An Urban EV Charging Demand Generator For 343 Cities in China Based on AOI Information

Lyuzhu Pan, Hongcai Zhang
State Key Laboratory of Internet of Things for Smart City, University of Macau, Macao, China
hczhang@um.edu.mo

***Abstract*—In recent years, research on electric vehicles (EVs) has expanded rapidly, driven by their widespread adoption, which has resulted in a pressing demand for EV demand datasets. To this end, we propose a bottom-up charging demand generation method based on area-of-interest (AOI) information to simulate daily charging demand for 343 cities in China. We first construct a behavioral base by fitting charging data from 3 data-rich cities. Building on the derived behavioral parameters, charging sessions for all stations in data-scarce cities are generated, where each target station references a counterpart selected according to its AOI type. Finally, station-level sessions are aggregated to obtain city-level charging demand. The results show that the proposed method can well reproduce charging behaviors in 3 data-rich cities. Moreover, the simulated results basically follow real-world charging demand in data-scarce cities, demonstrating the proposed method's feasibility. As the proposed framework relies solely on open-source datasets, it can be readily reproduced for other countries with corresponding inputs.**

***Index Terms*—electric vehicle, charging behavior, demand generation, AOI**

I. Introduction

The transportation sector is a major contributor of greenhouse gas emissions, where land transportation accounts for 12% of global emissions [1]. To mitigate transport emissions, transportation electrification has become a key decarbonization strategy worldwide. For example, by 2024, electric vehicle (EV) stock in China exceeded 31.4 million, with a sale share approaching 50% [2]. To accommodate the growing charging demand, public charging infrastructure has been continuously expanded. However, with the increasing adoption of fast and ultra-fast chargers, EV charging behavior may exert significant negative impacts on the power grid [3]. For instance, rated capacities of large fast-charging stations can easily exceed ten megawatts, and uncontrolled charging may lead to local overloads in urban power distribution networks, seriously threatening grid stability.

To figure out the impact of EV charging and operation and planning of charging stations, a real-world charging dataset is vital for studies. In the past decade, extensive datasets have been published. For instance, Li *et al.* [4] present a charging station occupancy dataset in Shenzhen, China, involving 1,682 public charging station covering several months. Zhang *et al.* [5] provide a charging session dataset in Jiaxing, China, covering 2 years with 13 charging stations. Apart from datasets in China, many countries publish related datasets as well. For example, Baek *et al.* [6] publish a dataset from South Korea involving the EV user index. Amara-Ouali *et al.* [7] gather multi-source datasets from several countries, including the UK, USA, and France, to facilitate real-world studies. Furthermore, Guo *et al.* [8] also integrate datasets from 6 countries into a standard format, including rich environment and electricity price information. With the reliable real-world datasets, data-driven methodologies can generate more charging data or predict future charging demand. Lahariya *et al.* [9] adopt a Netherlands dataset to generate charging sessions based on Monte Carlo simulation. Shang *et al.* [10] and Qu *et al.* [11] adopt learning-based techniques to predict the urban EV charging demand.

The above studies can effectively apply the real-world data to generate or predict the urban charging behavior and demand in the *data-rich cities*, while it is hardly reproducible for a *data-scarce city*. However, it is impossible for every city to release the charging dataset due to data privacy. As a result, research involving these data-scarce cities is hard to conduct. To this end, some studies focus on estimating the charging behavior using the geographic information, such as the area of interest (AOI). AOI is a type of identifier to understand the urban functional area, such as commercial, residential, and work areas. Ren and Sun [12] predict the charging demand of public stations in Jinan, China, using population, road network, and AOI information. Kang *et al.* [13] adopt AOI information to reveal the charging behavior heterogeneity in Beijing, China, indicating the correlation between EV charging demand and urban facilities. Nevertheless, these studies are limited to a relatively small scale.

To bridge the research gap, we propose a bottom-up city charging demand generation method based on AOI information. Specifically, we first fit the distribution parameters of station-level charging behavior for different AOIs using data-rich cities. We then transfer the obtained station-level distributions to the data-scarce cities according to AOI information. The main contributions of this paper are twofold:

1) We propose a bottom-up city-level charging demand generation method, in which station-level charging behavior is characterized by two probabilistic distributions: EV arrival and charging duration, and each station is labeled with AOI information. The charging behaviors obtained from data-rich cities are then transferred to

This paper is funded in part by the Science and Technology Development Fund, Macau SAR (File no. 0094/2024/AGJ and File no. 001/2024/SKL), and in part by the Collaborative Research Grant of University of Macau (File no. MYRG-CRG2025-00014-IOTSC).

979-8-3315-8138-1

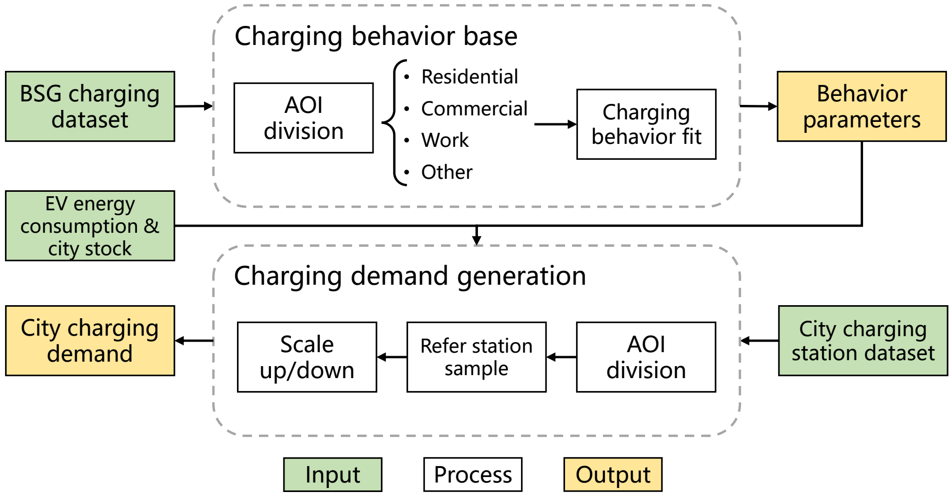


Fig. 1. Framework of the proposed method.

data-scarce ones. Unlike existing studies that estimate demand from the EV demand side or directly predict city-level demand, the proposed bottom-up approach first models station-level behavior and subsequently aggregates it to the city level, offering improved interpretability. Moreover, as the proposed framework relies entirely on open-source datasets, it is reproducible and can be readily adapted to other countries with corresponding data inputs.

2) We generate the typical weekday and weekend city charging demand with 5-minute intervals over 12 months for 343 Chinese cities, which has been released on GitHub [14]. The charging demand can be readily scaled to the future scenarios based on the China EV stock prediction.

II. Charging Behavior Model

We first present the basic framework of the proposed method. The proposed method consists of two major parts, as shown in Fig. 1: charging behavior base construction and charging demand generation for any city. The charging station will be divided into four AOIs: residential, commercial, work, and other areas. All charging stations in the data-rich city dataset will hold a set of charging behavior parameters, including arrival and charging duration. The first part will output a *charging behavior base*, which stores all behavior parameters with AOI labels. Then, for any city, we input all charging station information of this city, including pile scale, rated power, and AOI. For every station in the target city (a data-scarce city), we choose a reference station for the target station from the charging behavior base according to its AOI and station scale. Building on the reference station, we sample charging sessions for the target station. Repeating for all stations in the target city, we aggregate the charging sessions into charging demand and scale up/down according to the EV energy consumption dataset (i.e., EV energy consumption & city stock in Fig. 1). Finally, we can output the charging demand for any city in China.

A. Input Dataset

Three datasets are mainly used in this study, including the BSG charging dataset, EV energy consumption & city stock, and city charging station dataset.

1) BSG charging dataset: it is a charging session dataset published in reference [15], including over 700 thousand charging sessions of nearly 1,700 charging stations from 3 China first-tier cities: Beijing, Shanghai, and Guangzhou. The dataset consists of five columns: *gun_id, station_id, time_start, gun_power, time_end*. Besides, the AOI information of each station is provided.

2) EV energy consumption & city stock: we gather open-source reports and multi-source datasets to create this dataset, which has been published on GitHub [14]. It consists of two parts: China's monthly EV energy consumption (only for public charging stations), and China city EV stock.

3) City charging station dataset: it is collected by the API of the Chinese map platform. It includes the pile scale, rated power, AOI, and location of all public charging stations.

B. Charging Behavior Base Construction

When constructing the charging behavior base, we fit the charging behavior parameters for every station. We use $\mathcal{P} = \{p_{s,h,w}\}$ to denote charging behavior base, where $s \in \mathcal{S}^{\text{base}}$ is the index of all stations in the BSG charging dataset, $h \in \mathcal{H}$ denotes the hour index of 24 hours and $w \in \{\text{weekday}, \text{weekend}\}$ recognize the weekday/end. Note that the base station set $\mathcal{S}^{\text{base}}$ has three subsets $\mathcal{S}^{\text{bj/sh/gz}}$, denoting the station in different cities. The charging behavior includes EV arrival and charging duration. The parameter of arrival is denoted by $p^{\text{arr}}_{s,h,w}$, and the arrival numbers are denoted by $N_{s,h,w}$. Similarly, symbols $p^{\text{dur}}_{s,h,w}$ and $T_{n,s,h,w}$ are parameters and the time of the duration. Note that the charging duration for each arrival n is different, thus there is an additional subscript in $T_{n,s,h,w}$. Then, charging behavior parameter $p_{s,h,w} = \left[p^{\text{arr}}, p^{\text{dur}}\right]_{s,h,w}$. The process of charging behavior base construction is summarized in **Algorithm 1**.

We use the Poisson distribution to depict the arrival behavior:

$$N_{s,h,w} \sim \text{Poisson}(p^{\text{arr}}_{s,h,w}), \forall s, h, w, \tag{1}$$

$$p^{\text{arr}}_{s,h,w} = \frac{\sum_d \lambda_{s,h,d,w}}{|\mathcal{D}_w|}, \forall s \in \mathcal{S}^{\text{base}}, h, w, \tag{2}$$

where $\lambda_{s,h,d,w}$ is the arrival rate of the station s at hour h of day d in weekday/end w. The index $d \in \mathcal{D}$ denotes the date of the dataset, and we use a subscript w to distinguish the date in weekday or weekend. As shown in equation (2), we use the average arrival rate at hour h to denote the arrival parameter.

We use the Weibull distribution to depict the charging duration:

$$T_{n,s,h,w} \sim \text{Weibull}(p^{\text{dur}}_{s,h,w}), \forall s, h, w, \tag{3}$$

$$p^{\text{dur}}_{s,h,w} = \left[p^{\text{dur,k}}_{s,h,w}, p^{\text{dur,c}}_{s,h,w}\right], \forall s \in \mathcal{S}^{\text{base}}, h, w, \tag{4}$$

$$p^{\text{dur,k}}_{s,h,w} = \frac{\sum_d k_{s,h,d,w}}{|\mathcal{D}_w|}, \forall s \in \mathcal{S}^{\text{base}}, h, w, \tag{5}$$

$$p^{\text{dur,c}}_{s,h,w} = \frac{\sum_d c_{s,h,d,w}}{|\mathcal{D}_w|}, \forall s \in \mathcal{S}^{\text{base}}, h, w. \tag{6}$$

The Weibull distribution contains two parameters: shape parameter k and scale parameter c. Therefore, the charging duration parameter contains two parameters as formulated in equation (4). We use the average k and c to represent the charging duration parameters as formulated in equation (5) and (6), respectively.

Algorithm 1 Charging Behavior Model

```
Input dataset: BSG charging dataset
for ∀s ∈ S^base do
  for ∀w ∈ {weekday,weekend} do
    for ∀h ∈ H do
      Fit EV arrival parameters p^arr_{s,h,w} according to equation (2)
      Fit charging duration parameter p^dur_{s,h,w} according to equation (5) and (6).
    end for
  end for
end for
```

C. Charging Demand Generation

The charging station dataset stores station information for all cities in China. We select any city and use $\mathcal{S}^{\text{tar}}$ to denote the station set of the target city (data-scarce city). The charging demand generation can be divided into two steps: generate the charging demand curve and then scale up/down. The charging demand generation is summarized in **Algorithm 2**.

Firstly, we should select a reference city (i.e. $\mathcal{S}^{\text{bj/sh/gz}}$) for the target city. In this study, we select the reference city according to the latitude. For instance, the cities in northern China select Beijing as a reference city. Then, with any charging station in $\mathcal{S}^{\text{tar}}$, we randomly select a reference station from the $\mathcal{S}^{\text{base}}$ according to AOI type and pile scale. We categorize the charging stations into small, medium, or large stations, according to pile scales. For instance, a station in the residential area with small-scale piles will also randomly select the reference station from stations in the same AOI and small scale. Building on the charging behavior parameter from the reference station $p_{s,h,w}$, we should scale its $p^{\text{arr}}_{s,h,w}$ according to the pile ratio of the reference station and target station. This is because the pile number is not necessarily the same. When we sample the charging sessions, we first sample arrivals, i.e., $N_{s,h,w}$. For each arrival, we assign a pile for the arrival. The pile information (pile scale and rated power) is obtained from the City charging station dataset. Then, we sample the charging duration for the arrival. Finally, we output charging sessions, including station index, arrival time, charging power, and charging duration. The departure time can be easily calculated.

After we sample charging sessions for all stations in the target city, we can obtain a charging demand curve. As we know the China monthly charging demand and city EV stock, we can estimate the daily EV energy consumption for weekday and weekend, which is formulated below:

$$E^{\text{tar}}_m = \sum_w N^{\text{day}}_{m,w} \tilde{E}^{\text{tar}}_{m,w}, \forall m, \tag{7}$$

$$\tilde{E}^{\text{tar}}_{m,w=\text{weekday}} = r^{\text{tar}} \tilde{E}^{\text{tar}}_{m,w=\text{weekend}}, \forall m, \tag{8}$$

where E^{tar}_m is the monthly energy consumption for the target city in month m, and $N^{\text{day}}_{m,w}$ is the number of weekday or weekend for month m. Symbol r^{tar} is the ratio of daily energy consumption on weekday $\tilde{E}^{\text{tar}}_{m,w=\text{weekday}}$ and that on weekend $\tilde{E}^{\text{tar}}_{m,w=\text{weekday}}$. In equation (7) and (8), symbols $\tilde{E}^{\text{tar}}_{m,w}, \forall w$ are variables, while other symbols are constants. Therefore, we can solve $\tilde{E}^{\text{tar}}_{m,w}, \forall w$ by combining two equations. With the daily energy consumption, we can scale up or down the simulated charging demand by following the equation:

$$X^{\text{real}}_{m,w} = \frac{\tilde{E}^{\text{tar}}_{m,w}}{\tilde{E}^{\text{tar,sim}}_{m,w}} X^{\text{sim}}_{m,w}, \forall m, w, \tag{9}$$

where $X^{\text{real}}_{m,w}$ and $X^{\text{sim}}_{m,w}$ are the real and simulated charging demand curve, respectively, on weekday/end, month m. Symbol $\tilde{E}^{\text{tar,sim}}_{m,w}$ is the daily energy consumption of simulated charging demand.

III. Results

In this study, the charging demand curve involves 12 months. For brevity, we verify and illustrate results in January as an example.

A. Result Verification

We first verify whether the proposed charging behavior model can reflect the real-world behaviors. To this end, we simulate the charging sessions for 33 days (the same length as the BSG charging dataset), wherein the weekday and weekend during the period are separately simulated using associated parameters. The comparisons are illustrated in Fig. 2, the first

979-8-3315-8138-1

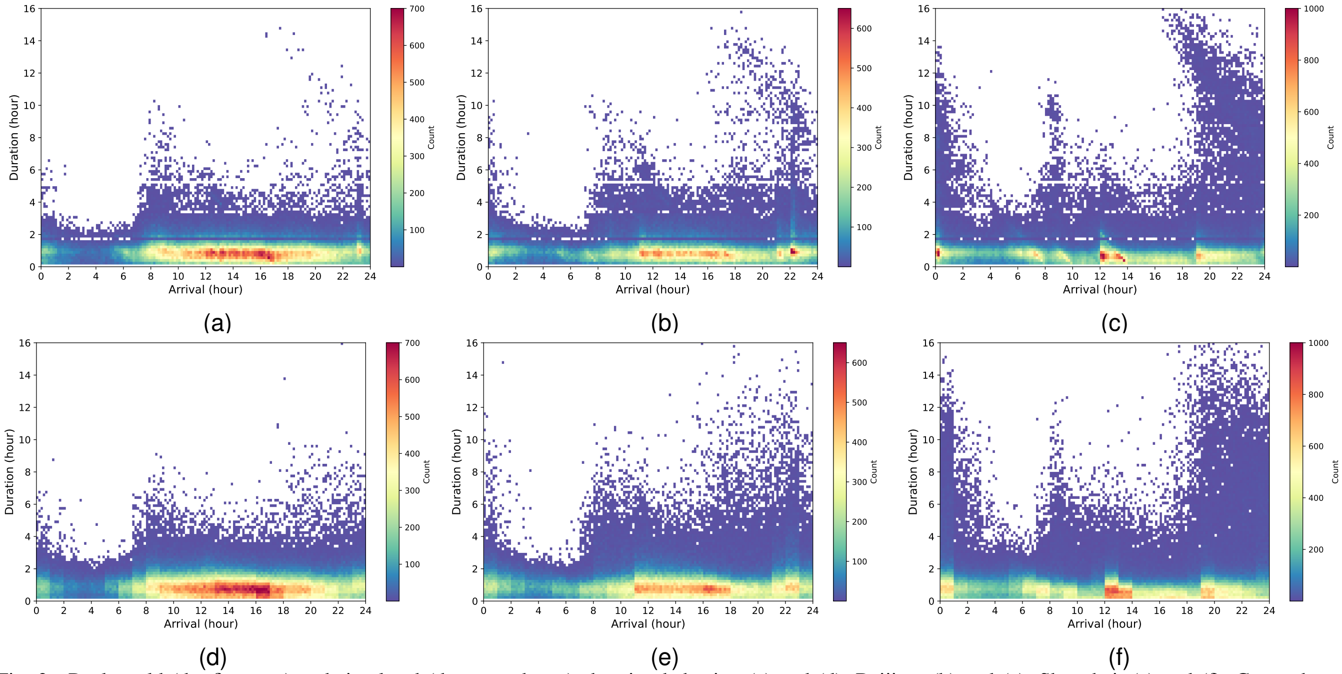


Fig. 2. Real-world (the first row) and simulated (the second row) charging behavior. (a) and (d): Beijing; (b) and (e): Shanghai; (c) and (f): Guangzhou.

Algorithm 2 Charging Demand Generation

```
Input dataset: EV energy consumption & city stock and City charging station dataset
Given a target city, select a reference city S^base
for ∀s ∈ S^tar do
    Select a reference station p_{s,h,w}, s ∈ S^base, ∀h according to AOI and pile scale
    for ∀h ∈ H do
        Scale p^arr_{s,h,w} according to pile scale
        Sample EV arrivals N_{s,h,w} according to equation (1)
        for ∀n ∈ {1, 2, ..., N_{s,h,w}} do
            Assign a pile for arrival n
            Sample charging duration T_{n,s,h,w} for arrival n according to equation (3)
        end for
    end for
end for
Scale up/down the charging demand curve from X^sim_{m,w} to X^real_{m,w} according to equation (9)
```

row is the real-world behavior, while the second row is the simulated one. In each sub-figure, the horizontal axis is the EV arrival time, and the vertical axis is the corresponding charging duration. The result shows that the simulated behavior can basically reflect the real-world charging behavior of three cities. Although the real-world behavior cannot be perfectly repeated, the simulated result can effectively depict the charging behavior during peak hours. We then verify whether the proposed model can be transferred to the data-scarce city. We only find a sufficient dataset in Shenzhen [4], China, due to limited data availability. The Shenzhen dataset is a charging occupancy dataset with a 5-minute interval, covering nearly 1,700 stations; thus, we believe that it can basically represent the charging demand curve of Shenzhen. The Shenzhen dataset includes the occupancy of slow and fast charging piles and the corresponding rated power. Therefore, we transform this dataset into a charging demand curve by multiplying the number of used piles and the corresponding rated power. The normalized charging demand curves from real-world and simulated datasets are presented in Fig. 3. The comparison shows that the simulated result can well follow the real-world curve in most of the time, apart from the early morning period. To explain the error during early morning, we should analyze from the bottom perspective. The charging demand curve of Shenzhen is based on the charging behavior from Guangzhou. Besides, two mega cities share the same time-of-use electricity price and similar AOI mix of stations. Therefore, a reasonable explanation of the error is the charging preference difference of Shenzhen and Guangzhou. However, the major limitation of this study lies in that it cannot take subtle behavior difference into consideration, leading to the existence of early morning error. Nevertheless, the simulated result can effectively reflect the demand peak in 00:00 and 13:00, and the platform after 20:00.

B. Urban Charging Demand Generation

We adopt **Algorithm 2** to generate charging demand for 343 cities in China. We plot the charging demand curve of the data-rich city and one corresponding data-scarce city, as shown in Fig. 4, Fig. 5, and Fig. 6. The data-scarce city will exhibit a similar feature to its reference city in terms of the start time of demand peaks. However, the details of the curve shape will be different because the AOI mix and rated power

979-8-3315-8138-1

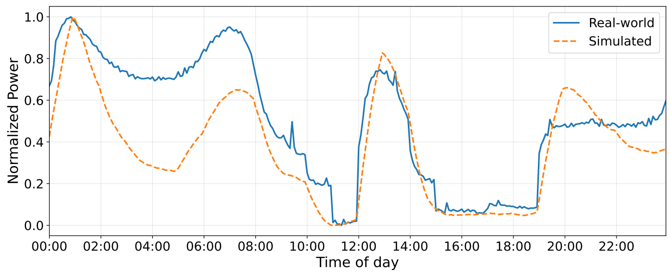
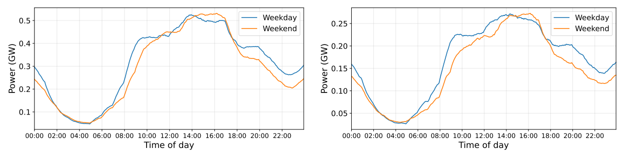


Fig. 3. Real-world and simulated charging demand curve in Shenzhen (data-scarce city).



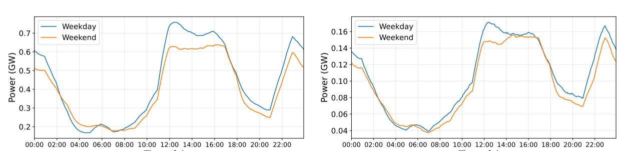
(a) (b)

Fig. 4. Simulated charging demand curve, (a) Beijing and (b) Tianjin.

distribution are generally different. Specifically, the AOI mix will affect the selection of the reference station in the charging behavior base, leading to a subtle difference in the shape of the curve. Besides, some cities may invest more in charging stations, thus they have more fast or ultra-fast stations. As a result, the demand will surge faster and the slope will be steeper.

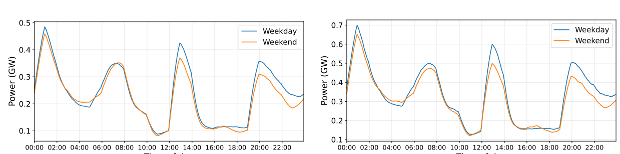
IV. Conclusion

In this study, we propose a bottom-up charging demand generation method based on AOI information. We first fit the charging behavior using the BSG charging dataset, where the charging arrival is depicted by the Poisson distribution, and the charging duration is depicted by the Weibull distribution. After constructing the charging behavior base, we generate the charging demand for any data-scarce city. We first find a reference city for the data-scarce city and select the reference station for every station in the data-scarce city. After that, we use the behavior parameter to sample the charging sessions for every station, then aggregate them into the city-level charging demand. Finally, the aggregated urban-level demand will be scaled up/down according to the real-world daily energy consumption of the public station. The result shows that the proposed method can basically reflect the real-world charging behavior in the data-rich cities (Beijing, Shanghai, Guangzhou). Besides, the verification in the data-scarce city (Shenzhen) shows that the charging demand by the proposed method can well follow the real-world one. The proposed method mainly uses open-source datasets; thus, it is easily reproducible, even for other countries. Besides, the generated charging demand can be readily scaled into future scenarios based on EV stock prediction, providing a friendly data source for researchers focusing on the nation-level EV studies.



(a) (b)

Fig. 5. Simulated charging demand curve, (a) Shanghai and (b) Suzhou.



(a) (b)

Fig. 6. Simulated charging demand curve, (a) Guangzhou and (b) Shenzhen.

References

[1] S. R. Department, “Topic: Transportation emissions worldwide.” https://www.statista.com/topics/7476/transportation-emissions-worldwide/.

[2] “Global EV Outlook 2025 – Analysis.” https://www.iea.org/reports/global-ev-outlook-2025, 2025.

[3] H. Zhang, X. Hu, Z. Hu, and S. J. Moura, “Sustainable plug-in electric vehicle integration into power systems,” *Nat. Rev. Electr. Eng.*, vol. 1, no. 1, pp. 35–52, 2024.

[4] H. Li, H. Qu, X. Tan, L. You, R. Zhu, and W. Fan, “UrbanEV: An Open Benchmark Dataset for Urban Electric Vehicle Charging Demand Prediction,” *Sci. Data*, vol. 12, no. 1, p. 523, 2025.

[5] Y. Zhang, T. Xu, T. Chen, Q. Hu, H. Chen, X. Hu, and Z. Jiang, “A high-resolution electric vehicle charging transaction dataset with multidimensional features in China,” *Sci. Data.*, vol. 12, no. 1, p. 643, 2025.

[6] K. Baek, E. Lee, and J. Kim, “A dataset for multi-faceted analysis of electric vehicle charging transactions,” *Sci. Data*, vol. 11, no. 1, p. 262, 2024.

[7] 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, p. 2233, 2021.

[8] Z. Guo, L. You, R. Zhu, Y. Zhang, and C. Yuen, “A City-scale and Harmonized Dataset for Global Electric Vehicle Charging Demand Analysis,” *Sci. Data*, vol. 12, no. 1, p. 1254, 2025.

[9] M. Lahariya, D. F. Benoit, and C. Develder, “Synthetic Data Generator for Electric Vehicle Charging Sessions: Modeling and Evaluation Using Real-World Data,” *Energies*, vol. 13, no. 16, p. 4211, 2020.

[10] Y. Shang, D. Li, Y. Li, and S. Li, “Explainable spatiotemporal multi-task learning for electric vehicle charging demand prediction,” *Appl. Energy*, vol. 384, p. 125460, 2025.

[11] H. Qu, H. Kuang, Q. Wang, J. Li, and L. You, “A Physics-Informed and Attention-Based Graph Learning Approach for Regional Electric Vehicle Charging Demand Prediction,” *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 10, pp. 14284–14297, 2024.

[12] Q. Ren and M. Sun, “Predicting the spatial demand for public charging stations for EVs using multi-source big data: An example from jinan city, china,” *Sci. Rep.*, vol. 15, no. 1, p. 6991, 2025.

[13] J. Kang, H. Kong, Z. Lin, and A. Dang, “Mapping the dynamics of electric vehicle charging demand within Beijing’s spatial structure,” *Sustain. Cities Soc.*, vol. 76, p. 103507, 2022.

[14] “Uban-ev-charging-demand-dataset-for-343-cities-in-china.” https://github.com/panda1225/Uban-EV-charging-demand-dataset-for-343-cities-in-China.

[15] Q. Yu, P. Zhao, J. Li, H. Wang, J. Yan, and H. Zhang, “China’s urban EV ultra-fast charging distorts regulated price signals and elevates risk to grid stability,” *Nat. Commun.*, vol. 16, no. 1, p. 8451, 2025.

979-8-3315-8138-1