

Supplementary Information for Unlocking vehicle-to-grid potential of load shifting in China's megacities considering comprehensive real-world behaviors

Kaisan Li¹, Xinxin Li¹, Zuxun Xiong², Shengyu Tao¹, Gucheng Zhao¹, Yi Jiang^{1,3}, He Qi⁴ and Yi Zhang^{1*}

^{1*}Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen, 518055, China.

²Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, UK.

³Department of Building Science, School of Architecture, Tsinghua University, Beijing, 100084, China.

⁴China Construction Science and Technology Group Co. Ltd, Beijing, 100195, China.

*Corresponding author(s). E-mail(s): zy1214@sz.tsinghua.edu.cn;

Supplementary Items List

Supplementary Figures

Supplementary Fig. 1: Extraction of stays	4
Supplementary Fig. 2: Distributions of n_w, β_1, β_2 from 2022 BPTD	5
Supplementary Fig. 3: Distributions of n_w, β_1, β_2 from 2022 SPTD	6
Supplementary Fig. 4: Circadian travel rhythms $P(t)$ of Beijing and Shenzhen P-PEV users	8
Supplementary Fig. 5: Grouped spatial parameters of P-PEVs in Beijing	9
Supplementary Fig. 6: Grouped spatial parameters of P-PEVs in Shenzhen	9
Supplementary Fig. 7: ToU information of Shenzhen and Beijing	13
Supplementary Fig. 8: Explanations of charging potential	14
Supplementary Fig. 9: Flowchart of the V2G participation mechanism	15
Supplementary Fig. 10: Selection of seasonal urban loads of Shenzhen	16
Supplementary Fig. 11: Comparison of results and computational costs between the mixed-integer linear programming (MILP) optimization and the linear programming (LP) optimization	18
Supplementary Fig. 12: Comparison of different forms of f_2 to eliminate SCD	19
Supplementary Fig. 13: Key mobility characteristics extracted from 2022 Beijing P-PEV Trajectories Dataset (2022 BPTD)	20
Supplementary Fig. 14: Validation of charging behavior simulation results of 2022 BPTD under Charge Only	21
Supplementary Fig. 15: Weekly peak load distribution of P-PEVs in Beijing, divided into stay regions	22
Supplementary Fig. 16: The load curve of Beijing P-PEVs, divided into charging places	23

Supplementary Notes

Supplementary Note. 1: Details of the mobility framework	3
--	---

Supplementary Tables

Supplementary Table. 1: Comparison of median $n_w, n_w \cdot \beta_1, n_w \cdot \beta_1$ between 2022 BPTD and 2022 SPTD	7
Supplementary Table. 2: Example of L-th most visited location for a given P-PEV user	10
Supplementary Table. 3: Comparison of average travel distance between 2022 BPTD and 2022 SPTD	11
Supplementary Table. 4: Numbers of parking lots in each district of Beijing	11
Supplementary Table. 5: Comparison of charging behavior parameters among different assumed cases	12
Supplementary Table. 6: China top-10 PEV market share data in 2019 (Estimated)	17
Supplementary Table. 7: Comparison of constraints applied according to user actions in different scenarios	23

Supplementary References

Supplementary References	24
--------------------------------	----

Supplementary Note. 1: Details of the mobility framework.

The mobility framework identifies stay points from trajectories. According to the data accuracy and individual traveling characteristics, different stay points may stand for the same activity location. Hence, we merge all the stay points into a stay region through grid-based clustering. Specifically, we discretize the city map into rectangular cells 300 meters wide, and we calculate the number of stay points in each cell. Then, starting with the cell with maximum stays, we repeatedly merge the unlabeled cell nearby into the labeled cell to get a new stay region. The maximum range of a stay region is set as 900 meters wide. In Supplementary Fig. 1, the three stay points $s1$, $s2$ and $s3$ are merged into a new stay region $r1$.

The mobility framework identifies Home and Work places for single P-PEV user. Based on the identification methods introduced in our article, Methods, we use some thresholds to avoid unreasonable labelling. For example, we assume that a valid Home label should account for at least 20% of user's total visits and at least 30% of the user's total stay duration. Additionally, the number of days that a valid Work label is visited should account for at least 40% of the total number of days of the user's travelling records.

The mobility framework determines individual travelling parameters n_w, β_1, β_2 . The weekly home-based tour number (n_w) is obtained by calculating the weekly average number of trips originating from Home for each user. Dwell rate (β_1) and burst rate (β_2) are obtained by finding the optimal parameter set that gives the minimum value of the following loss function¹:

$$L(\beta_1, \beta_2) = \int |P_G(\Delta t) - P_S(\Delta t|\beta_1, \beta_2)|d\Delta t + \eta|\bar{N}_G - \bar{N}_S(\beta_1, \beta_2)|, \quad (1)$$

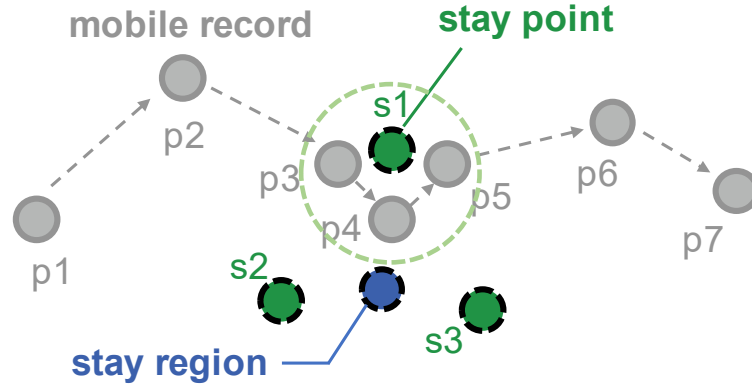
where $P_G(\Delta t)$ represents the distributions of individual empirical stay duration, and $P_S(\Delta t|\beta_1, \beta_2)$ is the simulated stay duration distribution given by the value of β_1, β_2 . \bar{N}_G and $\bar{N}_S(\beta_1, \beta_2)$ are the average daily number of visited locations of a user's empirical data and from the model simulation for a given parameter combination of β_1, β_2 . The parameter η is set to 0.05 to control the weight between the two losses. Since $L(\beta_1, \beta_2)$ is non-convex, we use a discrete set of β_1 and β_2 to estimate the optimal combination (β_1, β_2) that minimizes $L(\beta_1, \beta_2)$ for each user. Specifically, β_1 is set to 1, 2, ..., 20 and β_2 is set to 1, 6, ..., 101. Distributions of n_w, β_1, β_2 extracted from 2022 Beijing P-PEV Trajectories Dataset and 2022 Shenzhen P-PEV Trajectories Dataset are shown in Supplementary Fig. 2, Supplementary Fig. 3 and Supplementary Table. 1.

The mobility framework determines weekly average travel circadian rhythm $P(t)$ to make probabilistic travelling decisions. $P(t)$ is extracted from all P-PEV users in the ground truth travel records, and is further divided into commuter version (for users with Work labels) and non-commuter version (for users without ones). The patterns of $P(t)$ extracted from 2022 Beijing P-PEV Trajectories Dataset and 2022 Shenzhen P-PEV Trajectories Dataset are shown in Supplementary Fig. 4.

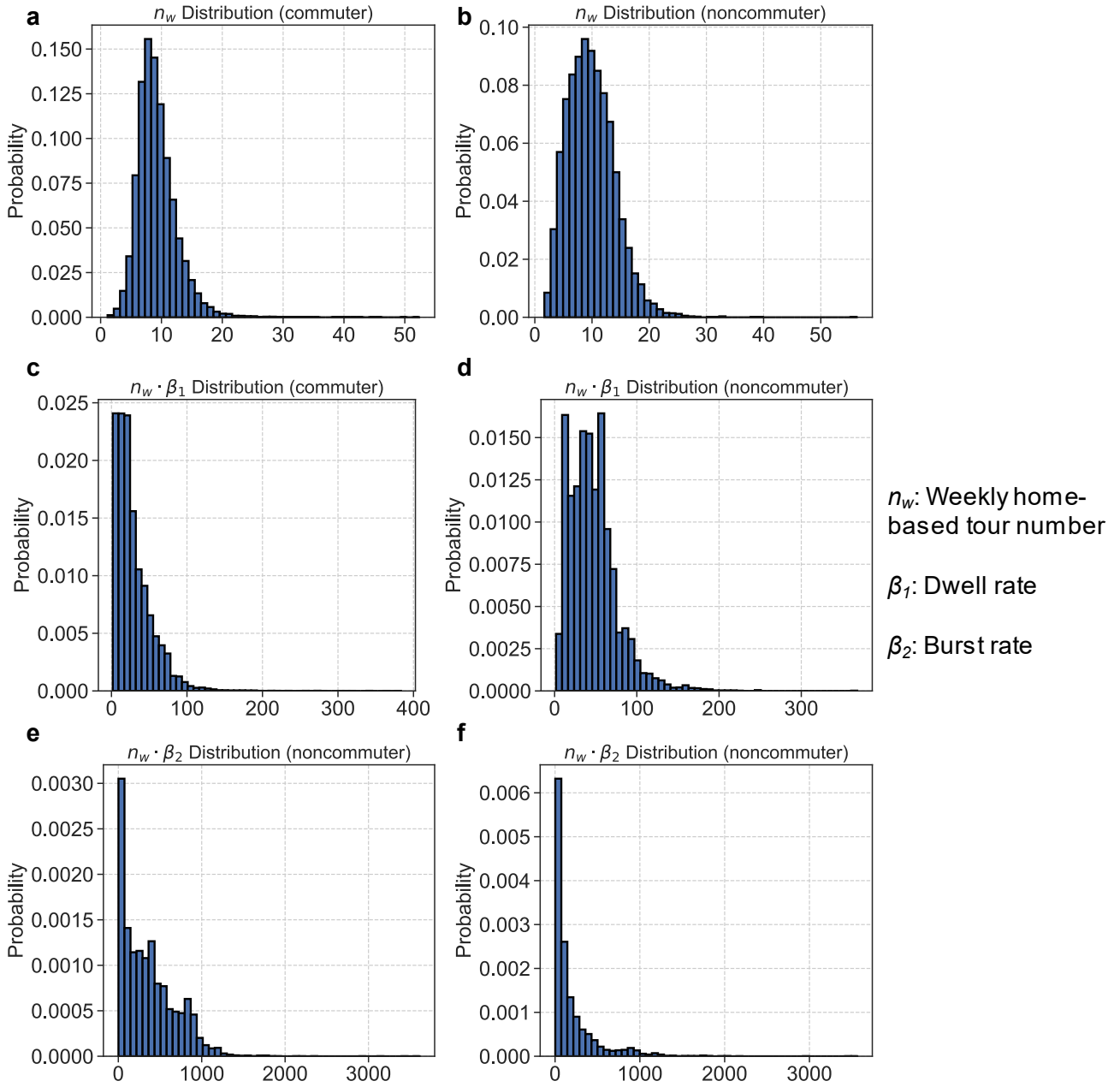
The L-th most visited location shown in the main text, Fig. 3d is explained as follows. For a P-PEV i , the mobility model generates M Stay-s in T time slots as introduced above. Each Stay has an index $I_{s1}, I_{s2}, \dots, I_{sM}$ indicating which stay region the Stay is located in. Intuitively, some Stay-s in specific stay regions such as Home, Work may be visited multiple times in the given time slots. Therefore, the cardinality of the set of visited stay regions $|\{I_{sm}\}|$ satisfies $|\{I_{sm}\}| \leq M$. We sort the number of visits of each stay region $V_{I_{st}}, I_{st} \in \{I_{sm}\}$ in a descending order. The L-th most visited location is introduced to describe the visit frequency of sorted $V_{I_{st}}$ of P-PEV i in T time slots. An example of L-th most visited location of a single user is provided in Supplementary Table. 2. Considering that the range of L may vary for different users depending on their travelling patterns, we calculate the summed normalized visit frequency of all P-PEV users in the article, main text, Fig. 3.

Since 2022 BPTD and 2022 SPTD only contain 30,000-40,000 P-PEV users, which are not comparable to actual Beijing and Shenzhen P-PEV holdings, we consider to expand the number of P-PEVs by considering the parking spot distributions of the cities. Specifically, we allocate the total P-PEV number to all streets based on their parking spot capacity. For Shenzhen, suppose a street i has N_p^i parking spots given by 2019 SPCD, and has N_e P-PEVs with their

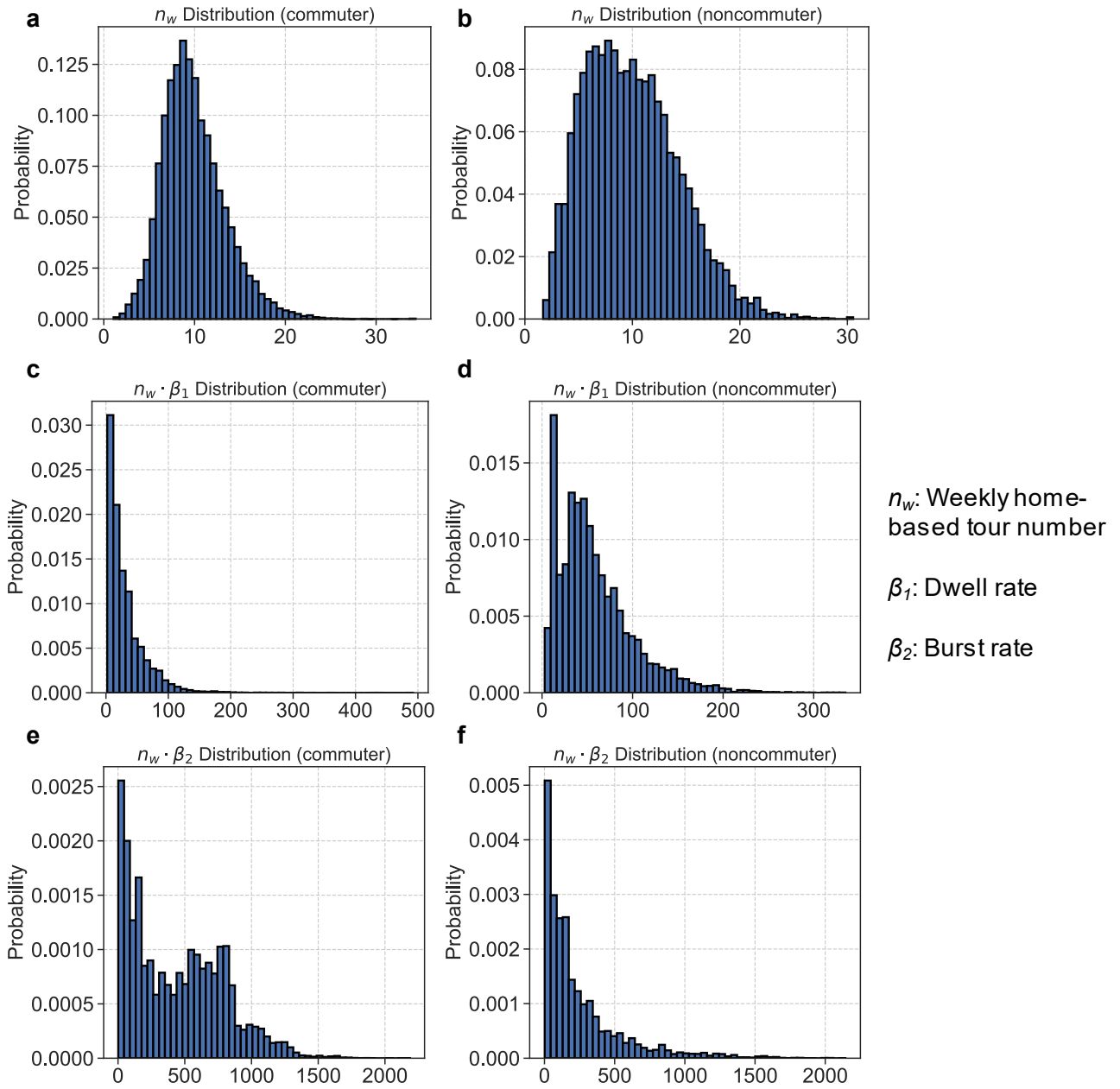
Home labels located within the street given by 2022 SPTD. We then expand the number N_e to $N_e * 480000 * \frac{N_p^i}{\sum_i N_p^l}$. Expanding the number to other values is similar. For Beijing, we do not have a fine-grained parking lot dataset like 2019 SPCD. Therefore, we refer to the statistics of parking lots in Beijing provided by Beijing Municipal Commission of Transport². The district-level numbers of parking lots are provided in Supplementary Table. 4.



Supplementary Fig. 1: Extraction of stays¹. Trajectories of P-PEV users consist of multiple mobile records. A stay point is identified from temporal consecutive and spatial adjacent mobile records. A stay region contains multiple neighboring stay points.



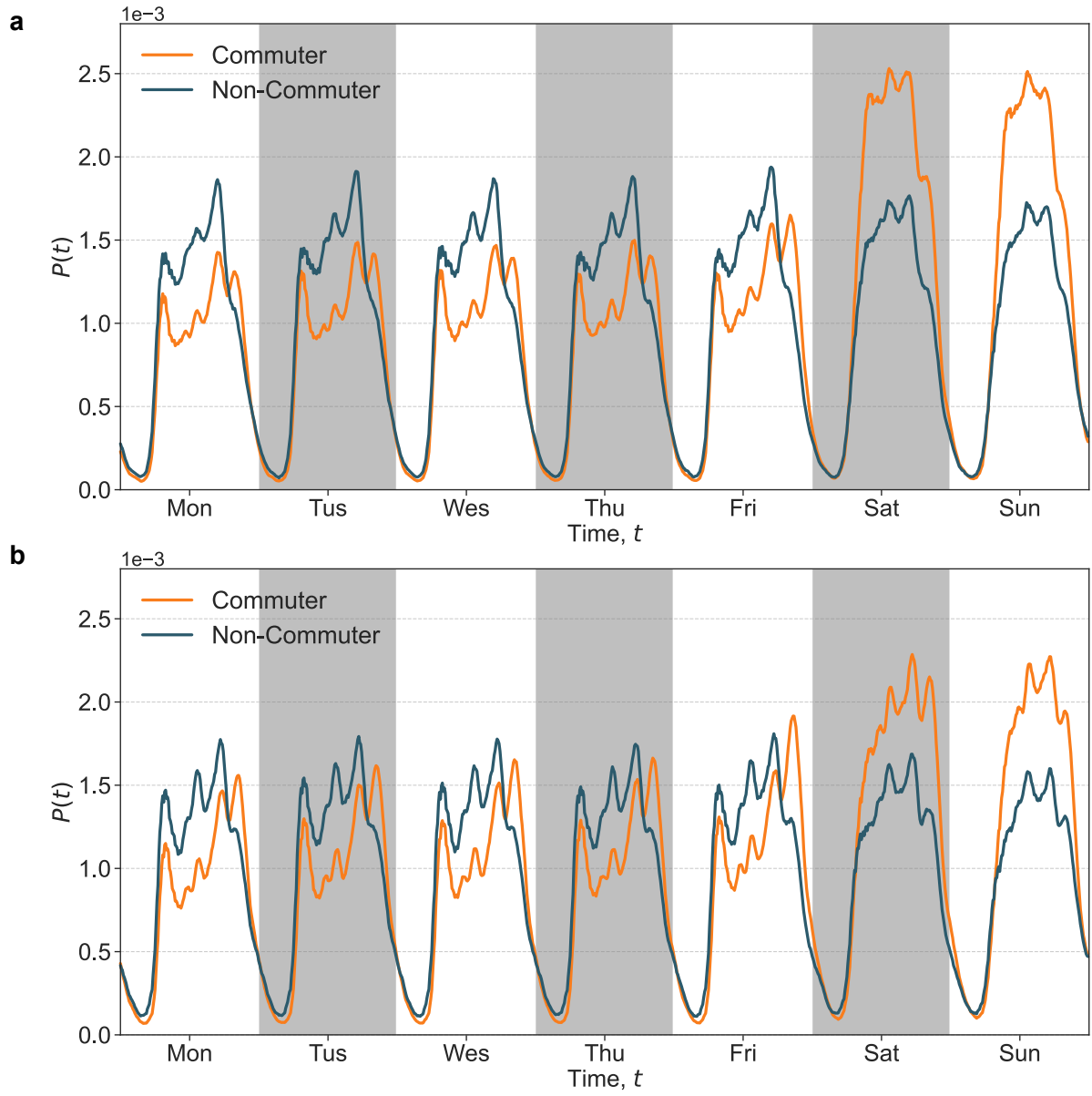
Supplementary Fig. 2: Distributions of n_w, β_1, β_2 from 2022 BPTD. **a**, n_w distribution of commuters. **b**, n_w distribution of noncommuters. **c**, $n_w \cdot \beta_1$ distribution of commuters. **d**, $n_w \cdot \beta_1$ distribution of noncommuters. **e**, $n_w \cdot \beta_2$ distribution of commuters. **f**, $n_w \cdot \beta_2$ distribution of noncommuters.



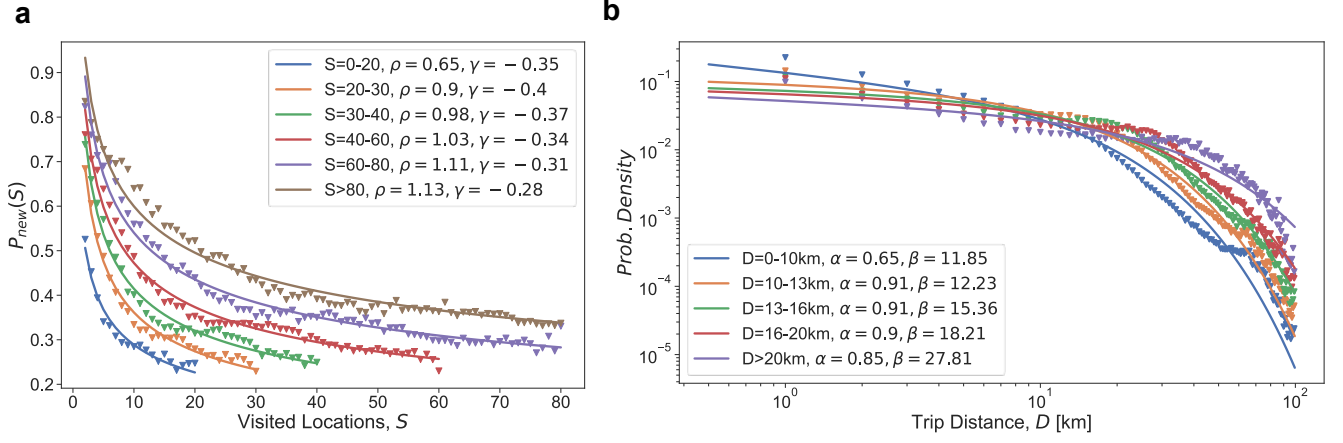
Supplementary Fig. 3: Distributions of n_w, β_1, β_2 from 2022 SPTD. **a**, n_w distribution of commuters. **b**, n_w distribution of noncommuters. **c**, $n_w \cdot \beta_1$ distribution of commuters. **d**, $n_w \cdot \beta_1$ distribution of noncommuters. **e**, $n_w \cdot \beta_2$ distribution of commuters. **f**, $n_w \cdot \beta_2$ distribution of noncommuters.

Supplementary Table. 1: Comparison of median n_w , $n_w \cdot \beta_1$, $n_w \cdot \beta_2$ between 2022 BPTD and 2022 SPTD.

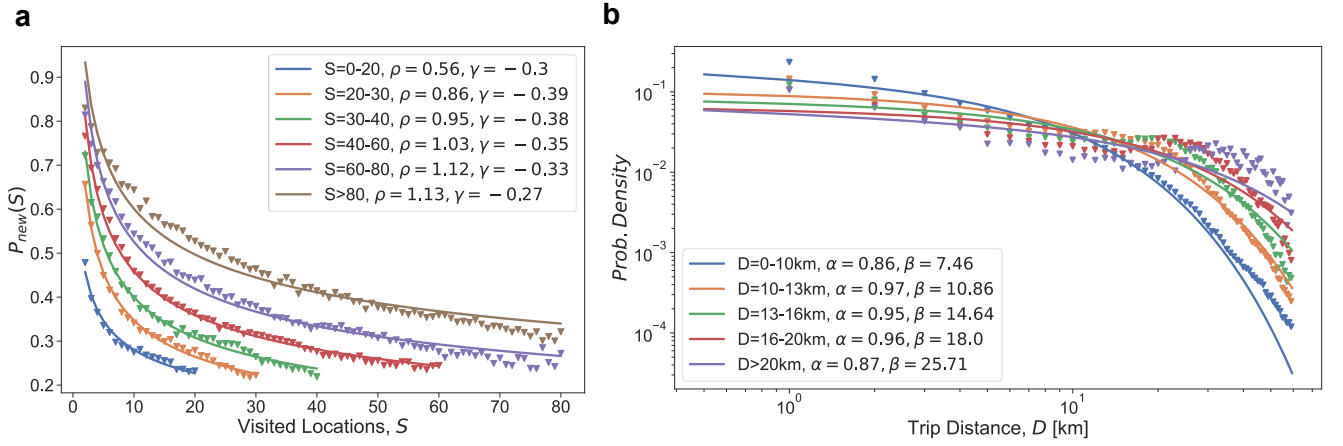
Median values	2022 BPTD	2022 SPTD
n_w (commuter)	8.77	9.45
n_w (noncommuter)	9.54	9.52
$n_w \cdot \beta_1$ (commuter)	22.18	20.78
$n_w \cdot \beta_1$ (noncommuter)	42.91	48.80
$n_w \cdot \beta_2$ (commuter)	302.00	408.36
$n_w \cdot \beta_2$ (noncommuter)	90.22	146.13



Supplementary Fig. 4: Circadian travel rhythms $P(t)$ of Beijing and Shenzhen P-PEV users. **a, $P(t)$ for commuters and noncommuters of 2022 BPTD. **b**, $P(t)$ for commuters and noncommuters of 2022 SPTD.**



Supplementary Fig. 5: Grouped spatial parameters of P-PEVs in Beijing. According to various travelling patterns of P-PEVs, we divide spatial parameters into several groups. Specifically, we fit different values of ρ, γ for $P_{new} = \rho S^{-\gamma}$, and the ones of α, β for $gamma(d) = \frac{\beta^\alpha}{\Gamma(\alpha)} d^{\alpha-1} e^{-\beta d}$. **a**, Distributions of visited locations, and the corresponding fitted parameters ρ, γ . **b**, Distributions of trip distance, and the corresponding fitted parameters α, β .



Supplementary Fig. 6: Grouped spatial parameters of P-PEVs in Shenzhen. The extraction of spatial parameters is identical to Supplementary Fig. 5. **a**, Distributions of visited locations, and the corresponding fitted parameters ρ, γ . **b**, Distributions of trip distance, and the corresponding fitted parameters α, β .

Supplementary Table. 2: Example of L-th most visited location for a given P-PEV user.

L	Stay region index	Visit frequency
1	2540	57
2	1224	40
3	2301	17
4	2	6
5	561	2
6	714	2
7	1363	2
8	1455	2
9	34	1
10	168	1
11	573	1
12	730	1
13	1090	1
14	1185	1

Supplementary Table. 3: Comparison of average travel distance between 2022 BPTD and 2022 SPTD.

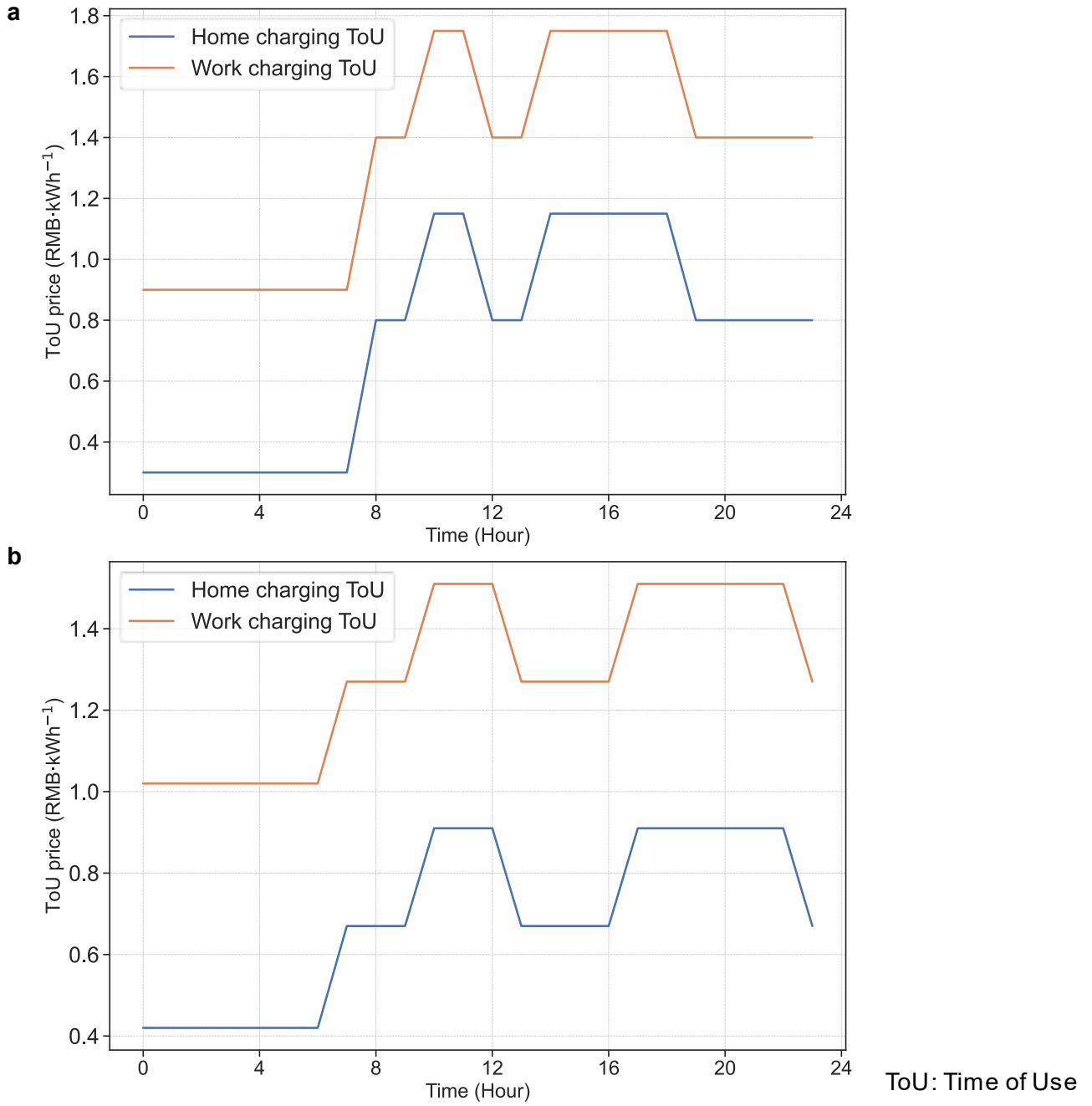
Average values	2022 BPTD	2022 SPTD
Travel distance per day (km)	53.26	40.35
Travel distance per time (km)	16.34	11.76

Supplementary Table. 4: Numbers of parking lots in each district of Beijing.

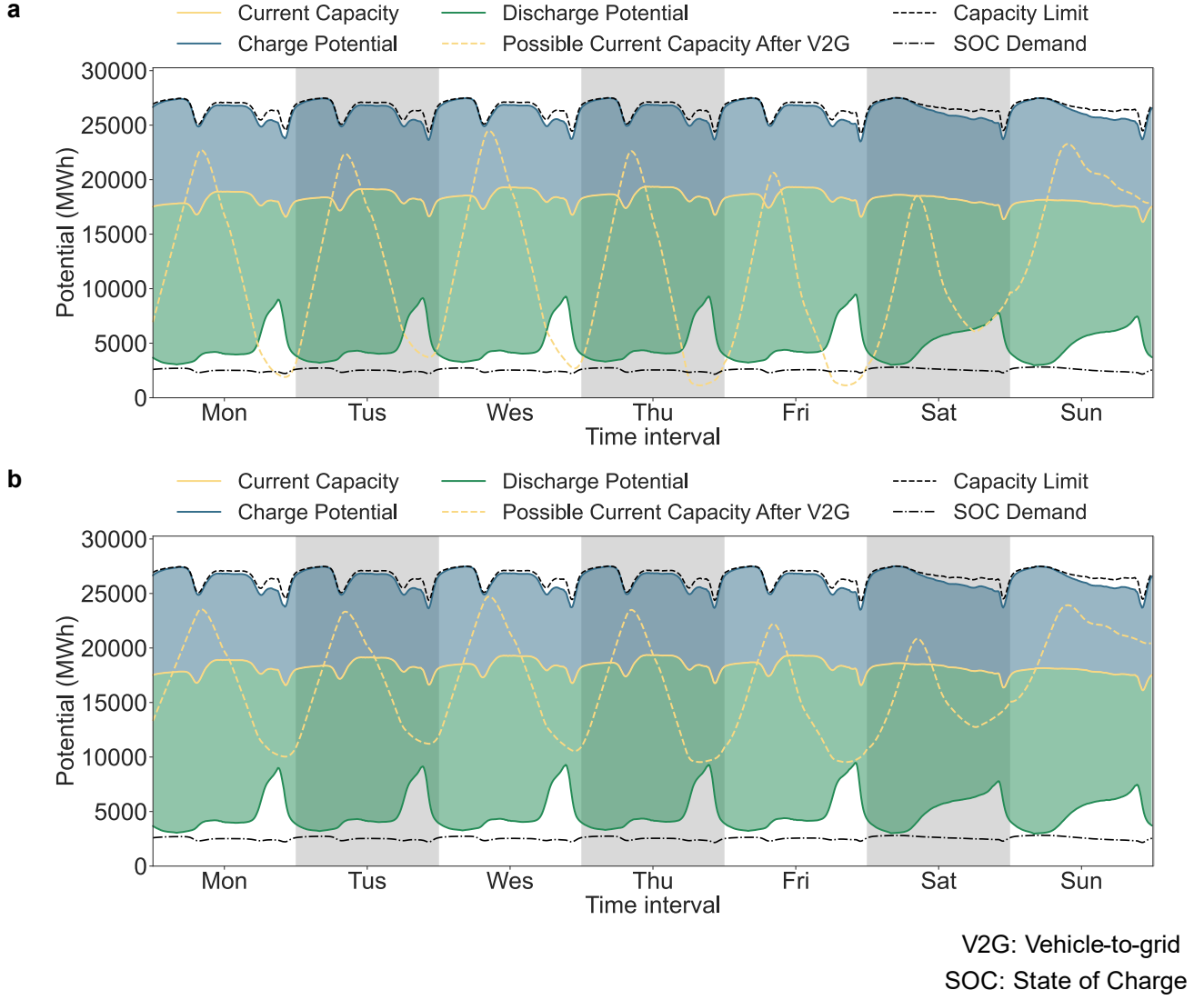
District	Number of parking lots
Chaoyang	628
Haidian	397
Dongcheng	265
Fengtai	254
Xicheng	234
Tongzhou	102
Shijingshan	100
Shunyi	96
Daxing	89
Fangshan	80
Yizhuang	79
Mentougou	58
Lvping	49
Yanqing	44
Huairou	39
Miyun	35
Pinggu	30

Supplementary Table. 5: Comparison of charging behavior parameters among different assumed cases.

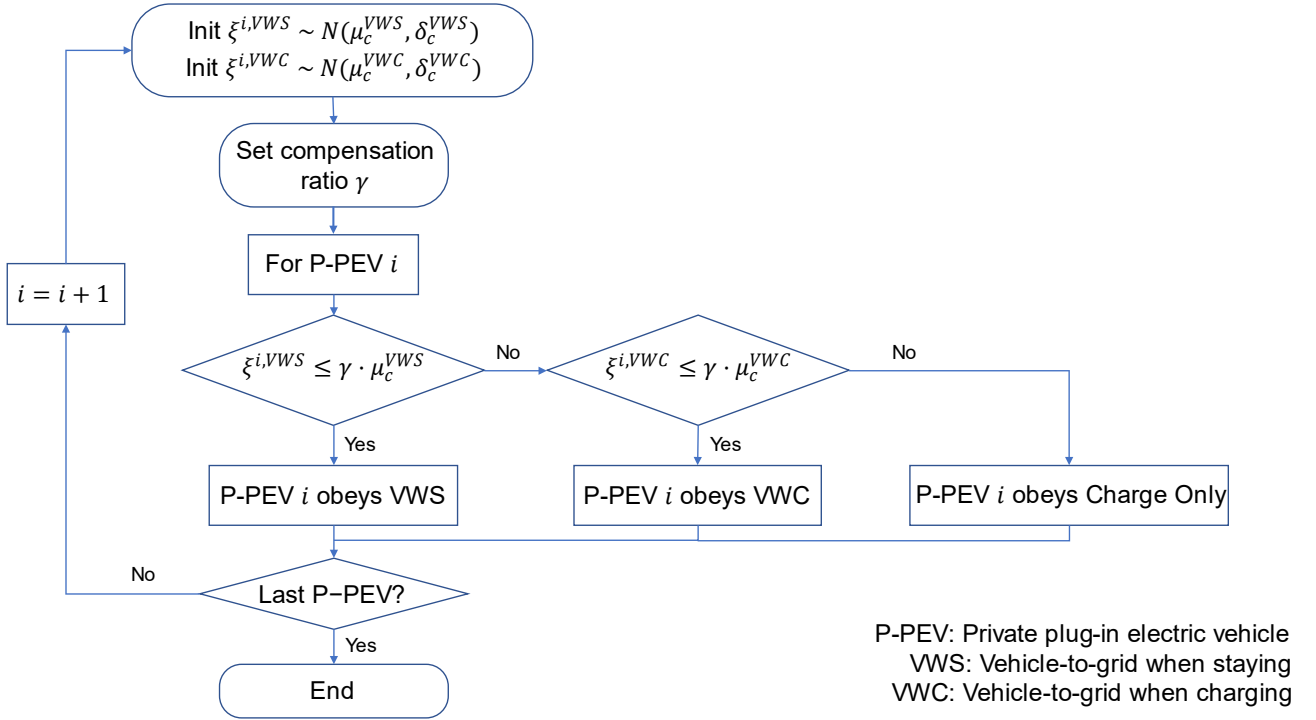
Parameter	Base case	Case 1	Case 2	Case 3	Case 4	Case 5
β_0	1	1	1	1	1	1
β_{SOC}	3	1	10	3	3	3
β_R	0	0	0	0.06	0	0
$\beta_{\Delta SOC}$	3	3	3	3	3	3
β_{Cost}	0.35	0.35	0.35	0.35	0.35	0
β_{Stay}	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2
$R_{Lcm,t}^{ToU, pub} - P_{Lcm,t}^{ToU, home}$	0.6	0.6	0.6	0.6	0	0.6



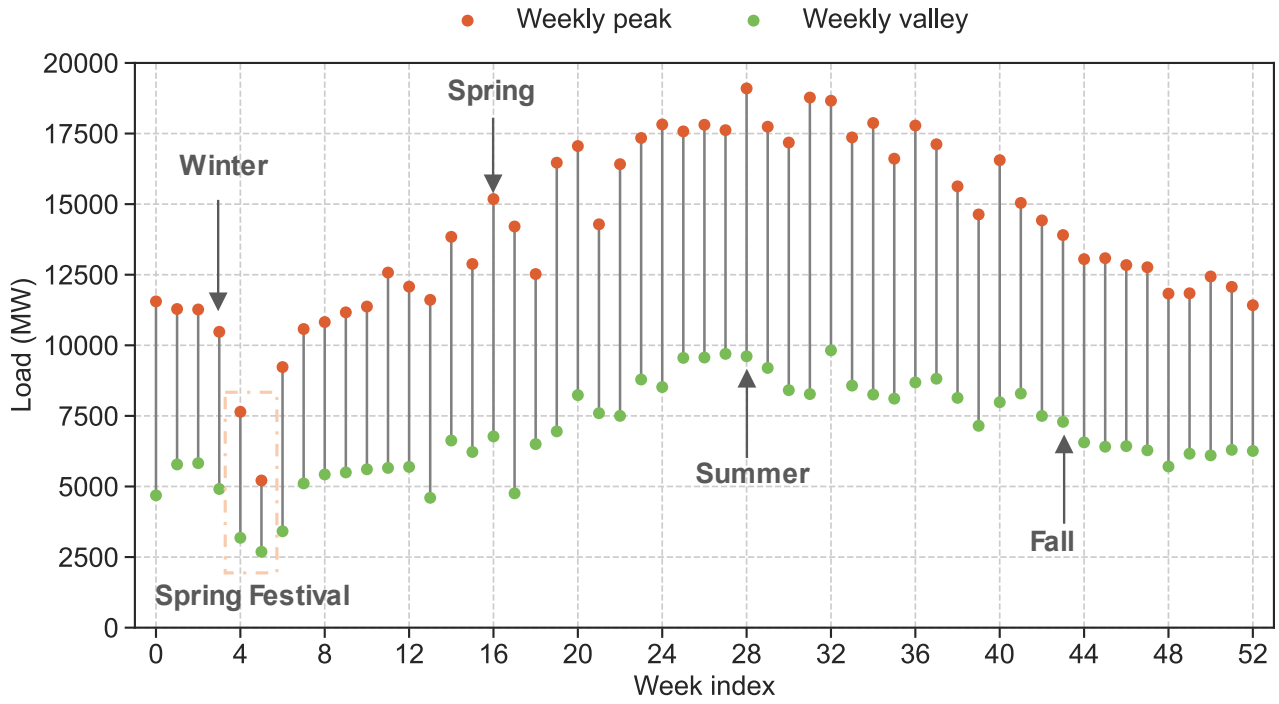
Supplementary Fig. 7: ToU information of Shenzhen and Beijing. The information is sourced from charging service providers³. **a**, ToU information of Shenzhen. The price difference between public and private charging piles ($R_{L_{cm},t}^{ToU, pub} - P_{L_{cm},t}^{ToU, home}$) is set as 0.6 RMB per kWh. **b**, ToU information of Beijing. $R_{L_{cm},t}^{ToU, pub} - P_{L_{cm},t}^{ToU, home}$ is set as 0.55 RMB per kWh.



Supplementary Fig. 8: Explanations of charging potential. The possible current capacity after V2G may exceed the range of discharge potential and SOC demand. The reasons are twofold. First, SOC demand $DP_{i,s}$ defined in Equation (39) of the manuscript, $DP_{i,s} = \max(SOC_{i,s}^{need}, SOC_{i,s}^e - P_i * T_{i,s})$, represents a rough estimate that remains constant throughout Stay s . However, our V2G optimization model only requires the SOC at departure time to satisfy $SOC_{i,s}^{need}$, as specified in Equation (24)(30). Second, the charging/discharging potential is based on the SOC information given by Charge Only. This assessment does not account for V2G. When the V2G optimization model is introduced, the SOC of P-PEV i at time slot t becomes variable, determined by the model's objective function (Equation (15) in the manuscript). Consequently, the actual values of charging/discharging potential at subsequent time slots will deviate from the initial values of Charge Only. To prevent potential misinterpretation, we modify the values of the yellow dashed line in the manuscript, Fig. 4d. We emphasize that these results serve solely as a schematic illustration and may not precisely reflect the actual optimization results. **a**, An actual result of possible current capacity after V2G. **b**, A modified result of possible current capacity after V2G, as shown in the manuscript, Fig. 4d.



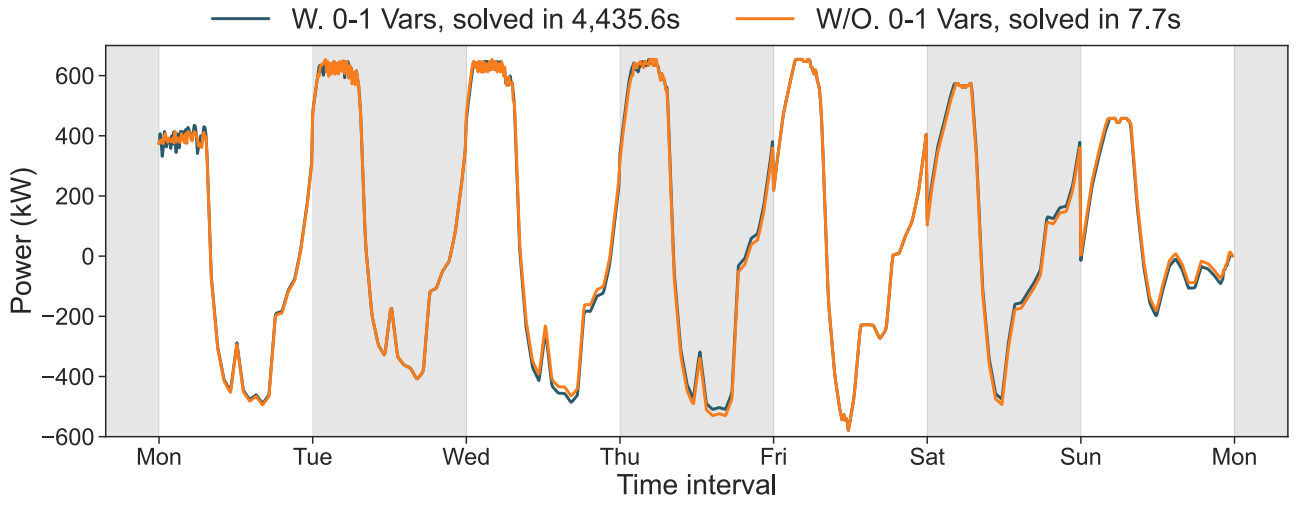
Supplementary Fig. 9: Flowchart of the V2G participation mechanism. For each P-PEV user, we randomly generate their desired compensation according to the normal distribution given in the manuscript, Equation (12, 13). For the two V2G scenarios, we have $\xi^{i,VWS} \sim N(\mu_c^{VWS}, \delta_c^{VWS})$ and $\xi^{i,VWC} \sim N(\mu_c^{VWC}, \delta_c^{VWC})$. Given a specific compensation rate γ , we first select users willing to participate in V2G When Staying. The unselected users further decide whether to join V2G When Charging according to their values of $\xi^{i,VWC}$. The last remaining users will not participate in either V2G scenarios.



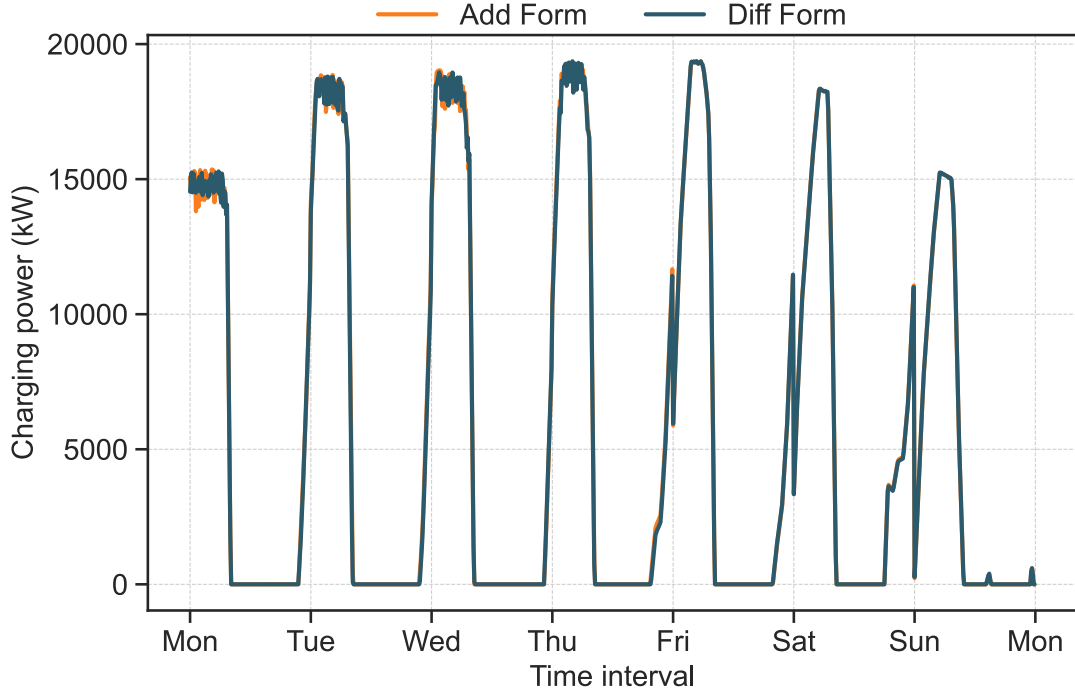
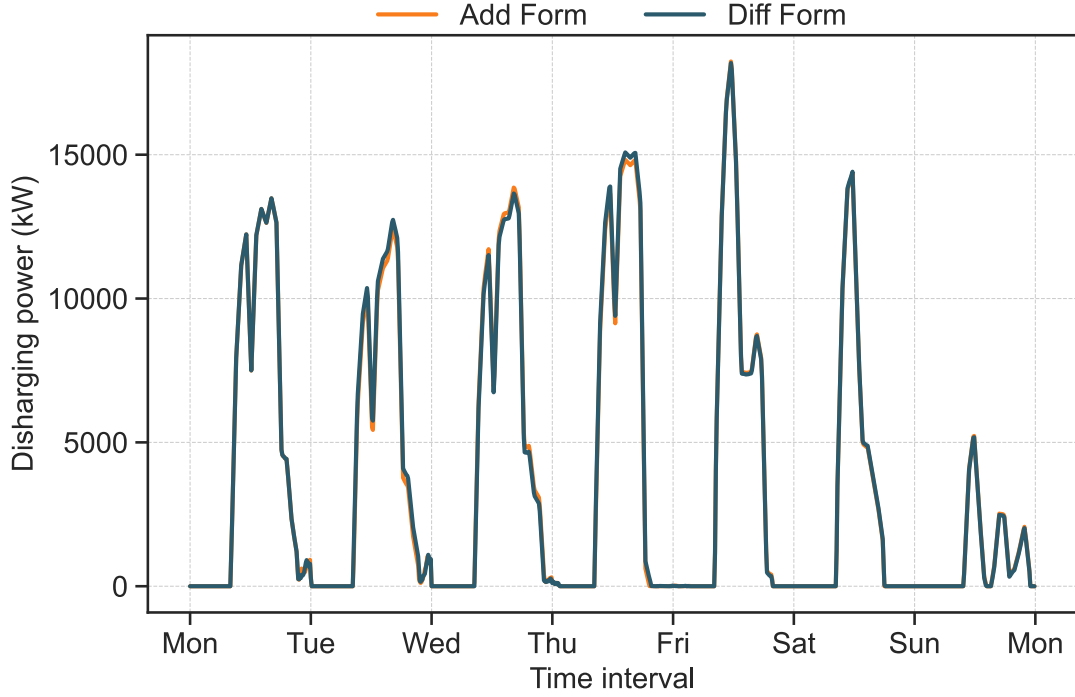
Supplementary Fig. 10: Selection of seasonal urban loads of Shenzhen. Since we do not have the permission to share the original load data provided by China Southern Power Grid, we illustrate the weekly peak and valley value throughout the year. We avoid weeks related to Chinese public holidays. The summer week has the annual peak value. The peak value of the selected fall week is the median of the peak values of 53 weeks. The spring week has similar patterns compared to the fall week but has greater fluctuations. The winter week represents an overall lower load of the year.

Supplementary Table. 6: China top-10 PEV market share data in 2019 (Estimated)

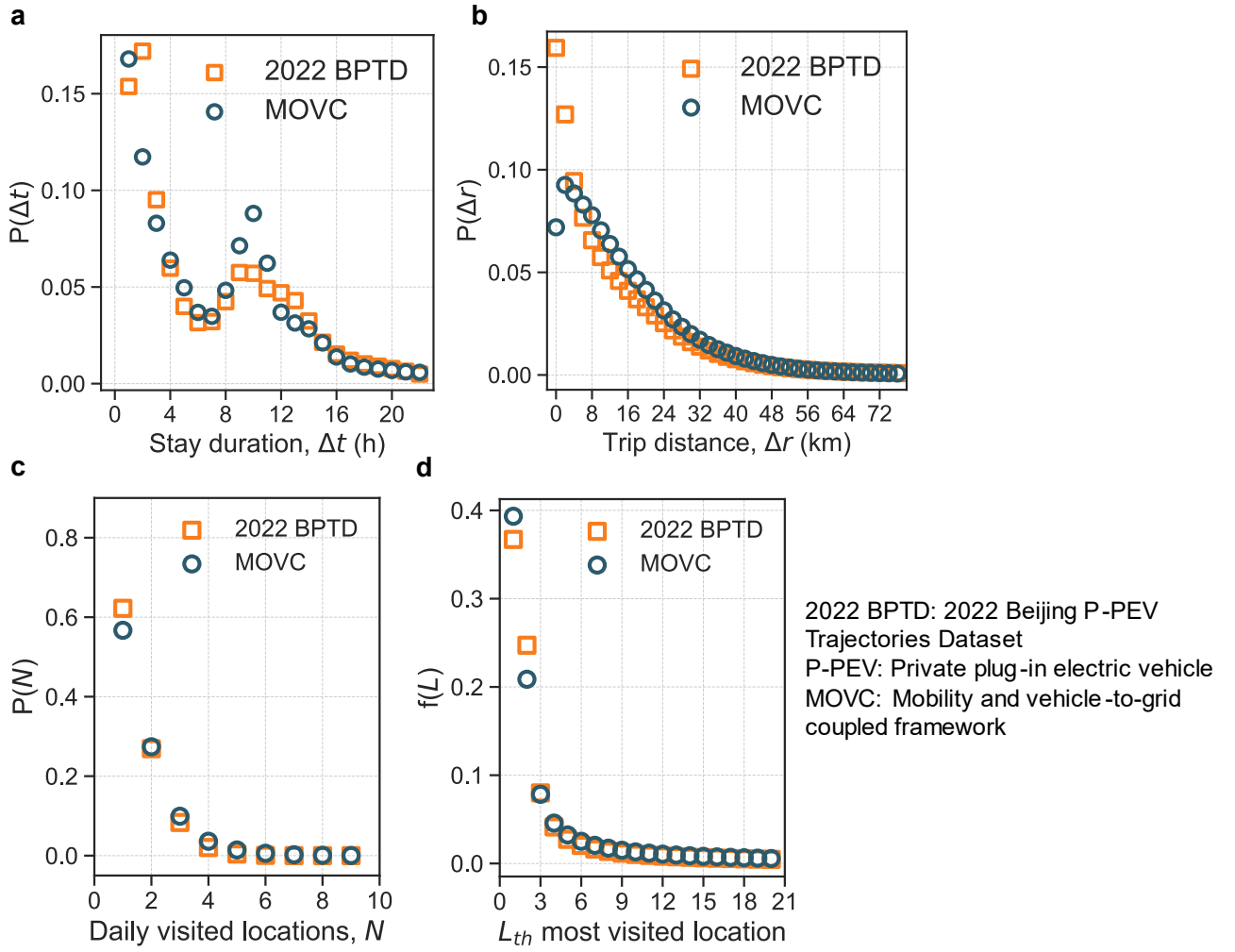
EV Model	Market share (%)	Power (kW)	QC Power (kW)	Capacity (kWh)
Tesla Model 3	29.23	7.0	125.0	78
BYD Qin	13.10	7.0	17.5	53
Aion S	11.66	5.0	15.0	58
Baojun E100	8.63	7.0	17.5	28
BAIC EU	7.51	7.0	17.5	45
NIO ES6	7.51	7.0	100.0	75
Chery eQ	6.23	5.0	12.5	35
ORA R1	5.91	5.0	25.0	33
BYD Yuan	5.43	7.0	40.0	43
WM EX5	4.79	7.0	40.0	52



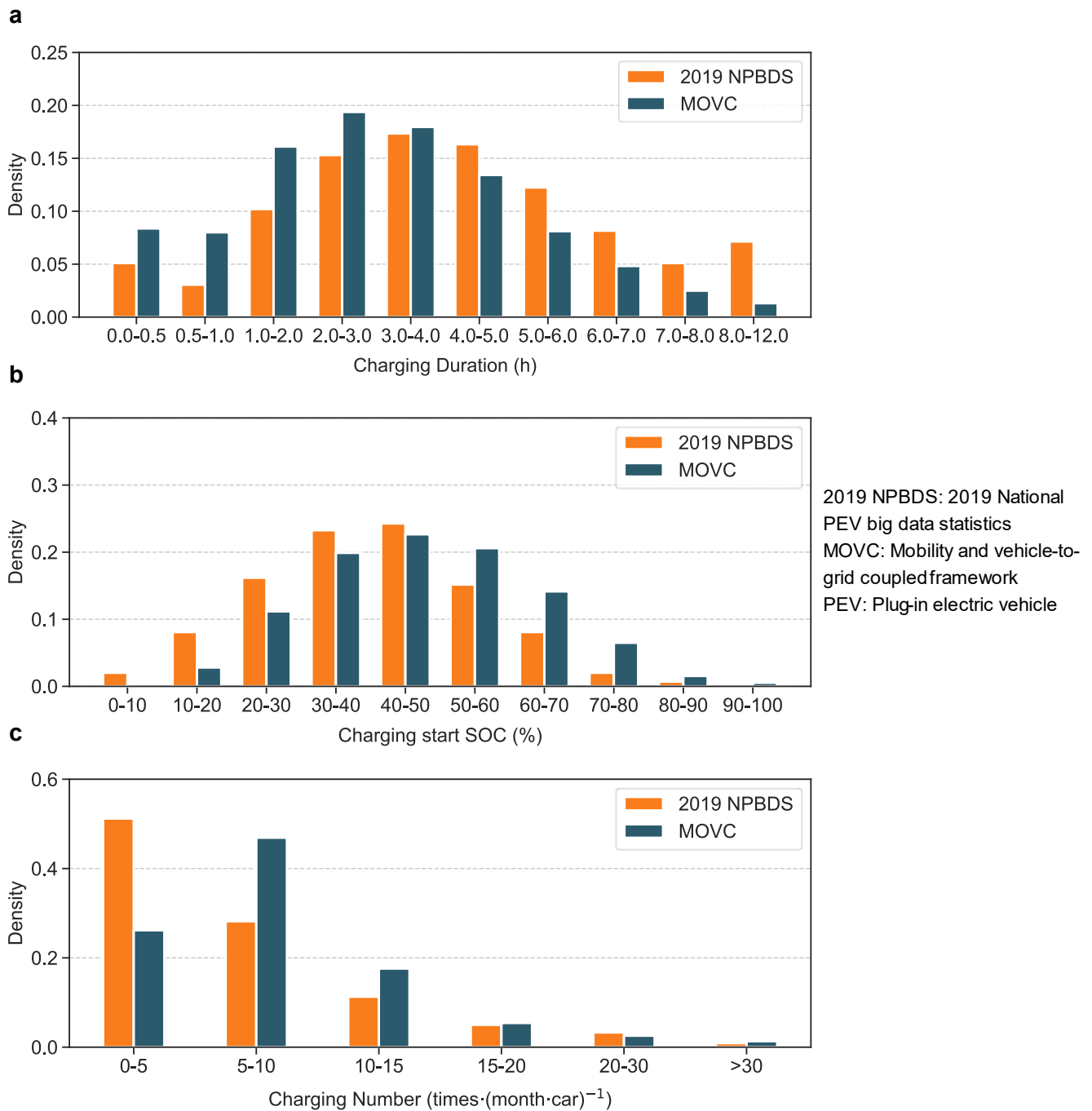
Supplementary Fig. 11: Comparison of results and computational costs between the mixed-integer linear programming (MILP) optimization and the linear programming (LP) optimization. As introduced in Methods in the main body of the paper, to balance between accuracy and efficiency, we design an LP optimization model with a appropriately designed objective function to avoid simultaneous charging and discharging (SCD). In this figure we optimize 100 P-PEVs and compare the optimized patterns of their power flows between our LP model and an MILP version with 0-1 variables added in the constraints, since empirically 100 is the maximum capacity to solve the problem with MILP in our platform. The two models give very close results, however the MILP model takes 4435.6 seconds to solve, while the LP model only takes 7.7 seconds.

a**b**

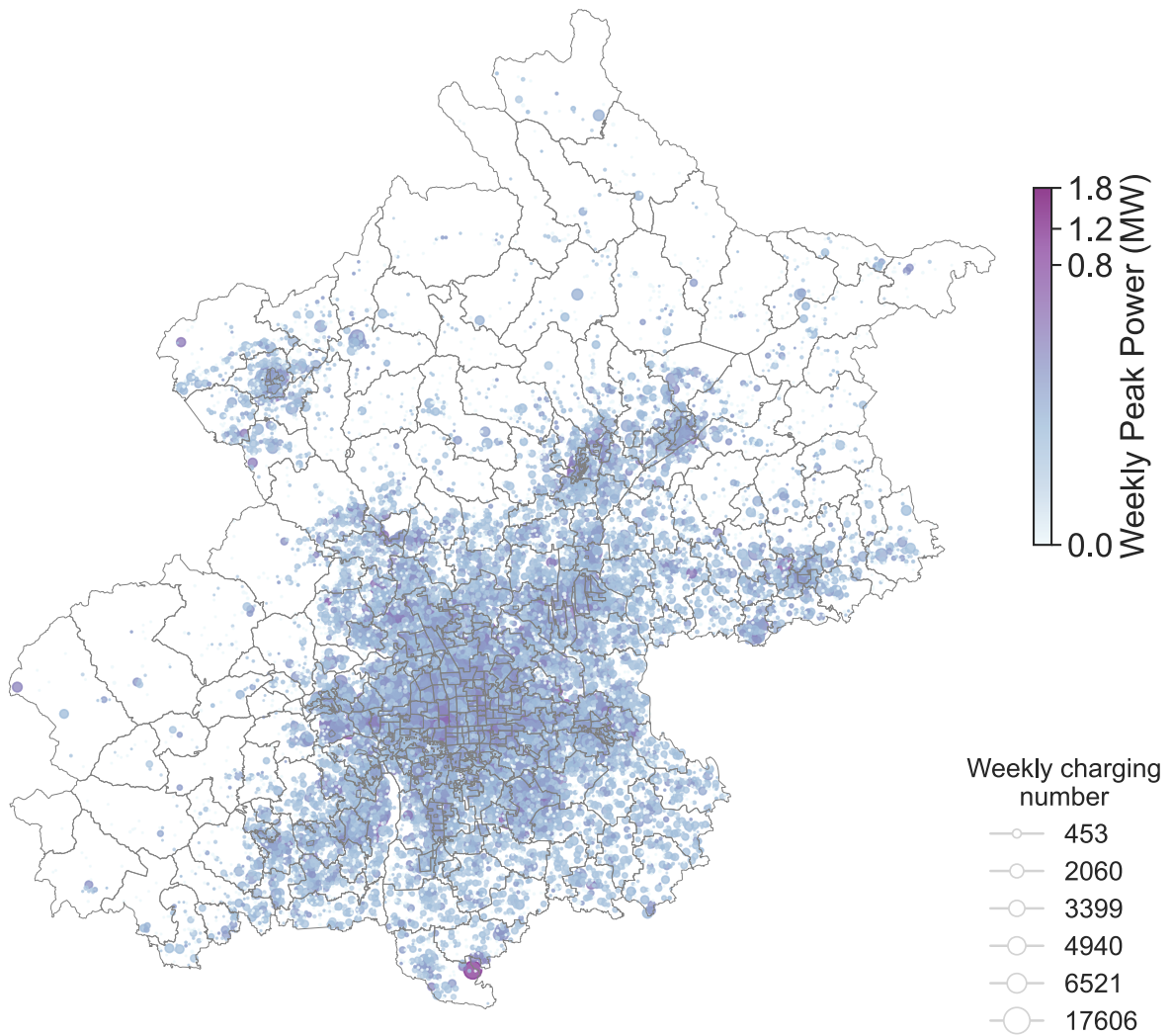
Supplementary Fig. 12: Comparison of different forms of f_2 to eliminate SCD. The second objective term f_2 presented in the manuscript, Equation (15) is introduced to eliminate SCD. We use $\lambda f_2 = \lambda \frac{1}{T} \sum_{i=1}^N \sum_{t=1}^T (P_{i,t}^{ch} + P_{i,t}^{dis})$ in our manuscript⁴⁻⁶, and we set $\lambda = 0.15$. The form is dubbed as ‘Add Form’. Another form of difference $\lambda' f_2' = \lambda' \frac{1}{T} \sum_{i=1}^N \sum_{t=1}^T (P_{i,t}^{ch} - P_{i,t}^{dis})$, dubbed as ‘Diff Form’, can also empirically yield almost identical results. **a**, Aggregated charging power of 3,000 P-PEVs. **b**, Aggregated discharging power of 3,000 P-PEVs.



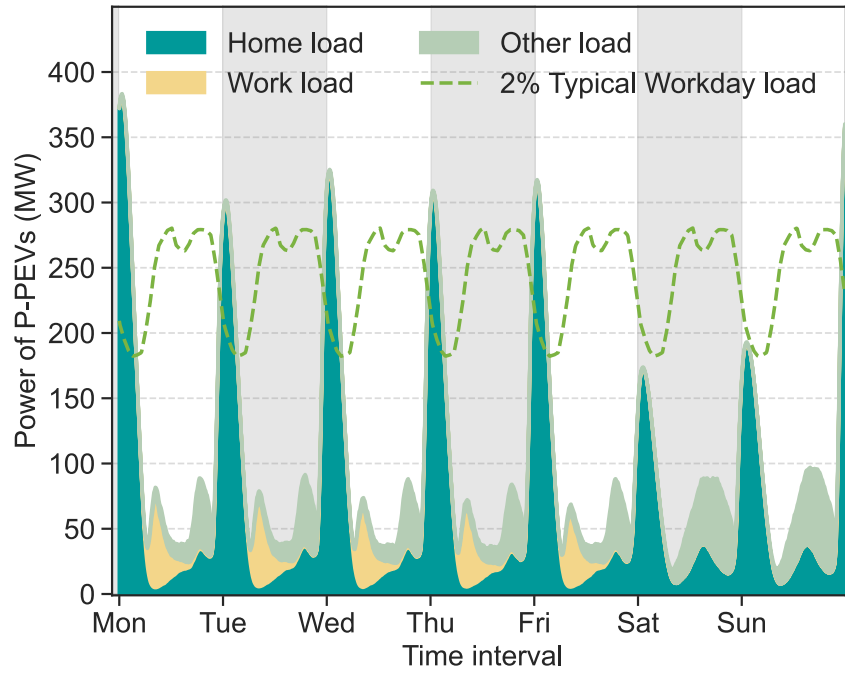
Supplementary Fig. 13: Key mobility characteristics extracted from 2022 Beijing P-PEV Trajectories Dataset (2022 BPTD). **a**, Probability distributions of single stay duration. **b**, Probability distributions of single trip distance. **c**, Probability distributions of the number of daily visited locations. **d**, Probability distributions of visiting the L -th most visited location.



Supplementary Fig. 14: Validation of charging behavior simulation results of 2022 BPTD under Charge Only. **a**, Probability distributions of charging duration. **b**, Probability distributions of charging start SOC. **c**, Probability distributions of monthly charging number.



Supplementary Fig. 15: Weekly peak load distribution of P-PEVs in Beijing, divided into stay regions. To make the results of Beijing comparable to Shenzhen as illustrated in the manuscript, Fig. 4a, we simulate 400,000 P-PEVs in Beijing, which is approximately the Beijing P-PEV holding in 2022.



Supplementary Fig. 16: The load curve of Beijing P-PEVs, divided into charging places. The parameters ($\beta_0, \beta_{SOC}, \beta_R, \beta_{\Delta SOC}, \beta_{Cost}, \beta_{Stay}$) for simulating charging results of Beijing are exactly the same as Shenzhen's case, as provided in Supplementary Table. 5.

Supplementary Table. 7: Comparison of constraints applied according to user actions in different scenarios.

Constraints \ Action \ Scenario	Park	Charge	QC
V2G When Charging	Equation (16-18)	Equation (19-24)	Equation (25-30)
V2G When Staying (Default)	Equation (19-24)	Equation (19-24)	Equation (25-30)
V2G When Staying (QC 100%)	Equation (25-30)	Equation (19-24)	Equation (25-30)

Supplementary References

1. Jiang, S. *et al.* The TimeGeo modeling framework for urban mobility without travel surveys. *Proceedings of the National Academy of Sciences* **113**, E5370-E5378 (2016).
2. Zhiyan Consulting. *Analysis of Beijing's parking facility construction in 2020*, <<https://www.chyxx.com/industry/202109/976573.html>> (2021).
3. Longshine Technology Group Co., L. *Xindiantu - Aggregated charging service platform*, <<https://www.longshine.com/solutionInfo/81>> (2024).
4. Yang, P. & Nehorai, A. Joint optimization of hybrid energy storage and generation capacity with renewable energy. *IEEE Transactions on Smart Grid* **5**, 1566-1574 (2014).
5. Li, Q. & Vittal, V. Non-iterative enhanced SDP relaxations for optimal scheduling of distributed energy storage in distribution systems. *IEEE Transactions on Power Systems* **32**, 1721-1732 (2016).
6. Yang, S., Gao, H. O. & You, F. Integrated optimization in operations control and systems design for carbon emission reduction in building electrification with distributed energy resources. *Advances in Applied Energy* **12**, 100144 (2023).