



Electric vehicle aggregation model: A probabilistic approach in representing flexibility

Mark Jenkins^{a,*}, Ivana Kockar^b

^a Institute of Energy and Environment, University of Strathclyde, Glasgow, Scotland, United Kingdom

^b Institute of Energy and Environment, University of Strathclyde, Glasgow, Scotland, United Kingdom

ARTICLE INFO

Keywords:

Aggregation
Behavioural modelling
Electric vehicles
Markov-chain
Probabilistic models

ABSTRACT

With the increased roll-out of electric vehicles (EV), the need, but also opportunity, for leveraging their flexibility in the context of grid support, ancillary services and local market energy trading. However, the uncertainty and variability in driving patterns and resultant charging profiles pose substantial risk for aggregators. Given this context, the paper demonstrates a method for producing a stochastic, socio-economically-differentiated aggregation model that determines the flexibility space of a realistic and diverse EV fleet. A probabilistic Monte Carlo Markov Chain model is developed that allows for the overlay and comparison of different technical, spatial and social-economic behavioural factors through clustering and correlation analyses. In turn, the model enables a statistically significant analysis of the 'Energy Space' available that captures the inherent risk and uncertainty when leveraging EV flexibility.

1. Introduction

Governments worldwide are setting ambitious targets to decarbonise the transport sector to mitigate the impacts of climate change and reduce air pollution [1, 2]. Electric Vehicles (EVs) are cited as key contributors in achieving these targets, resulting in their rapid uptake. In the UK during 2020, EVs sales were up by 185.9% versus 2019, at 6.6% of new-market cars and as of 2040, all sales of petrol and diesel vehicles will be prohibited [2, 3]. It is expected that, with such substantial change in transportation's energy carrier to electricity, the impacts on the electrical power system will be significant. Concerns have been raised by both network operators and regulators by the rapid uptake in EVs and the subsequent increase in domestic home charging on existing/weaker distribution networks [4].

With that said, this trend of EV uptake is indicative of general uptake in distributed energy resources (DERs) and the move towards a more decentralised power system [5]. As such, there is growing need, but also opportunity, for flexibility in the context of grid support and ancillary services. With the innate capabilities of EVs with vehicle-to-grid (V2G) technology, their increased levels of penetration can prove to be a potential solution to integrate higher shares of renewables and DERs [6]. The introduction of a single EV to current ancillary service markets is doubtful and will most likely have their participation managed and

interfaced through such actors as aggregators [7]. In the UK, support for the role of an independent aggregator has been growing, with Ofgem stating that their participation in market-based flexibility services can bring benefit to consumers [8]. Typically, aggregators create value by capitalising on economies of scale whilst identifying and managing risk. For EVs, the uncertainty and variability in driving patterns and resultant charging profiles (CPs) is a substantial risk, and as such, there has been significant strides towards methods of analysing and quantifying future CPs [9, 10].

Given that the prevalence of EVs across the population is still limited, common approaches to quantify CPs has broadly fallen into two methods – specific public trials and emulation/simulation. Public trials, typically part of a publicly funded EV project such as "My Electric Avenue and Network Revolution", are used to monitor individuals with the aim of drawing insights into driving and charging patterns of EV users. These insights can translate directly into the impacts on the power system and provide the opportunity to derive probability distribution functions (PDFs) of charging start time and energy requirement [11]. As for simulated models, the aim remains similar except that the individual behaviours of EV peers are emulated in a model typically derived from surveys such as household travel (HTSs) [10]. Both methods have distinct advantages, trial data captures the unique driving behaviours of EV users whereas models allow for the focusing of arbitrary yet

* Corresponding author.

E-mail address: mark.jenkins@strath.ac.uk (M. Jenkins).

<https://doi.org/10.1016/j.epsr.2022.108484>

Received 3 October 2021; Received in revised form 17 April 2022; Accepted 2 July 2022

Available online 2 August 2022

0378-7796/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

customisable scenarios and individual characteristics. Some studies have grounded simulation models from EV trials, however, such approaches are doubtful as trial data lacks population size and heterogeneity [12]. In fact, it can be argued that using data from EV trials poses a risk of capturing and emphasising EV charging characteristics and behaviours from specific, self-selecting, societal, spatial and technical factors. For example, the trial participants in “SwitchEV trial” were volunteers who paid a monthly participant fee and had access to specific charging infrastructure (such as off-street home charging). Studies based on such trial data may mislead the consequent impacts on electrical network infrastructure, or at least the efficacy of using such data may quickly become outdated and obsolete.

As such, it can be argued that the most suitable method currently to estimating EV CPs is with models/simulations, until the prevalence of EVs within a natural population subset can be observed. [13] shows a Markov-chain model that derives EV driving patterns based on the persons household presence and movements. In [14], PDFs are derived from the American national HTS that are used to model key EV statistics, that is, initial charging time, required energy and net power consumption of charging to assess aggregated charging impacts. In addition, both [15] and [16] used similar approaches using HTSs for Germany and Netherlands, respectively.

Furthermore, in [17] and [18] models were developed that capture the stochastic nature of the driving behaviour that include a model verification against the original HTS data –important when considering uncertainty in EV behaviour. A key advantage in using validated models grounded on real HTS data is that it allows for methods to accurately increase population size - needed for statistically significant analysis. With that said, the referenced literature emphasises the technical and spatial factors of CPs and lack the socio-economic. As stated in [19], age, income and educational background vary significantly in working and living patterns and thus will lead to diverse travel modes and energy consumptions. For example, in [20] the most statistically significant influencing factors on residential charging behaviour was economic status and occupation. Heterogeneity modelled within a population is important, and studies that withdraw impacts of demographics on the travel patterns risk the potential for hidden errors within their simulations.

Lastly, only a few studies have translated the uncertainty in EV CPs that include the aforementioned population heterogeneity into statistical analysis that highlights the energy ‘space’. It is this information that aggregators, the actor most likely to facilitate EVs within existing and new ancillary service markets, will require. Such information will enable aggregators to quantify not only the magnitude of available flexibility but also encapsulate the inherent risk and uncertainty when leveraging a technically, spatially and behaviourally diverse population.

As such, this paper utilizes a Monte Carlo Markov-chain (MCMC) model that can generate stochastic, socio- economically-differentiated EV driving patterns and charging profiles. This MCMC model is both derived and verified against a UK based HTS, namely the UK Time Use Survey (TUS), which in turn is used to synthetically and accurately increase the population size of the dataset. The use of the UK TUS allows for the overlay and comparison of different behavioural, spatial and social-economic factors through clustering and correlation analyses. In these analyses, conclusions are drawn that highlight the potential for overlap in certain characteristics, that if disregarded, can lead to misclassification of EV users and errors within driving models. Furthermore, with this model, different factors can be synthetically increased within a population to represent a variety of district level CPs for both network and population specific areas. The model outputs are then presented in a manner that is targeted for the use of district level aggregators that gives them an insight to the large space that is available for flexibility which is both high level but granular in its overall population make up, i.e., differentiation in aspects such as occupation and income as well as technical infrastructure etc. This can then inform aggregators of how to harness this flexibility (for services such as smart

charging, local energy trading etc.) whilst accounting for the population heterogeneity induced risk and uncertainty.

The paper is structured as follows. Section 2 describes the methodology to translate the raw TUS data into the simulation of EV mobility. Section 3 presents the construction and verification of the MCMC model used to increase the statistical significance of the dataset through synthetic population growth. Section 4 presents the methodology for the clustering and correction analysis to the highlight most prominent characteristics for EV driving behaviour. Section 5 summarises the most relevant findings and provides insights into the ‘energy space’ available for leveraging EV flexibility when accounting for risk and uncertainty. Lastly, Section 6 details the conclusions and future work.

2. Vehicle based activities from time of use survey data

This section describes the method where the raw TUS survey data is extracted and translated from “the movement of the participants to the movement of their vehicles”. An overview of significant factors for an EV model is provided followed by a discussion the preprocessing steps taken for the raw data as well as some assumptions made.

2.1. Influential factors in EV models

To fully assess the impact of EVs on the electrical power system, one must have a fundamental understanding of the CPs and its influencing factors. These factors can be divided into three main aspects – behavioural, technical and spatial. As stated previously, a lot of work has highlighted the significance and importance of the latter two and lacks emphasis on behaviours aspects. However, all three are not mutually exclusive and thus a more holistic view is required to accurately anticipate EV CPs. Understanding the behaviour of the individual can directly translate to the driving behaviour and thus the energy required for driving an EV. Many aspects of an individual’s behaviour affect driving, as shown in [20] where the most statistically significant influencing factors on residential charging behaviour was economic status and occupation. That is to say, driving characteristics such as the number and length of trips, arrival and departure times and parking locations are all heavily influenced by the individual’s social- economic status as well as household type and location. Generally speaking, the driving has the largest influence on the energy ‘required’ and less on the actual charging/connection decisions. These decisions are of course also essential in understanding an individual’s CP and involve decisions based on the State of Charge (SoC) of the EV as well as the time, frequency and availability of connections.

2.2. UK time use survey data

Given the high level of interdependencies on all three of the above factors, it can be concluded that to accurately anticipate EV CPs, one must consider a more holistic overview. Such an overview requires an analysis of an appropriate dataset that allows for a statistical analysis of technical, spatial and behavioural differentiation within a model. As concluded from the review of work done so far, simulated models, such as those derived from HTS datasets, are currently most appropriate. More generally, models that use HTS are referred to as direct use of observed activity-travel schedules (DUOATS) and, as the name suggests, details the movement of people during an allotted timeframe with information gathered at set intervals. The detailing of such a schedule inherently captures the realised and complex decisions that an individual would make, including implications this has on their mobility – i.e., provide spatial information. Furthermore, behavioural differentiation can then be extracted from the resultant mobility behaviour from the HTSs social-economic questions/entries, as well allowing for sensitivity analysis of varying technical infrastructure.

In this paper, the analysis and resultant model stems from the 2014–2015 UK TUS data [21]. Generally, TUSs measure the amount of

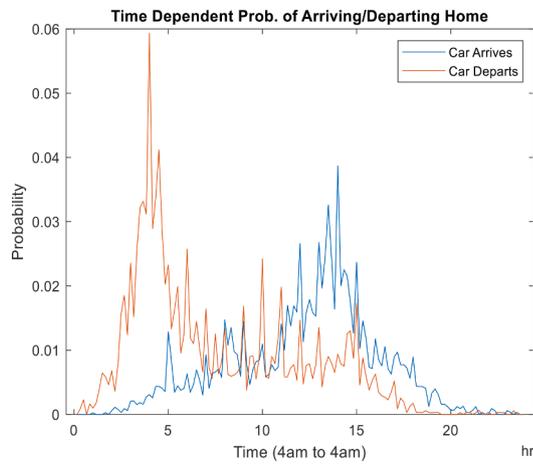


Fig. 1. Figure details the independent probability that the vehicle will either arrive or depart the home.

time people spend doing various activities, such as paid work, household and family care, personal care, voluntary work, social life, travel, and leisure activities. The survey consists of daily, 10 min interval diaries, with information detailing the respondent, sociodemographic background i.e., household income, age, occupations and employment status. In total, there are 4741 unique households within the survey with 11,421 individuals/diaries each with one weekday and weekend entry. Each day is recorded from 04:00 am to 03:50 am of the following day, where each respondent details their primary and secondary activity and location of activity. The main activity monitored for this paper is the individual use of their vehicle, time spent driving and time spent at home after returning with their vehicle.

2.2.1. Data format and extraction for vehicle movement

Before overlaying additional factors, the movement of people must first be translated to the movement of cars. Although the time and location of the individual is easily extracted, some assumptions of how this infers car use must be made. Initially, all entries that did not have an associated household car were removed. Secondly, entries where the survey participant did not at any point drive the vehicle as the main passenger were removed. Given that this may result in inaccuracies with a specific vehicles SOC, it was deemed necessary in order to appropriately keep track of fully observable car movements with the 24 h period. For example, using such diary entries could result in an ambiguous timeframe of car use if the respondent used an alternative mode of transport back home after travelling within the household car as a 'passenger'.

With the resultant diaries, of which there were 3124, more specific car movements can now be extracted. As an example, for one 10 min interval, the diaries can reflect the purpose and duration of the journey which can be supplemented with additional codes for method of transport – which in this case was 34 labelled as 'Travelling by passenger car as the driver'. Therefore, it is possible to extract a continuous catalogue of vehicle locations throughout the day. For this paper, home domestic charging will be the main focus as it is considered to be more reliable, cheaper and convenient whilst also being a vital component towards enabling consumer participating in the future flexibility services. It is expected that domestic home charging, which accounts for around 80% of all EV charging today, will continue to be a prominent part of future charging ecosystems [22].

These locational states will be discussed further in the MCMC discussion. It is possible to create different travel activities to monitor, for example, *to work*, *to school* and *to the shops*, however given the variance in trip length, as well as trip probability and start-time variance, it was decided these activities were less statistically valuable.

As well as being able to translate the diary respondents' moments to

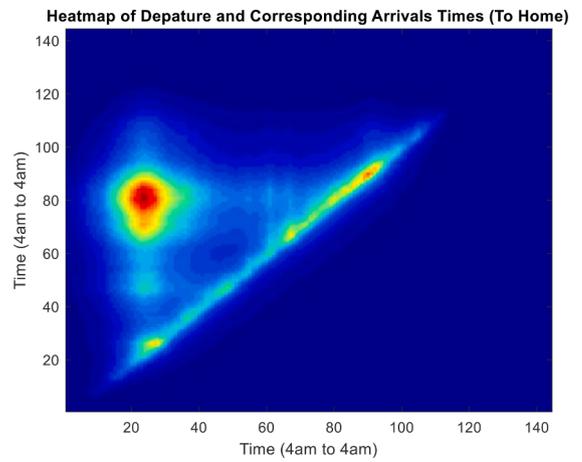


Fig. 2. This Figure couples round-trips to home with the higher density of red colour showing a larger share of the population.

car movements, population statistics on arrival and departure times, independent of the vehicles prior locations can be drawn, as shown in Figs. 1 and 2. The distinction here is that the former shows the decoupled probability distribution functions of departure time/arrival times to the home, that is, given the population within the TUS data, the likelihood that the vehicle was away from home prior and will now arrive back home and vice versa for departs. In this case, the probability value is obtained by dividing the associated activities by the total number of diaries considered for each 10 min time interval. However, for Fig. 2, the graph couples the round-trip journey of the vehicle leaving and departing the home (removing potential intermediate journeys before returning home). What this demonstrates is that, the majority of round-trip journeys are either relatively short (as seen by the prominent diagonal line) or longer that linked to the participant's round-trip journey to and from work and home.

3. Methodology – Monte Carlo Markov chain model

3.1. MCMC model description

As this point, due to the limited size of the extracted TUS data, sampling this data would result in less statistically significant metrics, particularly once the data is overlaid and subdivided according to additional influential factors such as occupation and income. Following the methodology shown in [23], vehicle movement from the TUS data can essentially be converted to a series of state transitions through the use of a Time-inhomogeneous Markov Chain model. Here, time-inhomogeneous refers to the time-varying nature of the state transition of a vehicle's movement. With knowledge of the vehicles previous state, one can determine the probability distribution of the current state which can be followed through the entire time horizon. These Markov properties are generally contained within a transition matrix P_t , where at time t holds (for that specific time interval) a matrix containing the state transition probabilities. More specifically, given a set rows and columns within Markov chain X the element number i, j of P_t is given by:

$$(P_t)_{ij} = P(X_{t+1} = j | X_t = i) \quad (1)$$

To satisfy the Markov property, the resultant summation of each row within P_t must equal 1 for any natural number t and states $i, j \in S$. For example, given a two state Markov chain at time step 2 with the previous state = 2, the first-row vector should be:

$$P_2 = P(X_3 = 1 | X_2 = 1) + P(X_3 = 2 | X_2 = 1) = 1 \quad (2)$$

As discussed in Section 2 the observable states of vehicle movement

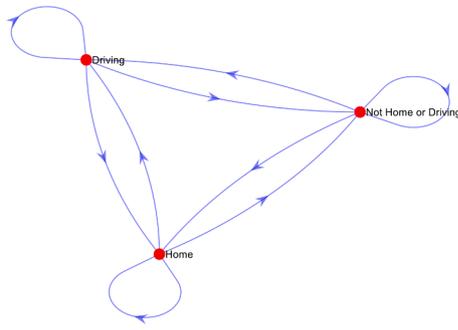


Fig. 3. Markov Chain diagram detailing the states surround vehicle use within one 10 min interval.

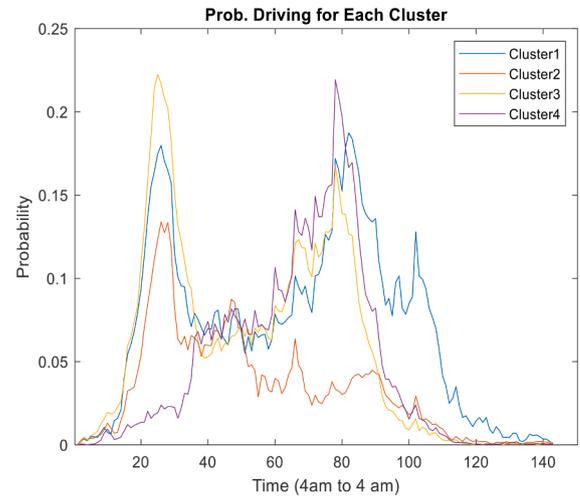


Fig. 5. Probability of driving for each of the 4 clusters.

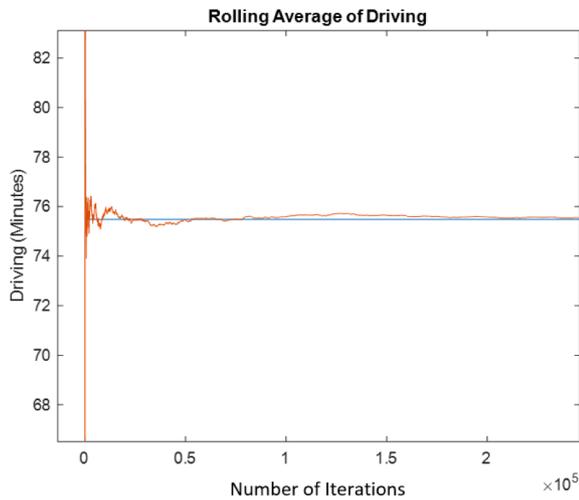


Fig. 4. Model verification of MCMC showing convergence to average of 75 min of driving time from TUS data.

Algorithm 1

Transitional matrix formulation.

- 1: Convert Diary Entries to Markov chain states
- 2: **for** Diary Entry=1,2,... **do**
- 3: **for** Timeinterval=1,2,...,144 **do**
- 4: Record current and previous markov chain states
- 5: **end for**
- 6: **end for**
- 7: Count number of repeated states transition at each time space
- 8: Divided by total number of Diary entries

that are of interest in this paper are shown in Fig. 3.

The three states shown are the vehicle parked at home, being driven and parked out with the home. The latter state here includes such locations as the vehicle parked at the work place or at commercial locations. It is important to distinguish these parked states separate from parked at home but at the same time, it is possible to expand the Markov chain to include more states. As shown in [23], additional states such as commercial parking can be added however, for the purpose of this paper, it was only necessary to observe when the vehicle is being driven, at home or otherwise.

Given the TUS data format, the model used in this paper will have a simulation period of 24 h with 144 discrete 10 min intervals. The states within the full Time- inhomogeneous Markov Chain will remain constant with the Vehicle movement through time varying i.e., the transitions between states, dictated by P_t . Taking into account the state space of 3, P_t will take the general form of:

$$P_t = \begin{pmatrix} P'_{11} & P'_{12} & P'_{13} \\ P'_{21} & P'_{22} & P'_{23} \\ P'_{31} & P'_{32} & P'_{33} \end{pmatrix} \quad (3)$$

For $t = 1$ to 144, where the subscript states of 1, 2 and 3 correspond to the states in Fig. 3 of driving, parked at home and not home/driving respectively. In order to achieve the elements with the transition matrix, the process outlined below was followed:

Following the process from Algorithm 1 results in 9 vectors, each containing the probability distribution function of the elements shown in Eq. (3).

3.2. MCMC verification

With the TUS data converted into a Time-inhomogeneous Markov Chain, a Monte Carlo simulation can be run to synthetically increase the population size. The probabilistic series vectors from Algorithm 1, that are stored in transition matrix P_t , can generate synthetic driving patterns using empirical PDF based Monte Carlo simulation as the sampling method. Having the ability to accurately generate a population that captures the stochastic and diverse driving behaviours of the TUS data is highly desirable but such model must be verified. The series of 100,000 simulations were run and compared against the original TUS data. Although a number of metrics can be used to ensure the efficacy of the data, for this paper the average number of minutes spent driving was used as this is the prominent factor used to determining EV flexibility.

4. Behavioural analysis

As discussed in Section 2.1, there are three prominent factors that must be considered within an EV model however there has been a lack of consideration into the significant of behavioural aspects (i.e., occupation/income) given its highly influential impact on mobility. As such this section details the process to disaggregate the TUS data into distinct clusters of participants coupling both behavioural aspects and mobility. It should be noted that in this paper, the process of vehicle mobility was extracted first before the overlaying of behavioural and technical aspects. It is true that heterogeneity modelled within a population is important and studies that withdraw impacts of demographics on the travel patterns risk the potential for hidden errors within their simulations. On the same token, extracting driving habits based solely on these metrics can also lead to hidden errors. For example, classifying groups of EV owners based on boundaries/limits of income brackets initially, may lead to a highly overlapping driving characteristics. Therefore, using machine learning, the dataset for driving behaviour broken down to

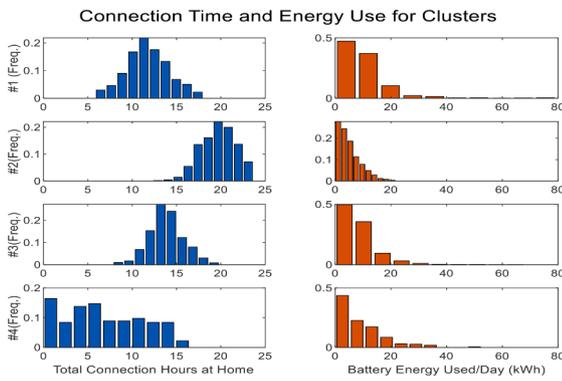


Fig. 6. Histograms of availability and energy used - The numbers calculated for the energy required uses the following assumptions (based on UK averages): battery efficiency of 12.9 kWh/100 km and average speed of 48 kmph.

distinct driving patterns first, followed by the disaggregation additional factors to identifying the key aspects resulting in distinct driving behaviour. Using, K-Means clustering, an unsupervised learning algorithm, the original TUS dataset can be partition data into 4 clusters. As the input data size is subject to the number of simulations produced by the MCMC simulation, the K-means clustering algorithm was seen as advantageous to this paper due to the algorithm having a linear time complexity, meaning it scales conveniently with larger datasets. Furthermore, the algorithm can be applied to ungrouped/unlabelled datasets that always converge to solution.

With the visualisation of the cluster outputs into driving probability, four distinct driving clusters were seen from the algorithm output, as shown in Fig. 5.

As stated, with the visualisation of the cluster outputs, 4 prominent clusters were identified. What Fig. 5 shows is total associated probability for each subgroup driving their vehicle throughout the time horizon. With that said, there are still significant overlap in time periods for some clusters such as Cluster 3 and 4 showing relatively similar patterns for the majority of the day. As an unsupervised machine learning technique was used, a deeper level of cross correlation is needed for further analysis. This involved deriving both simulated battery energy use of the EV as well as the numbers of hours connected at home.

In Fig. 6, for equal comparison, each histogram has been normalised and split between 10 bins to account for the different subgroup population sizes. Also, it is noted that sensitivity analysis on the ‘Energy Used’ assumptions can lead to different results however there would be a minimal impact the actual distribution spread of the chart as they have been normalised.

First, if the RHS energy used/day charts are examined, it shows for each cluster a relatively similar positively skewed distribution. For clusters 1, 3 and 4 the thicker bars represented a wider bracket for the bins as there is less of a distribution across the kWh range when compared to Cluster 2. With that said, sole comparison of these RHS charts again would lead to misleading results within driving behavioural models and classifications. Therefore, it was necessary to include the LHS chart on ‘availability’, that is the number of hours that a vehicle is home and available to charge. More specifically the percentage of time at home and available to charge is 28%, 58%, 82% and 49% for clusters 1, 2, 3 and 4 respectively. The issue discussed in the resulting similarity of driving behaviour in the clusters (from Fig. 5), when compared to that of the data shown in Fig. 6, is now more distinct. More specifically, the manner of grouping the subsets using the machine learning technique has also identified subgroups that follow a clear distribution in terms of number of hours connected at home. For example, the physical driving behaviour of clusters 3 and 4, although similar, show a significantly different distribution in terms of ‘availability’, hence the resultant classification differences. Such a distinction is of course important when considering the potential space for flexibility within an aggregators

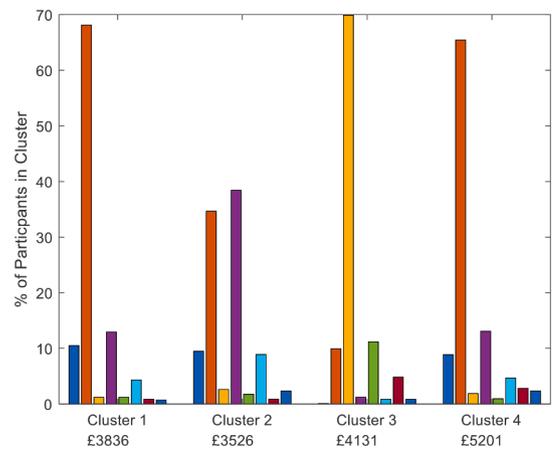


Fig. 7. Occupation status and net household income for each cluster: ‘Self-employed’, ‘Paid employment’, ‘Unemployed’, ‘Retired’, ‘Maternity’, ‘Carer’, ‘Student’, ‘Sick Leave’.

model, as discussed in the next section.

With the patterns and classifications identified above, it makes it possible target each subgroup with certain social- economic characteristics. With the advantage of using the UK TUS data, clusters can be further explored by extracting income and occupation status, to illustrate how these background variables heavily influence activity patterns (and thus driving behaviour).

Fig. 7 clearly shows a distinct make-up of varying occupational and income characteristics within each cluster. Visualisation of the data in this manner allows for the observation variety of insights that can draw the connection between factors that influence driving behaviour. Firstly, the overarching influential factors seems to be the number of participants in the clusters that are either ‘Self-employed’, ‘Paid employment’, ‘Unemployed’, or ‘Retired’ (the blue, red yellow and purple bars respectively). However, one distinction can be seen when comparing the occupational status of Cluster 1 and Cluster 4. Their respective occupational percentages are mirrored, however if you examine the net household income, this gives a deeper distinction in the resultant/varied driving pattern. This demonstrates a previous note made, that is, basing driving behaviour on that of solely income or occupational status, can lead to a mischaracterisation. Highlighting this insight is important when considering the heterogeneity of an EV population and lacking a more holistic disaggregation of additional factors can lead to errors in EV modelling and hide potential risks and uncertainties for aggregators.

5. Discussion: the energy space for district level aggregators

Given that a single EV’s participation within the current ancillary service markets is doubtful, the role of actors such as aggregators will play an increasingly important part in leveraging EV flexibility. With the model presented in this paper, EV datasets can be synthetically increased in population size, representing a variety of district level CPs for both network and population specific areas i.e., include area specific population heterogeneity. With the MCMC model, a sufficient number of simulations can also be run to provide statistically significant bounds of uncertainty, that is, a confidence interval (CI) of 99% for both energy requirements and availability, disaggregated within each EV cluster. Using CI analysis in this manner, can provide aggregators an insight into the space that is available to leverage EV flexibility whilst accounting for the inherent uncertainty for clusters of area specific population heterogeneity. With the disaggregation and analysis provided for each cluster, an aggregator can create a portfolio of locational specific assets, with a reduced degree of risk. An aggregator, even with direct load control (i.e., has physical infrastructure to control an EV charger) would still incur risk due to asset availability. Highlighting the ‘energy space’

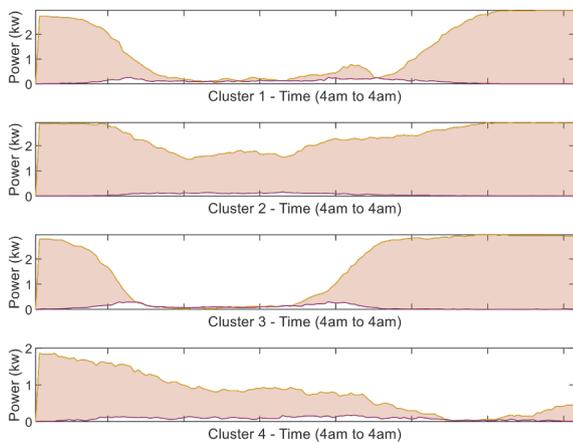


Fig. 8. Energy Space based on worst case 99% CI worst case scenarios for each cluster.

that encompasses the clustering and correlation analysis of population heterogeneity within one graph, can provide the necessary insights to reduce aggregator risk and uncertainty.

For each cluster, a series of 100,000 MCMC simulations were run to determine a CI of 99% for two statistics – energy used and availability throughout the time horizon. Using the ‘worst’ case bounds, that is, the upper 99% CI for energy needed and the lower for availability, an ‘energy space’ can be represented. More specifically, the ‘energy space’ refers to the area between the upper and lower boundaries that is used to characterize EV flexibility for an average EV within a given cluster group. This area defines the boundaries for the potential flexibility in discharging and charging an EV for ancillary service participation. Again, for equal comparison between each cluster, the graphs are represented by the ‘average’ of a single EV within each cluster’s population. Using the analysis and MCMC model from the previous section, Fig. 8 can be produced.

The shaded area between the two curves defines the energy space with a 99% CI. Defining the space for flexibility for each cluster in a manner that incorporates EV uncertainty is valuable information for aggregators participating in flexibility and ancillary services markets. An aggregator can use these graphs, for each cluster in two ways- for defining the boundary limits when 1) charging and 2) discharging the EV. In each case, it should be noted that the total energy displayed in the shade area is subject and bound to the total capacity within the EV battery. The scope of this paper was to define a methodology for aggregators and therefore the specific use cases is considered future work.

6. Conclusion

This paper has demonstrated a methodology for a probabilistic approach in representing flexibility for the use of EV aggregators. Developing a model derived from the UK TUS data has enabled both a verifiable method to produce statistically significant datasets as well as a process to effectively analyse key influential factors for EV characteristics such as CPs. This approach, using a stochastic MCMC model, gives aggregators an insight into the flexibility space that is available for market services whilst accounting for population heterogeneity induced risk and uncertainty. Analyses shows that a holistic model is required, one that includes an overlay and comparison of different technical, spatial and social-economic behavioural factors to ensure aggregators can interpret the data to understand and mitigate the risks associated with EV aggregation.

Funding

This research is supported by UK Engineering and Physical Sciences Research Council (EPSRC) through Centre for Doctoral Training in Future Power Networks and Smart Grids [EP/L015471/1]. All results can be fully reproduced using the methods and data described in this paper and provided references.

CRedit authorship contribution statement

Mark Jenkins: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – original draft. **Ivana Kockar:** Supervision, Writing – review & editing.

Declaration of Competing Interest

None

References

- [1] Committee of Climate. Change, Net zero: the UK’s contribution to stopping global warming,” in “Tech. Rep,” [Online], Tech. Rep (May 2019) [Online]. Available, <https://bit.ly/2LS8gkC>.
- [2] S. Errity, “Electric car sales UK,” Driving Electric. Available:2022. <https://www.drivingelectric.com/news/678/electric-car-sales-uk-market-down-but-continued-growth-for-evs-and-phevs>.
- [3] UK Dept. of Transport, “Government takes historic step towards net-zero with end of sale of new petrol and diesel cars by 2030,” 2022. Available: <https://www.gov.uk/government/news/government-takes-historic-step-towards-net-zero-with-end-of-sale-of-new-petrol-and-diesel-cars-by-2030>.
- [4] Ofgem, “Implications of the transition to electric vehicles,” July 2018. [Online]. Available:https://www.ofgem.gov.uk/sites/default/files/docs/2018/07/ofg1086_future_insights_series_5_document_master_v5.pdf.
- [5] C. Zhang, J. Wu, C. Long, M. Cheng, Review of existing peer-to-peer energy trading projects, Energy Procedia 105 (2017) 2563–2568, <https://doi.org/10.1016/j.egypro.2017.03.737>, 2017/05/01.
- [6] P. Meenakumar, M. Aunedi, G. Strbac, Optimal business case for provision of grid services through EVs with V2G capabilities, in: 2020 Fifteenth International Conference on Ecological Vehicles and Renewable Energies (EVER), 2020, pp. 1–10, <https://doi.org/10.1109/EVER48776.2020.9242538>, 10-12 Sept. 2020.
- [7] J.M. Clairand, J. Rodríguez-García, C. Álvarez-Bel, Assessment of technical and economic impacts of EV user behavior on EV aggregator smart charging, J. Mod. Power Syst. Clean Energy 8 (2) (2020) 356–366, <https://doi.org/10.35833/MPCE.2018.000840>.
- [8] Ofgem, “Independent aggregators and access to the energy market.” Available: <https://www.ofgem.gov.uk/publications/independent-aggregators-and-access-energy-market-ofgems-view>.
- [9] A.S. Al-Ogaili, et al., Review on scheduling, clustering, and forecasting strategies for controlling electric vehicle charging: challenges and recommendations, IEEE Access 7 (2019) 128353–128371, <https://doi.org/10.1109/ACCESS.2019.2939595>.
- [10] G. Pareschi, L. Küng, G. Georges, K. Boulouchos, Are travel surveys a good basis for EV models? Validation of simulated charging profiles against empirical data, Appl. Energy 275 (2020), 115318, <https://doi.org/10.1016/j.apenergy.2020.115318>, 2020/10/01.
- [11] J. Quiros-Tortós, A.N. Espinosa, L.F. Ochoa, T. Butler, Statistical representation of EV charging: real data analysis and applications, in: 2018 Power Systems Computation Conference (PSCC), 2018, pp. 1–7, <https://doi.org/10.23919/PSCC.2018.8442988>, 11-15 June 2018.
- [12] J. Brady, M. O’Mahony, Modelling charging profiles of electric vehicles based on real-world electric vehicle charging data, Sustain. Cities Society 26 (2016) 203–216, <https://doi.org/10.1016/j.scs.2016.06.014>, 2016/10/01.
- [13] P. Grahm, J. Munkhammar, J. Widén, K. Alvehag, L. Söder, PHEV home-charging model based on residential activity patterns, IEEE Trans. Power Syst. 28 (3) (2013) 2507–2515, <https://doi.org/10.1109/TPWRS.2012.2230193>.
- [14] Z. Darabi, M. Ferdowsi, Aggregated impact of plug-in hybrid electric vehicles on electricity demand profile, IEEE Trans. Sustain. Energy 2 (4) (2011) 501–508, <https://doi.org/10.1109/TSTE.2011.2158123>.
- [15] R.A. Verzijlbergh, Z. Lukszo, E. Veldman, J.G. Sloopweg, M. Ilic, Deriving electric vehicle charge profiles from driving statistics, in: 2011 IEEE Power and Energy Society General Meeting, 2011, pp. 1–6, <https://doi.org/10.1109/PES.2011.6039609>, 24-28 July 2011.
- [16] D. Fischer, J. Scherer, A. Flunk, N. Kreifels, K. Byskov-Lindberg, B. Wille-Haussmann, Impact of HP, CHP, PV and EVs on households’ electric load profiles, in: 2015 IEEE Eindhoven PowerTech, 2015, pp. 1–6, <https://doi.org/10.1109/PTC.2015.7232784>, 29 June-2 July 2015.
- [17] N.H. Tehrani, P. Wang, Probabilistic estimation of plug-in electric vehicles charging load profile, Electric Power Syst. Res. 124 (2015) 133–143, <https://doi.org/10.1016/j.epsr.2015.03.010>, 2015/07/01.

- [18] D. Wang, J. Gao, P. Li, B. Wang, C. Zhang, S. Saxena, Modeling of plug-in electric vehicle travel patterns and charging load based on trip chain generation, *J. Power Sources* 359 (2017) 468–479, <https://doi.org/10.1016/j.jpowsour.2017.05.036>, 2017/08/15.
- [19] J. Zhang, J. Yan, Y. Liu, H. Zhang, G. Lv, Daily electric vehicle charging load profiles considering demographics of vehicle users, *Appl. Energy* 274 (2020), 115063, <https://doi.org/10.1016/j.apenergy.2020.115063>, 2020/09/15.
- [20] D. Fischer, A. Harbrecht, A. Surmann, R. McKenna, Electric vehicles' impacts on residential electric local profiles – a stochastic modelling approach considering socio-economic, behavioural and spatial factors, *Appl. Energy* 233-234 (2019) 644–658, <https://doi.org/10.1016/j.apenergy.2018.10.010>, 2019/01/01.
- [21] The United Kingdom Time Use Survey, 2014-2015 (UKTUS) Available: <https://www.timeuse.org/survey-data>.
- [22] UK Dept. of Transport., "Electric vehicle charging in residential and non-residential buildings," ed, 2018. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/818810/electric-vehicle-charging-in-residential-and-non-residential-buildings.pdf.
- [23] Y. Wang, D. Infield, Markov Chain Monte Carlo simulation of electric vehicle use for network integration studies, *Int. J. Electrical Power Energy Syst.* 99 (2018) 85–94, <https://doi.org/10.1016/j.ijepes.2018.01.008>, 2018/07/01.