.
- [27] M. A. Ahmed, M. R. El-Sharkawy, Y.-C. Kim, Remote Monitoring of Electric Vehicle Charging Stations in Smart Campus Parking Lot, *Journal of Modern Power Systems and Clean Energy* 8 (1) (2020) 124–132. doi:10.35833/MPCE.2018.000502.
- [28] J. Zhu, Z. Yang, Y. Guo, J. Zhang, H. Yang, Short-Term Load Forecasting for Electric Vehicle Charging Stations Based on Deep Learning Approaches, *Applied Sciences* 9 (9) (2019) 1723. doi:10.3390/app9091723.
- [29] R. Garcia-Valle, J. G. Vlachogiannis, Letter to the Editor: Electric Vehicle Demand Model for Load Flow Studies, *Electric Power Components and Systems* 37 (5) (2009) 577–582. doi:10.1080/15325000802599411.
- [30] P. Olivella-Rosell, R. Villafafila-Robles, A. Sumper, J. Bergas-Jané, Probabilistic Agent-Based Model of Electric Vehicle Charging Demand to Analyse the Impact on Distribution Networks, *Energies* 8 (5) (2015) 4160–4187. doi:10.3390/en8054160.
- [31] S. Xydas, C. Marmaras, L. Cipcigan, A. Hassan, N. Jenkins, Electric Vehicle Load Forecasting using Data Mining Methods, in: *Hybrid and Electric Vehicles Conference 2013 (HEVC 2013)*, Institution of Engineering and Technology, London, UK, 2013, pp. 10.1–10.1. doi:10.1049/cp.2013.1914.
- [32] M. Neaimeh, R. Wardle, A. M. Jenkins, J. Yi, G. Hill, P. F. Lyons, Y. Hübner, P. T. Blythe, P. C. Taylor, A probabilistic approach to

- combining smart meter and electric vehicle charging data to investigate distribution network impacts, *Applied Energy* 157 (2015) 688–698. doi:10.1016/j.apenergy.2015.01.144.
- [33] M. Majidpour, C. Qiu, P. Chu, H. R. Pota, R. Gadh, Forecasting the EV charging load based on customer profile or station measurement?, *Applied Energy* 163 (2016) 134–141. doi:10.1016/j.apenergy.2015.10.184.
- [34] J. Li, J. Klee Barillas, C. Guenther, M. A. Danzer, A comparative study of state of charge estimation algorithms for LiFePO₄ batteries used in electric vehicles, *Journal of Power Sources* 230 (2013) 244–250. doi:10.1016/j.jpowsour.2012.12.057.
- [35] J. Topić, B. Škugor, J. Deur, Neural Network-Based Modeling of Electric Vehicle Energy Demand and All Electric Range, *Energies* 12 (7) (2019) 1396. doi:10.3390/en12071396.
- [36] M. H. Amini, M. P. Moghaddam, Probabilistic modelling of electric vehicles’ parking lots charging demand, in: 2013 21st Iranian Conference on Electrical Engineering (ICEE), 2013, pp. 1–4, iSSN: 2164-7054. doi:10.1109/IranianCEE.2013.6599716.
- [37] M. Amini, O. Karabasoglu, M. D. Ilic, K. G. Boroojeni, S. S. Iyengar, ARIMA-based demand forecasting method considering probabilistic model of electric vehicles’ parking lots, in: 2015 IEEE Power & Energy Society General Meeting, IEEE, Denver, CO, USA, 2015, pp. 1–5. doi:10.1109/PESGM.2015.7286050.
- [38] G. Li, X.-P. Zhang, Modeling of Plug-in Hybrid Electric Vehicle Charging Demand in Probabilistic Power Flow Calculations, *IEEE Transactions on Smart Grid* 3 (1) (2012) 492–499. doi:10.1109/TSG.2011.2172643.
- [39] H. Liang, I. Sharma, W. Zhuang, K. Bhattacharya, Plug-in electric vehicle charging demand estimation based on queueing network analysis, in: 2014

IEEE PES General Meeting | Conference & Exposition, IEEE, National Harbor, MD, USA, 2014, pp. 1–5. doi:10.1109/PESGM.2014.6939530.

- [40] P. Paevere, A. Higgins, Z. Ren, M. Horn, G. Grozev, C. McNamara, Spatio-temporal modelling of electric vehicle charging demand and impacts on peak household electrical load, *Sustainability Science* 9 (1) (2014) 61–76. doi:10.1007/s11625-013-0235-3.
- [41] J. Huber, D. Dann, C. Weinhardt, Probabilistic forecasts of time and energy flexibility in battery electric vehicle charging, *Applied Energy* 262 (2020) 114525. doi:10.1016/j.apenergy.2020.114525.
- [42] H. Wang, Energy consumption of electric vehicles based on real-world driving patterns: A case study of Beijing, *Applied Energy* (2015) 10.
- [43] National Household Travel Survey.
URL <https://nhts.ornl.gov/> [accessed 2020-04-28]
- [44] A. Lojowska, D. Kurowicka, G. Papaefthymiou, L. van der Sluis, Stochastic Modeling of Power Demand Due to EVs Using Copula, *IEEE Transactions on Power Systems* 27 (4) (2012) 1960–1968. doi:10.1109/TPWRS.2012.2192139.
- [45] T. Hong, S. Fan, Probabilistic electric load forecasting: A tutorial review, *International Journal of Forecasting* 32 (3) (2016) 914–938. doi:10.1016/j.ijforecast.2015.11.011.
- [46] P. Gaillard, Y. Goude, R. Nedellec, Additive models and robust aggregation for GEFCom2014 probabilistic electric load and electricity price forecasting, *International Journal of Forecasting* 32 (3) (2016) 1038–1050. doi:10.1016/j.ijforecast.2015.12.001.
- [47] E. Xydas, C. Marmaras, L. M. Cipcigan, N. Jenkins, S. Carroll, M. Barker, A data-driven approach for characterising the charging demand of electric vehicles: A UK case study, *Applied Energy* 162 (2016) 763–771. doi:10.1016/j.apenergy.2015.10.151.

- [48] M. B. Arias, S. Bae, Electric vehicle charging demand forecasting model based on big data technologies, *Applied Energy* 183 (2016) 327–339. doi:10.1016/j.apenergy.2016.08.080.
- [49] Y. Li, Y. Huang, M. Zhang, Short-Term Load Forecasting for Electric Vehicle Charging Station Based on Niche Immunity Lion Algorithm and Convolutional Neural Network, *Energies* 11 (5) (2018) 1253. doi:10.3390/en11051253.
- [50] E. Valsera-Naranjo, A. Sumper, R. Villafila-Robles, D. Martínez-Vicente, Probabilistic Method to Assess the Impact of Charging of Electric Vehicles on Distribution Grids, *Energies* 5 (5) (2012) 1503–1531. doi:10.3390/en5051503.
- [51] S. Huang, D. Infield, The potential of domestic electric vehicles to contribute to Power System Operation through vehicle to grid technology, in: 2009 44th International Universities Power Engineering Conference (UPEC), 2009, pp. 1–5.
- [52] J. C. Kelly, J. S. MacDonald, G. A. Keoleian, Time-dependent plug-in hybrid electric vehicle charging based on national driving patterns and demographics, *Applied Energy* 94 (2012) 395–405. doi:10.1016/j.apenergy.2012.02.001.
- [53] Y. Wang, Q. Chen, T. Hong, C. Kang, Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges, *IEEE Transactions on Smart Grid* 10 (3) (2019) 3125–3148. doi:10.1109/TSG.2018.2818167.
- [54] Open Data Inception, library Catalog: data.opendatasoft.com.
URL <https://data.opendatasoft.com/explore/dataset/open-data-sources@public/table/> [accessed 2020-04-22]
- [55] Y. Amara-Ouali, EV load Open Data, library Catalog: github.com.

- URL <https://github.com/yvenn-amara/EV-Load-Open-Data> [accessed 2020-05-04]
- [56] M. Salmon, `riem`: Accesses Weather Data from the Iowa Environment Mesonet (Sep. 2016).
URL <https://CRAN.R-project.org/package=riem> [accessed 2020-04-23]
- [57] NOAA, National Oceanic and Atmospheric Administration, Data Access | National Centers for Environmental Information (NCEI) formerly known as National Climatic Data Center (NCDC).
URL <https://www.ncdc.noaa.gov/data-access> [accessed 2020-07-15]
- [58] International Energy Agency, Electric car market shares in Electric Vehicle Initiative countries, library Catalog: www.iea.org (Nov. 2019).
URL <https://www.iea.org/data-and-statistics/charts/electric-car-market-shares-in-electric-vehicle-initiative-evi-countries> [accessed 2020-04-22]
- [59] Electric vehicle market statistics 2020 - How many electric cars in UK ?
URL <https://www.nextgreencar.com/electric-cars/statistics/> [accessed 2020-05-26]
- [60] NOBIL, NOBIL Database.
URL <https://info.nobil.no/eng> [accessed 2020-04-28]
- [61] DataNorge, Traffic Volumes.
URL <https://data.norge.no/data/statens-vegvesen/trafikkmengde-%C3%A5rsd%C3%B8gntrafikk-%C3%A5dt-p%C3%A5-esri-geodatabase-format-20130806> [accessed 2020-04-28]
- [62] StatisticsNorway, Road traffic volumes, by main type of vehicle, type of fuel and age of vehicle 2005 - 2019, library Catalog: www.ssb.no.
URL <http://www.ssb.no/en/statbanken/statbank/table/12578/> [accessed 2020-04-28]

- [63] StatisticsNorway, Registered Vehicles, library Catalog: www.ssb.no.
URL <https://www.ssb.no/en/transport-og-reiseliv/statistikker/bilreg/aar/2020-03-31> [accessed 2020-04-28]
- [64] Reykjavík, Reykjavík Transport, library Catalog: opingogn.is.
URL <https://opingogn.is/dataset/reykjavik-samgongur> [accessed 2020-04-28]
- [65] StatisticsIceland, Registered Motor Vehicles, library Catalog: www.statice.is.
URL <https://www.statice.is/statistics/business-sectors/transport/vehicles/> [accessed 2020-04-28]
- [66] ElbilSverige, Charging point locations in nordic countries, library Catalog: www.elbilsverige.se.
URL <http://www.elbilsverige.se/hitta-laddplats/> [accessed 2020-04-28]
- [67] StatisticsSweden, Registered vehicles, library Catalog: www.statistikdatabasen.scb.se.
URL http://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START__TK__TK1001__TK1001A/PersBilarDrivMedel/ [accessed 2020-04-28]
- [68] RotterdamOpenData, Charging point locations and usage in Rotterdam, library Catalog: rotterdamopendata.nl.
URL <http://rotterdamopendata.nl/dataset/oplaadpalen-voor-elektrische-auto-s> [accessed 2020-04-28]
- [69] EindhovenOpenData, Charging Point locations in Eindhoven, library Catalog: data.eindhoven.nl.
URL <https://data.eindhoven.nl/explore/dataset/oplaadpalen/information/?disjunctive.adres> [accessed 2020-04-28]
- [70] Elaad NL, Data Analytics.

- URL <https://www.elaad.nl/research/data-analytics/> [accessed 2020-04-28]
- [71] CBS, Centraal Bureau Voor De Statistiek: Mobility by person, transport mode and region.
URL https://opendata.cbs.nl/statline/portal.html?_la=nl&_catalog=CBS&tableId=84709NED&_theme=400 [accessed 2020-04-28]
- [72] CBS, Centraal Bureau Voor De Statistiek: Onderzoek Onderweg in Nederland - ODiN 2018, medium: application/pdf,.sav,.por,.dta,.ods type: dataset (2019). doi:10.17026/DANS-XN4-Q9KS.
URL <https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:156525> [accessed 2020-04-28]
- [73] CBS, Centraal Bureau Voor De Statistiek: Motor vehicles; type, age class.
URL https://opendata.cbs.nl/statline/portal.html?_la=en&_catalog=CBS&tableId=82044ENG&_theme=1089 [accessed 2020-04-28]
- [74] StatisticsFinland, Average commuting distance by Year, library Catalog: pxnet2.stat.fi.
URL https://pxnet2.stat.fi:443/PXWebPXWeb/pxweb/en/StatFin/StatFin__lii__tyoma/statfin_tyoma_pxt_001.px/ [accessed 2020-04-28]
- [75] StatisticsFinland, Number of vehicles registered by Year, Vehicle class and Information, library Catalog: pxnet2.stat.fi.
URL https://pxnet2.stat.fi:443/PXWebPXWeb/pxweb/en/StatFin/StatFin__lii__mkan/statfin_mkan_pxt_11ib.px/ [accessed 2020-04-28]
- [76] Data Gov Hong-Kong, Details of HK Electric EV charging stations.
URL https://data.gov.hk/en-data/dataset/hkelectric-tnd_cs_ci-hkelectric-ev-location [accessed 2020-04-28]

- [77] StatsGov, China Statistical Yearbook 2018.
URL <http://www.stats.gov.cn/tjsj/ndsj/2018/indexeh.htm> [accessed 2020-04-28]
- [78] Data Gov Hong-Kong, Statistics on Passenger, Visitor and Vehicular Traffic.
URL <https://data.gov.hk/en-data/dataset/hk-immd-set5-statistics-passenger-visitor-vehicular-traffic> [accessed 2020-04-28]
- [79] Data Gov Hong-Kong, Traffic Surveys in Kowloon East.
URL <https://data.gov.hk/en-data/dataset/hk-devb-traf-traf> [accessed 2020-04-28]
- [80] DadosGOV, Charging Stations E-mobility in Lisbon, library Catalog: dados.gov.pt.
URL <https://dados.gov.pt/datasets/postos-de-carregamento-mobi-e/> [accessed 2020-04-28]
- [81] PORDATA, Traffic Statistics in Portugal.
URL <https://www.pordata.pt/Subtema/Portugal/Rodovi%3%a1rio-405> [accessed 2020-04-28]
- [82] USDepartmentofEnergy, Alternative Fuels Data Center.
URL https://afdc.energy.gov/data_download [accessed 2020-04-28]
- [83] L. Makram, Electric Vehicle Charging Stations: Energy Consumption & Savings.
URL <https://bouldercolorado.gov/open-data/electric-vehicle-charging-stations/> [accessed 2020-04-28]
- [84] CityofPaloAlto, Electric Vehicle Charging Station Usage (July 2011 - Dec 2017) · Open Data, library Catalog: data.cityofpaloalto.org.
URL <http://data.cityofpaloalto.org/dataviews/244892/>

electric-vehicle-charging-station-usage-july-2011-dec-2017/
[accessed 2020-04-28]

- [85] CityofEvanston, City-owned Electric Vehicle Charging Station Usage January 2016 to August 2017 | Open Data, library Catalog: data.cityofevanston.org.

URL <https://data.cityofevanston.org/dataset/City-owned-Electric-Vehicle-Charging-Station-Usage/nx7w-jb8v/data> [accessed 2020-04-28]

- [86] ACN-Data – A Public EV Charging Dataset.

URL <https://ev.caltech.edu/dataset> [accessed 2020-06-16]

- [87] CityofHouston, Traffic Counts in the City of Houston, library Catalog: cohgis-mycity.opendata.arcgis.com.

URL https://cohgis-mycity.opendata.arcgis.com/datasets/38ecbcfd8d084e8d80a14c2c9aed7074_0/data [accessed 2020-04-28]

- [88] B. Skerpan, Resident Mobility Survey.

URL <https://bouldercolorado.gov/open-data/resident-mobility-survey/> [accessed 2020-04-28]

- [89] NewYorkState, Electric Vehicles per County | Open Data NY, library Catalog: data.ny.gov.

URL <https://data.ny.gov/Transportation/Electric-Vehicles-per-County/uu25-czyc> [accessed 2020-05-04]

- [90] CityofEdmonton, Public Charging Stations for Electric Vehicles.

URL <https://data.edmonton.ca/Environmental-Services/Public-Charging-Stations-for-Electric-Vehicles/xzhy-xe8z>
[accessed 2020-04-28]

- [91] OpenGovOntario, Electric vehicle home charging program applicant data, library Catalog: data.ontario.ca.

URL <https://data.ontario.ca/dataset/>

- `electric-vehicle-home-charging-program-applicant-data` [accessed 2020-04-28]
- [92] TrafficDataProgram, Annual Traffic Volumes 2004-2010 in British Columbia, library Catalog: `catalogue.data.gov.bc.ca`.
URL <https://catalogue.data.gov.bc.ca/dataset/annual-traffic-volumes-2004-2010> [accessed 2020-04-28]
- [93] CanadaGov, Canadian travel survey, travel in Canada, by travel duration, library Catalog: `open.canada.ca`.
URL <https://open.canada.ca/data/en/dataset/19f7dc90-a073-4000-8180-a4bc0aa290f8> [accessed 2020-04-28]
- [94] G. Milot, Electric and Hybrid Vehicles - CKAN, library Catalog: `donnees.ville.montreal.qc.ca`.
URL <http://donnees.ville.montreal.qc.ca/dataset/vehicules-electriques-et-hybrides> [accessed 2020-04-28]
- [95] IRVE, Charging point locations for Electric Vehicles, library Catalog: `opendata.reseaux-energies.fr`.
URL <https://opendata.reseaux-energies.fr/explore/dataset/bornes-irve/information/?disjunctive.region> [accessed 2020-05-04]
- [96] DataGouv, Charging sessions Apr-May 2017 in Paris, library Catalog: `www.data.gouv.fr`.
URL <https://www.data.gouv.fr/fr/datasets/belib-reseau-parisien-de-bornes-de-recharges-accelerees-22-kw-ac-dc-pour-vehicules-el> [accessed 2020-04-28]
- [97] SAPLabs, Electric Vehicle charging transactions of SAP Labs France company fleet.
URL <https://opendata.reseaux-energies.fr/explore/dataset/conso-ve-sap0/information/>

?disjunctive.stop_inactivitystatus&dataChart=
eyJxdWVyaWVzIjpbeyJjaGFydHMiO1t7ImFsaWduTW9udGgiOnRydWUsInR5cGUiOiJjb2x1bW4iLCJmdW5jIjpo
[accessed 2020-11-01]

[98] ParisData, Belib' availability in real-time, library Catalog: open-
data.paris.fr.

URL https://opendata.paris.fr/explore/dataset/prises_electriques_vehicules_tps_reel/information/?disjunctive.static_name&disjunctive.plugs_outletmodel&disjunctive.plugs_status&disjunctive.static_brand&disjunctive.static_owner&disjunctive.geolocation_postalcode&disjunctive.geolocation_city [accessed 2020-04-28]

[99] ParisData, Historical Traffic Counts in Paris, library Catalog: open-
data.paris.fr.

URL <https://opendata.paris.fr/explore/dataset/comptages-routiers-permanents-historique/information/> [ac-
cessed 2020-04-28]

[100] RennesMétropole, Real-time traffic in Rennes, library Catalog:
data.rennesmetropole.fr.

URL <https://data.rennesmetropole.fr/explore/dataset/etat-du-traffic-en-temps-reel/information/> [accessed 2020-04-
28]

[101] INSEE, Travel Survey 2015: home to work, library Catalog:
data.grandparissud.fr.

URL <https://data.grandparissud.fr/explore/dataset/mobilites-professionnelles-en-2015-deplacements-domicile-lieu-de-travail0/information/> [accessed 2020-04-28]

[102] INSEE, Registered Vehicles in France from 1990 to 2018.

URL <https://www.insee.fr/fr/statistiques/2045167> [accessed
2020-04-28]

- [103] NewPlymouthDistricCouncil, Traffic Counts in New Plymouth, library Catalog: data-npdc.opendata.arcgis.com.
URL http://data-npdc.opendata.arcgis.com/datasets/7358e6a46ddf49e19c1867469b742ddc_8/data [accessed 2020-04-28]
- [104] StatsNZ, Motor Vehicles Registered.
URL http://archive.stats.govt.nz/infoshare/default.aspx?RedirectReason=session_expired [accessed 2020-04-28]
- [105] OfficeforLowEmissionVehicles, National Charge Point Registry, library Catalog: data.gov.uk (Feb. 2012).
URL <https://data.gov.uk/dataset/1ce239a6-d720-4305-ab52-17793fedfac3/national-charge-point-registry> [accessed 2020-04-28]
- [106] SunderlandCityCouncil, Charge your car.
URL <https://www.sunderland.gov.uk/chargeyourcar> [accessed 2020-04-28]
- [107] TransportTeam, Electric Vehicle Charging Sessions Dundee, library Catalog: data.dundee.gov.uk.
URL <https://data.dundee.gov.uk/dataset/ev-charging-data> [accessed 2020-04-28]
- [108] OpenDataTeam, Electric Vehicle Charging Station Usage in Perth and Kinross, library Catalog: data.pkc.gov.uk.
URL <https://data.pkc.gov.uk/dataset/ev-charging-data> [accessed 2020-04-28]
- [109] A. Radford, Birmingham and West Midlands real-time traffic data, library Catalog: data.birmingham.gov.uk.
URL <https://data.birmingham.gov.uk/dataset/wm-utmc> [accessed 2020-04-28]
- [110] DepartmentforTransport, Traffic Flows, Borough, library Catalog: data.london.gov.uk.

- URL <https://data.london.gov.uk/dataset/traffic-flows-borough>
[accessed 2020-04-28]
- [111] UKDataService, National Travel Survey 2002-2017.
URL <https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=5340> [accessed 2020-04-28]
- [112] O. Smith, How Many Left?
URL <https://www.howmanyleft.co.uk/?utf8=%E2%9C%93&q=&commit=Search> [accessed 2020-05-04]
- [113] OffenesDatenPortal, Charging point locations in Moers, library Catalog:
www.offenesdatenportal.de.
URL <https://www.offenesdatenportal.de/dataset/autoladestationen-in-moers> [accessed 2020-04-28]
- [114] OpenDataBonn, Charging point locations and usage in real-time in Bonn.
URL <https://opendata.bonn.de/dataset/standorte-elektrotankstellen-e-lades%C3%A4hlen-realtime-belegung> [accessed 2020-04-28]
- [115] OpenDataBonn, Real-time traffic in Bonn.
URL <https://opendata.bonn.de/dataset/stra%C3%9Fenverkehrslage-realtime> [accessed 2020-04-28]
- [116] OpenDataBayern, Traffic Volumes in Bayern.
URL <https://opendata.bayern.de/detailansicht/datensatz/baysis-verkehrsmengen-wfs?0> [accessed 2020-04-28]
- [117] FDZ Ruhr am RWI, Research Data Center.
URL <https://www.rwi-essen.de/forschung-und-beratung/fdz-ruhr/datenzugang/> [accessed 2020-10-15]
- [118] L. Ecke, German Mobility Panel - Startseite, archive Location: KIT
Publisher: Lisa Ecke (Aug. 2020).

- URL <https://mobilitaetspanel.ifv.kit.edu/english/index.php>
[accessed 2020-10-15]
- [119] Bochum, Registered Vehicles in Bochum, library Catalog:
www.bochum.de.
URL [https://www.bochum.de//Open-Data/Datensaetze/
Transport-und-Verkehr](https://www.bochum.de//Open-Data/Datensaetze/Transport-und-Verkehr) [accessed 2020-04-28]
- [120] MLIT, Travel Survey in Japan from April 2019.
URL <https://www.mlit.go.jp/k-toukei/jidousya.html> [accessed
2020-04-28]
- [121] e Stat, Registered Vehicles in Japan, library Catalog: www.e-stat.go.jp.
URL [https://www.e-stat.go.jp/stat-search?page=1&toukei=
00600700&bunya_1=10](https://www.e-stat.go.jp/stat-search?page=1&toukei=00600700&bunya_1=10) [accessed 2020-04-28]
- [122] Passenger car sales by country, around the world, library Catalog:
www.theglobaleconomy.com.
URL [https://www.theglobaleconomy.com/rankings/passenger_
cars_sales/](https://www.theglobaleconomy.com/rankings/passenger_cars_sales/) [accessed 2020-04-22]
- [123] NationaleDatabankWegverkeersgegevens, Real-time and historical traffic
in the Netherlands.
URL [https://www.ndw.nu/pagina/nl/103/datalevering/120/open_
data/](https://www.ndw.nu/pagina/nl/103/datalevering/120/open_data/) [accessed 2020-04-28]
- [124] Z. J. Lee, T. Li, S. H. Low, ACN-Data: Analysis and Applications of an
Open EV Charging Dataset, in: Proceedings of the Tenth ACM Interna-
tional Conference on Future Energy Systems, ACM, Phoenix AZ USA,
2019, pp. 139–149. doi:10.1145/3307772.3328313.
- [125] Cerema, Road traffic open data, last Modified: 2020-05-27 (May 2020).
URL [http://trafic-routier.data.cerema.fr/
acces-aux-donnees-de-traffic-routier-open-data-r25.html](http://trafic-routier.data.cerema.fr/acces-aux-donnees-de-traffic-routier-open-data-r25.html) [ac-
cessed 2020-11-01]

- [126] DepartmentforTransport, Domestic Chargepoint Analysis 2017, library Catalog: data.gov.uk (Dec. 2018).
URL <https://data.gov.uk/dataset/5438d88d-695b-4381-a5f2-6ea03bf3dcf0/electric-chargepoint-analysis-2017-domestics> [accessed 2020-06-09]
- [127] DepartmentforTransport, Rapid Chargepoint Analysis 2017, library Catalog: www.gov.uk (2018).
URL <https://www.gov.uk/government/statistics/electric-chargepoint-analysis-2017-local-authority-rapids> [accessed 2020-06-09]
- [128] DepartmentforTransport, Fast Chargepoint Analysis 2017, library Catalog: www.gov.uk (2018).
URL <https://www.gov.uk/government/statistics/electric-chargepoint-analysis-2017-public-sector-fasts> [accessed 2020-06-09]
- [129] LondonDatastore, Low Carbon London Electric Vehicle Load Profiles, library Catalog: data.london.gov.uk.
URL <https://data.london.gov.uk/dataset/low-carbon-london-electric-vehicle-load-profiles> [accessed 2020-06-15]
- [130] Laddinfra, Statistics on EV public charge, library Catalog: www.laddinfra.se.
URL <https://www.laddinfra.se/statistik> [accessed 2020-04-28]
- [131] Tampere, Tampere real-time traffic, library Catalog: data.tampere.fi.
URL <https://data.tampere.fi/data/fi/dataset/tampereen-kaupungin-liikennetiedoterajapinta> [accessed 2020-04-28]
- [132] P. Breidenbach, L. Eilers, RWI-GEO-GRID: Socio-economic data on grid

level, *Jahrbücher für Nationalökonomie und Statistik* 238 (6) (2018) 609–616, publisher: De Gruyter Oldenbourg Section: *Jahrbücher für Nationalökonomie und Statistik*. doi:10.1515/jbnst-2017-0171.

- [133] L. Yuanyuan, China installed more than 1000 EV charging stations per day in 2019, library Catalog: www.renewableenergyworld.com Section: Infrastructure (Jan. 2020).
URL <https://www.renewableenergyworld.com/2020/01/13/china-installed-more-than-1000-ev-charging-stations-per-day-in-2019/>
[accessed 2020-05-04]
- [134] J. R. Helmus, M. H. Lees, R. van den Hoed, A data driven typology of electric vehicle user types and charging sessions, *Transportation Research Part C: Emerging Technologies* 115 (2020) 102637. doi:10.1016/j.trc.2020.102637.
- [135] M. Lahariya, D. Benoit, C. Develder, Poster: Defining a synthetic data generator for realistic electric vehicle charging sessions, in: *Proceedings of the Eleventh ACM International Conference on Future Energy Systems*, 2020, p. 2. doi:<https://doi.org/10.1145/3396851.3403509>.
- [136] M. Liang, W. Li, J. Yu, L. Shi, Kernel-based electric vehicle charging load modeling with improved latin hypercube sampling, in: *2015 IEEE Power & Energy Society General Meeting*, IEEE, Denver, CO, USA, 2015, pp. 1–5. doi:10.1109/PESGM.2015.7285758.
- [137] L. Chen, X. Huang, H. Zhang, Modeling the Charging Behaviors for Electric Vehicles Based on Ternary Symmetric Kernel Density Estimation, *Energies* 13 (7) (2020) 1551. doi:10.3390/en13071551.
- [138] B. Khaki, Y.-W. Chung, C. Chu, R. Gadh, Nonparametric User Behavior Prediction for Distributed EV Charging Scheduling, in: *2018 IEEE Power & Energy Society General Meeting (PESGM)*, IEEE, Portland, OR, 2018, pp. 1–5. doi:10.1109/PESGM.2018.8585744.

- [139] B. Khaki, Y.-W. Chung, C. Chu, R. Gadh, Probabilistic Electric Vehicle Load Management in Distribution Grids, in: 2019 IEEE Transportation Electrification Conference and Expo (ITEC), IEEE, Detroit, MI, USA, 2019, pp. 1–6. doi:10.1109/ITEC.2019.8790535.
- [140] Y.-W. Chung, B. Khaki, C. Chu, R. Gadh, Electric Vehicle User Behavior Prediction Using Hybrid Kernel Density Estimator, in: 2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), IEEE, Boise, ID, 2018, pp. 1–6. doi:10.1109/PMAPS.2018.8440360.
- [141] P. Sokorai, A. Fleischhacker, G. Lettner, H. Auer, Stochastic Modeling of the Charging Behavior of Electromobility, *World Electric Vehicle Journal* 9 (3) (2018) 44. doi:10.3390/wevj9030044.
- [142] M. H. Amini, A. Kargarian, O. Karabasoglu, ARIMA-based decoupled time series forecasting of electric vehicle charging demand for stochastic power system operation, *Electric Power Systems Research* 140 (2016) 378–390. doi:10.1016/j.epsr.2016.06.003.
- [143] S. Bae, A. Kwasinski, Spatial and Temporal Model of Electric Vehicle Charging Demand, *IEEE Transactions on Smart Grid* 3 (1) (2012) 394–403. doi:10.1109/TSG.2011.2159278.
- [144] O. Häggström, O. H. G. M, *Finite Markov Chains and Algorithmic Applications*, Cambridge University Press, 2002, google-Books-ID: hpLxIJ9LwRgC.
- [145] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, G. M. Ljung, *Time series analysis: forecasting and control, fifth edition Edition*, Wiley series in probability and statistics, John Wiley & Sons, Inc, Hoboken, New Jersey, 2016.
- [146] F. Baskett, K. M. Chandy, R. R. Muntz, F. G. Palacios, Open, Closed, and

Mixed Networks of Queues with Different Classes of Customers, *Journal of the ACM (JACM)* 22 (2) (1975) 248–260. doi:10.1145/321879.321887.

- [147] The EV Project - Overview (Nov. 2010).
URL <https://web.archive.org/web/20101128153912/http://theevproject.com/overview.php> [accessed 2020-05-02]
- [148] M. Majidpour, C. Qiu, P. Chu, R. Gadh, H. R. Pota, A novel forecasting algorithm for electric vehicle charging stations, in: 2014 International Conference on Connected Vehicles and Expo (ICCVE), IEEE, Vienna, Austria, 2014, pp. 1035–1040. doi:10.1109/ICCVE.2014.7297504.
- [149] Y. Xiong, C.-c. Chu, R. Gadh, B. Wang, Distributed optimal vehicle grid integration strategy with user behavior prediction, in: 2017 IEEE Power & Energy Society General Meeting, IEEE, Chicago, IL, 2017, pp. 1–5. doi:10.1109/PESGM.2017.8274327.
- [150] O. Frendo, N. Gaertner, H. Stuckenschmidt, Improving Smart Charging Prioritization by Predicting Electric Vehicle Departure Time, *IEEE Transactions on Intelligent Transportation Systems* (2020) 1–8doi:10.1109/TITS.2020.2988648.
- [151] J. Mies, J. Helmus, R. van den Hoed, Estimating the Charging Profile of Individual Charge Sessions of Electric Vehicles in The Netherlands, *World Electric Vehicle Journal* 9 (2) (2018) 17. doi:10.3390/wevj9020017.
- [152] Y. Lu, Y. Li, D. Xie, E. Wei, X. Bao, H. Chen, X. Zhong, The Application of Improved Random Forest Algorithm on the Prediction of Electric Vehicle Charging Load, *Energies* 11 (11) (2018) 3207. doi:10.3390/en11113207.
- [153] J. Zhu, Z. Yang, M. Mourshed, Y. Guo, Y. Zhou, Y. Chang, Y. Wei, S. Feng, Electric Vehicle Charging Load Forecasting: A Comparative Study of Deep Learning Approaches, *Energies* 12 (14) (2019) 2692. doi:10.3390/en12142692.

- [154] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer New York, New York, NY, 1995. doi:10.1007/978-1-4757-2440-0.
- [155] L. Breiman, Random Forest, *Machine Learning* 45 (1) (2001) 5–32. doi:10.1023/A:1010933404324.
- [156] L. Buzna, P. D. Falco, S. Khormali, D. Proto, M. Straka, Electric vehicle load forecasting: a comparison between time series and machine learning approaches, in: *Proceeding of the International Conference on Energy Transition in the Mediterranean Area*, 2019, p. 5.
- [157] F. Rosenblatt, The perceptron: A probabilistic model for information storage and organization in the brain., *Psychological Review* 65 (6) (1958) 386–408. doi:10.1037/h0042519.
- [158] D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning Internal Representations by Error Propagation, Technical Report ICS-8506, Institute for Cognitive Science (1985).
- [159] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel, Backpropagation Applied to Handwritten Zip Code Recognition, *Neural Computation* 1 (4) (1989) 541–551, publisher: MIT Press. doi:10.1162/neco.1989.1.4.541.
- [160] S. Chakraborty, R. Tomsett, R. Raghavendra, D. Harborne, M. Alzantot, F. Cerutti, M. Srivastava, A. Preece, S. Julier, R. M. Rao, T. D. Kelley, D. Braines, M. Sensoy, C. J. Willis, P. Gurram, Interpretability of deep learning models: A survey of results, in: *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI)*, IEEE, San Francisco, CA, 2017, pp. 1–6. doi:10.1109/UIC-ATC.2017.8397411.

- [161] Q. Zhang, S.-C. Zhu, Visual Interpretability for Deep Learning: a Survey, arXiv:1802.00614 [cs]ArXiv: 1802.00614.
- [162] O. F. Vynakov, E. V. Savolova, A. I. Skrynnyk, Modern Electric Cars of Tesla Motors Company, Automation of technological and business processes 8 (2). doi:10.15673/atbp.v8i2.162.
- [163] S. Wood, mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation (Nov. 2019).
URL <https://CRAN.R-project.org/package=mgcv> [accessed 2020-06-30]
- [164] D. Burchart-Korol, S. Jursova, P. Folega, J. Korol, P. Pustejovska, A. Blaut, Environmental life cycle assessment of electric vehicles in Poland and the Czech Republic, Journal of Cleaner Production 202 (2018) 476–487. doi:10.1016/j.jclepro.2018.08.145.
- [165] General Data Protection Regulation (GDPR) – Official Legal Text, library Catalog: gdpr-info.eu.
URL <https://gdpr-info.eu/> [accessed 2020-04-30]
- [166] eCO2mix, Electrical consumption in cities open data.
URL <https://www.data.gouv.fr/fr/datasets/donnees-eco2mix-consommation-des-metropoles-temps-reel/>
[accessed 2020-11-06]
- [167] P. Zhang, C. Zhou, B. G. Stewart, D. M. Hepburn, W. Zhou, J. Yu, An Improved Non-Intrusive Load Monitoring Method for Recognition of Electric Vehicle Battery Charging Load, Energy Procedia 12 (2011) 104–112. doi:10.1016/j.egypro.2011.10.015.
- [168] S. Wang, L. Du, J. Ye, D. Zhao, A Deep Generative Model for Non-intrusive Identification of EV Charging Profiles, IEEE Transactions on Smart Grid (2020) 1–1doi:10.1109/TSG.2020.2998080.

- [169] N. Du, H. Dai, R. Trivedi, U. Upadhyay, M. Gomez-Rodriguez, L. Song, Recurrent Marked Temporal Point Processes: Embedding Event History to Vector, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16, ACM Press, San Francisco, California, USA, 2016, pp. 1555–1564. doi:10.1145/2939672.2939875.
- [170] S. Xiao, J. Yan, X. Yang, H. Zha, S. M. Chu, Modeling the Intensity Function of Point Process via Recurrent Neural Networks, in: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, 2017, p. 7.
- [171] N. Fallah, H. Gu, K. Mohammad, S. A. Seyyedsalehi, K. Nourijelyani, M. R. Eshraghian, Nonlinear Poisson regression using neural networks: a simulation study, Neural Computing and Applications 18 (8) (2009) 939–943. doi:10.1007/s00521-009-0277-8.
- [172] C. Capezza, B. Palumbo, Y. Goude, S. N. Wood, M. Fasiolo, Additive stacking for disaggregate electricity demand forecasting, arXiv:2005.10092 [stat]ArXiv: 2005.10092.
- [173] N. Sadeghianpourhamami, J. Deleu, C. Develder, Achieving Scalable Model-Free Demand Response in Charging an Electric Vehicle Fleet with Reinforcement Learning, in: Proceedings of the Ninth International Conference on Future Energy Systems - e-Energy '18, ACM Press, Karlsruhe, Germany, 2018, pp. 411–413. doi:10.1145/3208903.3212042.
- [174] Enedis, Grid map open data.
 URL <https://data.enedis.fr/pages/accueil/?id=dataviz-cartographie-des-reseaux> [accessed 2020-11-06]
- [175] S. Makonin, F. Popowich, Nonintrusive load monitoring (NILM) performance evaluation, Energy Efficiency 8 (4) (2015) 809–814. doi:10.1007/s12053-014-9306-2.

- [176] H. Salem, M. Sayed-Mouchaweh, A Semi-supervised and Online Learning Approach for Non-Intrusive Load Monitoring, in: U. Brefeld, E. Fromont, A. Hotho, A. Knobbe, M. Maathuis, C. Robardet (Eds.), Machine Learning and Knowledge Discovery in Databases, Lecture Notes in Computer Science, Springer International Publishing, Cham, 2020, pp. 585–601. doi:10.1007/978-3-030-46133-1_35.
- [177] Pecan Street Data, library Catalog: www.pecanstreet.org.
URL <https://www.pecanstreet.org/dataport/about/> [accessed 2020-06-16]
- [178] P. J. Laub, T. Taimre, P. K. Pollett, Hawkes Processes, arXiv:1507.02822 [math, q-fin, stat]ArXiv: 1507.02822.
- [179] O. Frendo, J. Graf, N. Gaertner, H. Stuckenschmidt, Data-driven smart charging for heterogeneous electric vehicle fleets, Energy and AI 1 (2020) 100007. doi:10.1016/j.egyai.2020.100007.
- [180] Department of Transport, Creating a market: Victorian Electric Vehicle Trial Mid-Term Report, Tech. rep., State Government of Victoria (2013).
- [181] Engineering Products and Training Courses | EA Technology.
URL <https://eatechnology.com/> [accessed 2020-11-23]
- [182] GOV.UK, National Travel Survey.
URL <https://www.gov.uk/government/collections/national-travel-survey-statistics> [accessed 2020-11-23]
- [183] PJM, Markets & Operations, District of Columbia.
URL <https://www.pjm.com/markets-and-operations> [accessed 2020-11-23]
- [184] A. D. Dominguez-Garcia, G. T. Heydt, S. Suryanarayanan, Implications of the Smart Grid Initiative on Distribution Engineering, Tech. rep., Power Systems Engineering Research Center (2011).

- [185] Ministerie van Verkeer en Waterstaat, Rijkswaterstaat, Dienst Verkeer en Scheepvaart, Mobiliteitsonderzoek Nederland 2008 (2008). doi:10.17026/DANS-232-ZCQR.
URL <https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:34107> [accessed 2020-11-23]
- [186] Tesla Motors, Electric Cars, Solar & Clean Energy.
URL <https://www.tesla.com/> [accessed 2020-11-23]
- [187] K. Leung, W. Massey, W. Whitt, Traffic models for wireless communication networks, IEEE Journal on Selected Areas in Communications 12 (8) (1994) 1353–1364. doi:10.1109/49.329340.
- [188] UCDavis, Plug-In Hybrid & Electric Vehicles Research Center, institute of Transportation Studies.
URL <https://phev.ucdavis.edu/> [accessed 2020-11-23]
- [189] NewYorkState, Traffic Data Viewer.
URL <https://www.dot.ny.gov/tdv> [accessed 2020-11-23]
- [190] J. Huber, J. Höffer, J. Thumm, C. Weinhardt, Parking events derived from trip data from MiD2008, type: dataset (2019). doi:10.5445/IR/1000098024.
URL <https://publikationen.bibliothek.kit.edu/1000098024> [accessed 2020-11-23]