Electrical vehicle load modelling for distribution system considering future scenarios

Hanshan Qing[1], Abhinav Kumar Singh[1], Stratis Batzelis[1]
[1] *School of Electronics and Computer Science, University of Southampton, Southampton, United Kindom*
H.Qing@soton.ac.uk, a.k.singh@soton.ac.uk, e.batzelis@soton.ac.uk

***Abstract*—Electric vehicles (EVs) are a valuable means of reducing our reliance on traditional fossil fuel based transportation. In recent years, the market share of EVs is increasing, which raises some important questions: what is the impact of EVs on the electrical load of our electricity distribution systems, and how can we adequately model it? Some researches have been done to model the EV loads and influence, but they are limited in applications due to complexity and requirements of data. This paper proposes a model for EV loads which takes into account diversity of EV loads and difficulties in applications, and provides analysis of the EV loads in modern and future grids. Efficacy of the proposed model has been demonstrated using real EV-datasets, which provides valuable statistical analysis.**

***Index Terms*—Electrical vehicles load modelling , load profile generation**

I. INTRODUCTION

The number of Electrical Vehicles (EV) in many countries has been rapidly increasing. For UK, it increases to over 530,000 battery EVs plus 405,000 plug-in hybrid vehicles (PHV), which in turn raises the need for more charging infrastructure. According to Zap-Map database, there are 21378 charging locations, 35778 charging devices, 59059 connectors in UK for EVs by November 2022, which only considers public chargers [1].

The modelling of EV-loads becomes a interesting topic with more EVs on road. However, this is not an easy task, as there are few available datasets with insufficient information being recorded. One study provides a detailed review on open databases, but most of them do not contain enough information [2]. In spite of the lack of datasets, there are several works on existing literature which model the mobility and load of EVs for varied applications [3]–[15].

The main approaches to model EV are Markov chain based, models with assumptions based on distributions and spatial temporal models. The basic Markov chain and process are applied in [6]–[10]. Markov chain/process can describe the EVs well due to its nature which reflects the state transitions. However, it does not contain the parking duration in the model, which has been improved in [7] by introducing nonhomogeneous semi-Markov process which considers the sojourn times, that is the parking time for EVs. In [10], a Markov chain with graph theory has been applied to describe the EV behaviour in a real world distribution system. However, the common drawback of all these methods is that a clear and complete tracking record of EVs is required, which is not available in most cases.

It has been a widely used and a proven fact that the distributions which describe well the events related to EVs are normal and log-normal distributions. Many articles have used them to model the daily travel events start time and end time, daily charging events start time and end time, daily state of charge (SOC) of EVs, daily EV loads, and distribution of hourly EV loads, such as [3], [5], [7], [9], [10], [15]. This approach has fewer requirements of datasets, but still needs a large size dataset to extract the distribution. There are some datasets given by [2] that fulfil the requirement, such as [16] which gives over 30,000 data points in each year.

A dataset which meets the requirements could be hard to find, therefore, there are some research papers which endeavour to generate datasets from available datasets and empirical knowledge. UK national household travel survey (NHTS) has been used to generate EV charging load using travel record in [3]. They first derive the probabilities for different purposes from annual to hourly, then apply a cumulative density function for each purpose to transfer the number of travel events in each hour to travel distances for each EV in those travel events. Subsequently, the battery model they propose is used to generate the EV loads. A similar method has also been applied in [6]. These two works present approaches which have less requirements of datasets and produce various travel patterns and records rather than being limited by the input datasets (as these approaches use the NHTS which describes travel patterns of all kinds). However, some of the constituent modelling details of these approaches are quite complicated because of which these approaches may not be directly used to develop a 'general' EV load model considering future scenarios.

A simpler and general model which could describe the variable EV-loads would be much valuable. Another research paper applies variational auto-encoder to generate EV load data, in which the daily load profile is processed as an image, its features are extracted and EV load model is validated after decoding [14]. This method is easy to use after a network being pre-trained, but its outputs are limited by the inputs. In other words, load patterns of all kinds of EVs should be introduced to the network to get expected outputs, therefore, a large amount of datasets are required.

In this paper, a novel and simple EV load modelling method proposed in Section 2 replaces the complex subsequent models in [3], [6] with two simple steps and adds an up-sampling step, but preserves the advantage that various kinds of EV loads can be included. The up-sampling step also expands the potential applications of the model. Analysis based on open datasets are given in later sections for load profiles which are used for modelling EV loads. Analysis considering the statistical representations of EV loads are rare but important in some applications, such as generating pseudo measurements in power system state estimation problems. The analysis provides better view of characteristics of EV loads, and also proves the efficiency of the proposed method.

II. Travel pattern model

This section briefly introduces the method used to generate EV load dataset based on UK NHTS 2021.

A. Travel distance derivations

The starting steps of the proposed method remain same as [3] and [6]. The number of annual trips for different purposes are given in the [17], therefore, the probability $P_1(x_1)$, probability of trips for different purposes in years, could be derived as (1).

$$P_1(x_1) = \prod_{b=1}^{n_p} (\lambda_{(i)})^{x_1(b)} \tag{1}$$

x_1 is the purpose matrix which are zeros except for the qth element representing qth purpose, while v is the bth element in the vector. n_p is the number of purposes, and $\lambda_{(i)}$ is the probability of trips for purpose 'b'. Then the number of travel for a specific purpose could be calculated by (2).

$$N(q) = N_t \times N_v \times P_1(x_1(q)) \; \forall q = 1, 2, ...n_p \tag{2}$$

, where $N(q)$, N_t, and N_v are the number of travel for purpose q, number of travel in a year and number of EVs, respectively. Similarly the $P_2(x_2)$ and $P_3(x_3)$, which are monthly and weekly probabilities of a trip, can be calculated by replacing n_p by 12 and 7, respectively, to distribute the annual trips of each purpose into 12 months, and to distribute the weekly trips of each purpose into 7 days, respectively, as:

$$P_2(x_2) = \prod_{m=1}^{12} (\lambda_{(q,m)})^{x_2(m)} \tag{3}$$

$$N(q, M) = N(q) \times P_{2(q)(x_2)} \; \forall M = 1, 2, ...12 \tag{4}$$

$$P_3(x_3) = \prod_{d=1}^{7} (\lambda_{(q,d)})^{x_3(d)} \tag{5}$$

Then N_q is multiplied with $P_2(x_2)$ to get the number of trips in a specific month for purposes "q", $N(q, M)$. Then (6) is applied to calculate the daily trip by purposes.

$$N(q, M, D) = \frac{N(q, M) \times P_{3(q)(x_3)}}{\frac{nd(M)}{7}} \; \forall D = 1, 2, ...7 \tag{6}$$

, where nd is the number of days in the month "M". Subsequently, the hourly trip could be derived as:

$$P_4(x_4) = \prod_{h=1}^{24} (\lambda_{(q,h)})^{x_4(h)} \tag{7}$$

$$N(q, M, D, H) = N_{(q,M,D)} \times P_{4(q)}(x_4) \; \forall H = 1, 2, ...24 \tag{8}$$

The above method is applied to increase the resolution of travel record. 60 Gaussian random numbers are generated for each hour in which there are more than 60 travel events. Then the normalized value of these numbers is multiplied with total number of travel events and rounded-off to integers, which gives the travel events in every minute.

$$P_5(x_5) = \prod_{t=i}^{60} (\lambda_{(q,t)})^{x_5(t)} \tag{9}$$

, where $\lambda_{(q,t)} = \frac{e_i}{\sum t = 1^{60} e_i}$, 60 random numbers from Gaussian distribution that are normalized to make the cumulative probability equal to 1.

If there are travel events less than 60, random numbers between 0 and 59 with total amount corresponding to the number of events in each hour would be generated and aggregated at each minute as events number.

$$N(q, M, D, H, t) = N_{(q,M,D,H)} \times P_{5(q)}(x_5) \; \forall t = 1, 2, ...60 \tag{10}$$

Another step is added to this data generation process to simulate the charging decision making, which uses log-normal distribution to generate random numbers to represent decision making (since, EV charging events are found to follow log-normal distribution, as explained in Section I). The decision process is given as:

$$P_6(z) = \frac{1}{zs\sqrt{2\pi}} exp \left[\frac{-(ln \; z - u)}{2s^2} \right] \tag{11}$$

Different from the definition used in [3] for trip per vehicle, it can be assumed that the random number z is the number of travels left before the charging events for each EV. u and s are, respectively, the mean, defined as $N_{q,M,D}/N_v$, and standard deviation, assumed to be 1.

Trips per vehicle have not been considered in EV load modeling since the focus here is the total load of EVs which would be aggregated to one load at each bus in the distribution network. But another step needs to be added (just like [3]) to derive the travel distance and to translate it to energy consumption and SOC. The distance travelled in miles for a trip finished in hour t of the annual travel model of vehicle v under purpose q, ML(q,v,t), is calculated as (12).

$$ML_{(q,v,t)} = C_{7(q)}^{-1}(y) \tag{12}$$

The inverse of cumulative density function, C^{-1}, which is the probability of a trip of purpose 'q' to be less than a certain distance. y is a random number generated between 0 and 1, which follows normal distribution.

B. EV loads derivations

As travel distances in each travel events are defined, charging events should be derived, in which the SOC and energy consumption would be derived. The EV model here is Tesla

Model X which has 100kWh battery capacity, 325 kM All Electric Range (AER) [18]. In UK, there are four different ranges of charging power, which are slow (3-6kW), fast (7-22kW), rapid (25-99kW) and ultra-rapid (100kW+) [1]. The maximum charging power for ultra-rapid is assumed to be 250kW here, which is the charging power of Tesla X [18]. A random number is selected from these four ranges as charging power that remains constant during charging and follows normal distribution. Then all charging events are mapped to 84 days which is 7 days a month multiplied by 12 months.

III. Datasets comparisons

This section introduces the real datasets from UK governments for Leeds city and for Perth and Kinross council from [16], [19] and an open dataset from project "Custom-Lead Network Revolution" from [20] and compares these real datasets to generated dataset generated based on UK 2021 NHTS by method shown in Section 2. Further analysis has also be given.

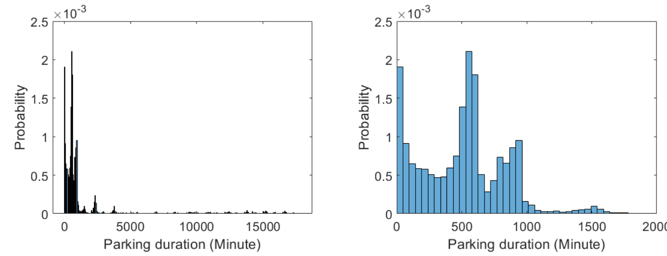
A. Leeds city and Perth and Kinros council datasets

The charging events record in [16], [19] includes charging start and end time, energy charged, and locations, which should be transferred to load record first. Because no specific data is given, constant charging power is assumed during charging to generate the load record. The dataset of Leeds city has few data points, which has the total number of events from 1900 to 4060 for every half year. That means that only events around 10 and 20 are recorded in each day. Therefore, the records in recent 3 years are aggregated to increase the density of charging events, which will make the load curve more practical.

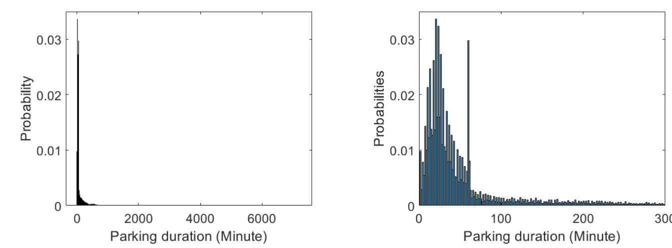
However, there is another issue that is the duration of events is long, which has mean value of 1515 minutes and maximum value of 17833 minutes. For the bin in the histogram, Fig. 1(a), largest probability is between [530, 578] which is shown in Fig.1(b). That makes it more like a parking events record at car parks rather than charging events at public stations, comparing with the one for Perth, which has mean value of 73.5 minutes and maximum value of 7256 minutes. The highest probability is at bin with range [20.70 23.28] in Fig.1(c) and 1(d).

The above hypothesis is proven by the summary provided in [19], which says "Usage data for Electric Vehicle charge points in Council car parks in Leeds", while the summary for Perth Kinross says "Datasets for Perth & Kinross Council's EV charging stations under the Charge Place Scotland scheme." [16]. Therefore, these datasets give the characteristics of charging at parking lots and stations, respectively.

Gaussian Mixture Models (GMM) is the weighted sum of some Gaussian distributions, which is widely used to model the historical load data in power systems and generate load profiles [21]. The following figures give the histogram of both the datasets, and GMM used to fit them, which has 20 components respectively. Besides, the number of bins is calculated following the Freedman-Diaconis rule [22]. The shape of histogram of Leeds' case is similar to a log-normal distribution, while the one for Perth Kinros looks like a



(a) Histogram of charging event duration of Leeds city's data. (b) Histogram of charging event duration of Leeds city's data after zoom.



(c) Histogram of charging event duration of Perth Kinros' data. (d) Histogram of charging event duration of Perth Kinros' data after zoom.

Fig. 1. Histogram of duration of charging events of two real datasets.

Weibull distribution. Despite the differences in histograms, both the datasets have similar shapes for their daily load profile.

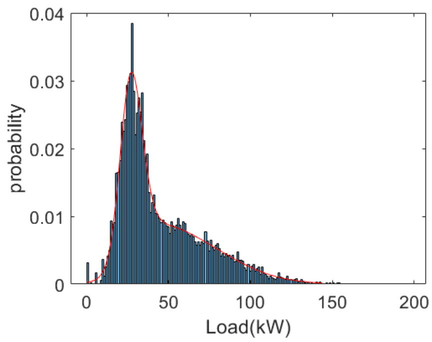
B. Home charging and EV charging data

There are 1,144 users who participated in "Custom-Lead Network Revolution" project [20]. In the experiment for EV, they placed one measurement for EV load and one for other household load, respectively, for some of the participants: 144 in the load-record. In this dataset, the energy consumed in every 10 minutes is provided for every participant. The histogram of it, Fig. 3(a), has a similar shape as the other two real-datasets, but some peaks at lower values could be observed. The daily load profile of home charging still follows normal distribution, but the peak is around 20:00. Besides, there is a small peak near the 50th data point, 8:20, which probably refers to the start of residents' early charging.

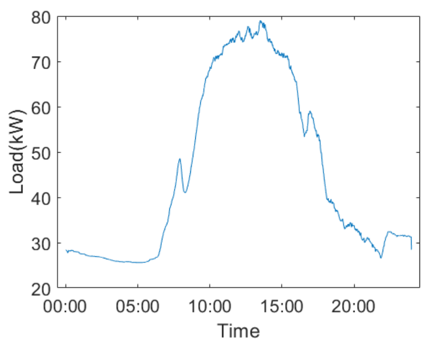
C. Datasets generated from UK 2021 NHTS

The dataset presented in this section is generated by applying the model described in Section 2. The survey used here is the UK NHTS 2021, which gives the number of annual, monthly, weekly and hourly travel events. These are translated to probability by normalization, which are the λs. The simulation describes 10,000 EVs' load record, which has 395 travels annually.

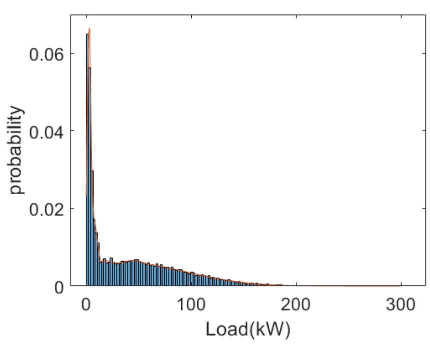
Since the survey contains all kinds of travel, the generated data contains all kinds of charging events, which is the advantage of the proposed method. Additionally, the total number of charging events could be modified by setting the annual numbers at start, which makes the datasets controllable and sufficiently large. Its histogram is similar to Perth Kinross'



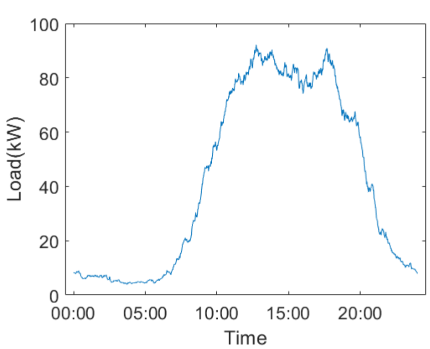
(a) Histogram of Leeds city's data and the GMM fitted.



(b) Daily load curve of Leeds city's data.

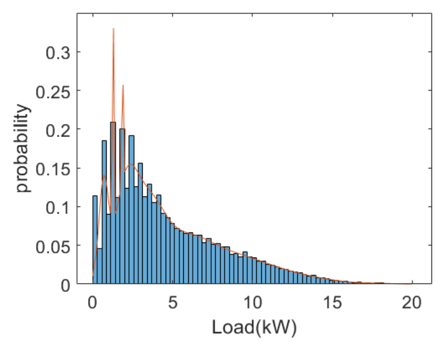


(c) Histogram of Perth and Kinros' data and the GMM fitted.

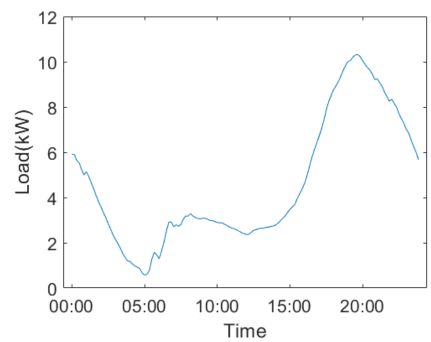


(d) Daily load curve of Perth and Kinros' data.

Fig. 2. Histogram of load record and daily load curve of two real datasets.



(a) Histogram of home charging data and the GMM fitted.



(b) Daily load curve of home charging data.

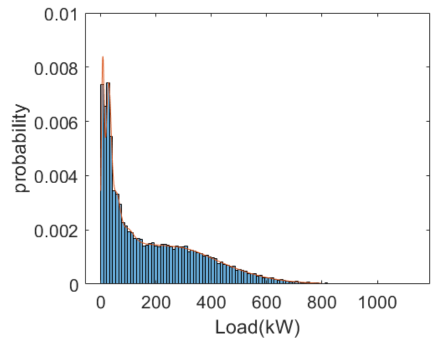
Fig. 3. Histogram of load record and daily load profile of home charging data.

data, but two peaks are observed in daily load curve, which is totally different from others.

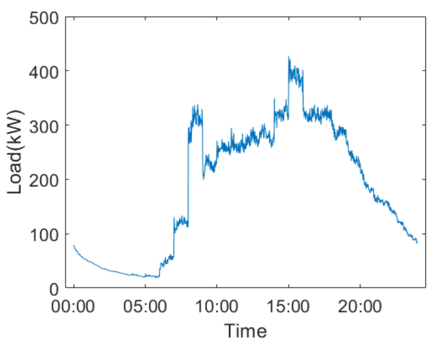
D. Comparison

The figure Fig. 5 plots all load profile of four datasets. The first peak of generated data could refer to the start of residents' activities of home charging, early arrivals charging at workplace, and some other activities. The kinds of travel recorded in the two datasets for Perth and Kinross and Leeds could be few due the limitation of locations and number of charging stations. Besides, the research in [13] gives a similar peak around 7 o'clock due to change in charging behaviours, which is the early charging in the morning and in the last several years of experiments. The middle portion between the first and second peaks could be considered as charging at car parks in workplaces after commuting as the Leeds' curves shows an increase in public activities and public station charging, while the second peak which represents the peak of public activities, similar to the curve of Perth Kinross shows. Similar to the home charging curve, a small peak is observed after 17:00 in the generated data's curve, which refers to the home charging becoming active but small due to power limitations at home.

The charging scenarios in the future would include charging at home, car parks and stations, which corresponds to the three real datasets discussed: project dataset, Leeds' dataset and Perth Kinross' dataset. The generated data could reflect the future load of EVs well, which contains two large peaks and one small peak for all three scenarios, if we assume the travel preference remains the same. However, it is possible the



(a) Histogram of generated data from NHTS and GMM fitted.



(b) Daily load curve of generated data from NHTS.

Fig. 4. Histogram of load record and daily load profile of generated datasets.

preference might change, but related discussion would be out of the scope of this article.

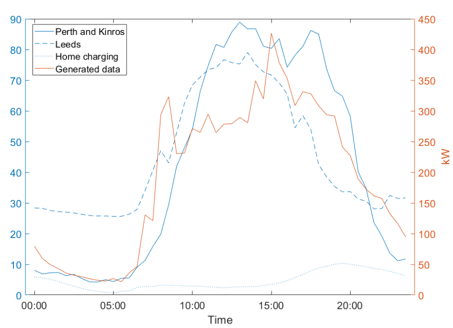


Fig. 5. Daily load profile of all datasets.

IV. Conclusion

In this paper, a novel and simple method for EV loads record generation was proposed. The load record was compared with three real datasets in Section 2. The comparison has demonstrated that the proposed modeling and data generation method is practical and is able to describe all kinds of EV loads, considering both current and future EV penetration scenarios. The influence of EVs at different penetration levels has not been discussed in this paper due to the limitation of time and length, which would be one of the future works. Additionally, realistic modelling of battery charging of EVs could be introduced.

References

[1] M. S. (Zap-Map), “Electric vehicle market statistics 2022,” Nov 2022. [Online]. Available: https://www.zap-map.com/ev-market-statistics/

[2] Y. Amara-Ouali, Y. Goude, P. Massart, J.-M. Poggi, and H. Yan, “A review of electric vehicle load open data and models,” *Energies*, vol. 14, no. 8, 2021. [Online]. Available: https://www.mdpi.com/1996-1073/14/8/2233

[3] M. F. Shaaban, Y. M. Atwa, and E. F. El-Saadany, “Pevs modeling and impacts mitigation in distribution networks,” *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1122–1131, 2012.

[4] M. Neaimeh, G. Hill, P. Blythe, R. Wardle, J. Yi, and P. Taylor, “Integrating smart meter and electric vehicle charging data to predict distribution network impacts,” in *IEEE PES ISGT Europe 2013*, 2013, pp. 1–5.

[5] K. Qian, C. Zhou, M. Allan, and Y. Yuan, “Modeling of load demand due to ev battery charging in distribution systems,” *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 802–810, 2011.

[6] A. Ul-Haq, C. Cecati, and E. El-Saadany, “Probabilistic modeling of electric vehicle charging pattern in a residential distribution network,” *Electric Power Systems Research*, vol. 157, pp. 126–133, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779617304765

[7] J. Rolink and C. Rehtanz, “Large-scale modeling of grid-connected electric vehicles,” *IEEE Transactions on Power Delivery*, vol. 28, no. 2, pp. 894–902, 2013.

[8] X. Wei, L. Lin, X. Yue, H. Yuan, and L. Yuanxi, “Optimal allocation for charging piles in multi-areas considering charging load forecasting based on markov chain,” in *2016 China International Conference on Electricity Distribution (CICED)*, 2016, pp. 1–7.

[9] X. Gao, L. Wei, B. Wang, G. Chen, and X. Wu, “Spatial load prediction considering spatiotemporal distribution of electric vehicle charging load,” in *E3S Web of Conferences*, vol. 256. EDP Sciences, 2021, p. 01001.

[10] M. B. Arias, M. Kim, and S. Bae, “Prediction of electric vehicle charging-power demand in realistic urban traffic networks,” *Applied energy*, vol. 195, pp. 738–753, 2017.

[11] J. Liu, R. Singh, and B. C. Pal, “Distribution system state estimation with high penetration of demand response enabled loads,” *IEEE Transactions on Power Systems*, vol. 36, no. 4, pp. 3093–3104, 2021.

[12] M. S. Islam and N. Mithulananthan, “Daily ev load profile of an ev charging station at business premises,” in *2016 IEEE Innovative Smart Grid Technologies-Asia (ISGT-Asia)*. IEEE, 2016, pp. 787–792.

[13] J. Schäuble, T. Kaschub, A. Ensslen, P. Jochem, and W. Fichtner, “Generating electric vehicle load profiles from empirical data of three ev fleets in southwest germany,” *Journal of Cleaner Production*, vol. 150, pp. 253–266, 2017.

[14] Z. Pan, J. Wang, W. Liao, H. Chen, D. Yuan, W. Zhu, X. Fang, and Z. Zhu, “Data-driven ev load profiles generation using a variational auto-encoder,” *Energies*, vol. 12, no. 5, p. 849, 2019.

[15] D. Zeng, k. Wang, Y. Li, X. Guo, and X. Jiang, “Load cluster characteristic analysis and modeling of electric vehicles,” *Engineering*, vol. 05, pp. 24–29, 01 2013.

[16] “Electric vehicle charging station usage,” Dec 2021. [Online]. Available: https://data.pkc.gov.uk/dataset/ev-charging-data

[17] D. for Transport, “National travel survey,” Aug 2022. [Online]. Available: https://www.gov.uk/government/collections/national-travel-survey-statistics

[18] “Model x.” [Online]. Available: https://www.tesla.com/en_gb/modelx

[19] L. C. Council, May 2021. [Online]. Available: https://www.data.gov.uk/dataset/5bb5c097-0e2f-42a3-8aee-1c0189a39082/electric-vehicle-chargepoints-operated-by-leeds-city-council

[20] “Project data download.” [Online]. Available: http://www.networkrevolution.co.uk/project-data-download/?dl=TC6.zip#

[21] R. Singh, B. C. Pal, and R. A. Jabr, “Statistical representation of distribution system loads using gaussian mixture model,” *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 29–37, 2010.

[22] D. Freedman and P. Diaconis, “On the histogram as a density estimator: L 2 theory,” *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 57, no. 4, pp. 453–476, 1981.