

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

Corresponding Author: Professor Yi Zhang

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The manuscript titled "Unlocking the Potential of Private EVs for Shaving the Electricity Peak Load of the City" presents a Mobility and Vehicle-to-Grid Coupled (MOVC) framework to quantify and optimize the utilization of private electric vehicles (P-PEVs) for urban electricity load balancing. Leveraging extensive datasets and a linear programming optimization model, the study investigates the potential of V2G technology in a large-scale setting (Shenzhen, China) to reduce peak-valley differences in electricity loads. The authors' approach integrates high-resolution mobility patterns and charging/discharging behaviors to derive actionable policy insights.

While the study presents a methodologically rigorous framework, the core research question—leveraging EVs for load balancing—has been extensively studied in the literature. The manuscript's primary novelty lies in its fine-grained, city-specific analysis, which, while valuable, may limit its broader applicability and appeal to a global audience. Therefore, in its current format, I do not find the insights provided in the study sufficient for publication in Nature Communications.

To be considered for publication, the authors should address the following issues:

- The concept of using EVs for peak shaving and load balancing is well-established, and while the granular modeling approach is noteworthy, the study does not introduce fundamentally new methods or paradigms. The authors should clearly explain if their proposed approach provides significant novelty and results in new general insights.
- The study's findings are specific to Shenzhen, and the manuscript does not adequately demonstrate the framework's transferability to other urban contexts. Broader applicability to cities with differing EV penetration rates, energy grids, or transportation systems should be explored.
- The charging behavior model assumes a probabilistic approach based on stay duration and state of charge (SOC), but it does not account for real-world behavioral factors, such as user cost sensitivity or range anxiety. Incorporating behavioral insights through surveys or literature-based models would enhance the realism of the framework.

(Remarks on code availability)

The code can be run as a "Reproducible Run" on codeocean.com. The data is also available.

However, there is no README file to give instructions on how to use the code. Including such a file can greatly increase the usability of the code.

Reviewer #2

(Remarks to the Author)

This paper aims to explore the potential of V2G for peak shaving by proposing a mobility and V2G coupled framework for charging and discharging strategies. The city of Shenzhen is taken as the case study. This study is valuable. The authors need to clarify the following issues:

1. On Line 87-89, the statements are incorrect. Many studies have existed regarding the assessment of load balancing/leveling, peak shaving, and valley filling, such as: "High efficient valley-filling strategy for centralized coordinated charging of large-scale electric vehicles", "A systematic methodology for mid-and-long term electric vehicle charging load forecasting: The case study of Shenzhen, China", "The peak load shaving assessment of developing a user-oriented vehicle-to-grid scheme with multiple operation modes: The case study of Shenzhen, China". The authors are suggested to summarize the shortcomings of existing studies regarding peak-load shaving.

2. The MOVC framework combines the Timegeo modeling framework in ref. (17) for urban mobility with the V2G optimization for charging and discharging strategies. It is a nice combination. My question is how to determine the probability for move or stay and for home or other places?

3. In Figure 1, there are the statements "car charges the building and building charges the car". I wonder if only the V2B is considered in this study.

4. In Figure 1 and Figure 2, the V2G action falls into three types: charge only, V2G when charging, and V2G when staying. In the V2G model, how do you distinguish these three scenarios for charging and discharging strategies?

5. In Figure 2, at the node "stay", with no charge, why is there V2G off or V2G on? When the charge occurs, there is also V2G off or V2G on. What is the difference between them?

6. The analysis of PEV load profiles is limited to two weeks representing summer and winter. How do these results account for the significant seasonal and temporal variations in electricity demand throughout the year? In addition, the number of PEVs would have a significant increase in the next years. The scale of PEVs should be considered for the result analysis.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

The manuscript "Unlocking the Potential of Private EVs for Shaving the Electricity Peak Load of the City" presents a study on mobility and V2G coordination to optimize the charging and discharging patterns of EVs in a mega-city in China. While the data analysis is interesting, the V2G strategy has some drawbacks, the manuscript is difficult to read, and several concepts are unclear. Based on these concerns, I do not recommend accepting the paper in its current form. Below are my specific comments:

Introduction: The identified research gap related to private V2G does not seem entirely accurate. Previous studies have already addressed this issue in various ways (e.g., 10.1109/OJVT.2023.3323087, 10.1109/ACCESS.2022.3185236, 10.1016/j.scs.2021.102744). Please refine the discussion on research gaps and clarify the contributions of your study.

Quick Charging (QC): The manuscript considers the use of quick charging, but the assumed power level is not specified. Please provide this information. Additionally, you state: "Assumed that all vehicles can perform QC on-site when needed, the number of QC activities takes possession of 10.34% of all charging events." This statement is unclear—please clarify the reasoning behind this assumption.

Variable Definitions & Notation: Several variables are not explicitly defined, including SOC_{in} , ΔD , and C_{uafter} . Additionally, after Figure 3, you mention the L -th most visited location, but its meaning is unclear. Does depot refer to charging stations? If so, please use consistent terminology. Moreover, the values of α (alpha) and β (beta) are not justified—please explain their origin.

Optimization Formulation: Why is the battery capacity not constrained? Additionally, in function f_2 , which is a sum of differences, please explain how this prevents simultaneous charging and discharging.

Figures & Data Presentation:

Provide numerical error values for Figures 3a, 3b, 3c, and 3d.

The computation method for Figure 4c is unclear. How is the charging potential considered? Does the charger not inject full power? Please provide a detailed explanation.

The scenarios are not clearly defined. Please introduce them explicitly before presenting the results.

V2G Considerations: The paper assumes that V2G increases energy, but how is the discharged energy accounted for? Are EVs allowed to discharge at any moment? Additionally, how is V2G implemented in the model? Please clarify aspects such as scheduling, user willingness, and battery degradation effects.

Data Validation: Is there a reference for the dataset used for validation? Please provide details on its source and reliability.

Manuscript Organization & Writing:

Improve the manuscript structure to ensure that figures appear in the order they are mentioned in the text.

Proofread the paper to correct typos. For example, on page 5, line 137, you reference Figures 1a and 1b, but these figures do not exist. Additionally, Table 1 does not present EV parameters, despite the statement in lines 389–392 on page 15.

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

I appreciate the authors' substantial effort in addressing my previous concerns through extensive revisions, and I recognize the significant improvements made to the manuscript. Specifically, the expanded scope to include multiple Chinese megacities (Beijing, Shanghai, Guangzhou, Shenzhen), the enhanced realism in the charging behavior model, and the clarified methodological novelty in terms of detailed, large-scale, user-centric modeling are commendable and reflect a meticulous and rigorous research process.

Nevertheless, after careful consideration, my primary concern remains regarding the level of novelty and the transformative insights provided by this manuscript. While the authors clearly demonstrate detailed modeling rigor, systematic analysis, and careful validation, the fundamental conceptual framework—leveraging electric vehicles (EVs) and vehicle-to-grid (V2G) technologies for urban grid load balancing—is well-established and extensively explored in existing literature. Consequently, the incremental insights provided, despite their depth, may not sufficiently meet the high-impact and broad audience criteria typically expected for publication in Nature Communications.

Therefore, I regretfully conclude that the manuscript, though methodologically sound and significantly improved, does not offer the necessary level of general or impactful insights for this particular journal.

However, the quality and detailed analysis presented would certainly be well-received in specialized or topical journals on energy or transportation. Hence, I strongly encourage the authors to consider submitting this refined manuscript to journals.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

Thanks for the efforts of the authors to revise this manuscript. All my comments have been well addressed.

(Remarks on code availability)

permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Responses to Reviewer Comments

Title	Unlocking the potential of private EVs for shaving the electricity peak load of the city
Revised Title	Unlocking vehicle-to-grid potential of load shifting in China's megacities considering comprehensive real-world behaviors
Authors	Kaisan Li, Xinxin Li, Gucheng Zhao, Zuxun Xiong, Yi Jiang, He Qi, and Yi Zhang
Revised Authors	Kaisan Li, Xinxin Li, Zuxun Xiong, Shengyu Tao, Gucheng Zhao, Yi Jiang, He Qi, and Yi Zhang
Journal	Nature Communications
Manuscript ID	NCOMMS-24-79794A

Table of Contents

Response to Reviewers.....	3
Response to Reviewer #1	3
Response to Comment 1	3
Response to Comment 2	8
Response to Comment 3	19
Response to Comment 4	27
Response to Reviewer #2	29
Response to Comment 1	29
Response to Comment 2	30
Response to Comment 3	34
Response to Comment 4	36
Response to Comment 5	39
Response to Comment 6	41
Response to Reviewer #3	52
Response to Comment 1	52
Response to Comment 2	53
Response to Comment 3	57
Response to Comment 4	65
Response to Comment 5	69
Response to Comment 6	76
Response to Comment 7	92
Response to Comment 8	92
References	94

Response to Reviewers

Response to Reviewer #1

The manuscript titled "Unlocking the Potential of Private EVs for Shaving the Electricity Peak Load of the City" presents a Mobility and Vehicle-to-Grid Coupled (MOVC) framework to quantify and optimize the utilization of private electric vehicles (P-PEVs) for urban electricity load balancing. Leveraging extensive datasets and a linear programming optimization model, the study investigates the potential of V2G technology in a large-scale setting (Shenzhen, China) to reduce peak-valley differences in electricity loads. The authors' approach integrates high-resolution mobility patterns and charging/discharging behaviors to derive actionable policy insights. While the study presents a methodologically rigorous framework, the core research question—leveraging EVs for load balancing—has been extensively studied in the literature. The manuscript's primary novelty lies in its fine-grained, city-specific analysis, which, while valuable, may limit its broader applicability and appeal to a global audience. Therefore, in its current format, I do not find the insights provided in the study sufficient for publication in Nature Communications.

Dear Respected Reviewer,

Thank you very much for your time and valuable feedback on our work. We are delighted that you recognize the merits of our study in fine-grained city-level analysis. Your comments also highlight critical limitations regarding the generalizability and modeling realism of our approach, which we have carefully addressed in the following revisions:

- (a) Clarification of the study's novelty,
- (b) Additional experiments to validate the method's generalizability, and
- (c) Enhanced literature-based charging behavior modeling with more realistic assumptions.

We sincerely hope these revisions meet Nature Communications' standards for a qualified research article. In the following, we highlight the cited content from the manuscript and Supplementary files in blue and the corresponding revisions in the manuscript and Supplementary files in yellow. Thank you again for your constructive critique.

Comment 1

The concept of using EVs for peak shaving and load balancing is well-established, and while the granular modeling approach is noteworthy, the study does not introduce fundamentally new methods or paradigms. The authors should clearly explain if their proposed approach provides significant novelty and results in new general insights.

Response to Comment 1

Thank you very much for your comment. The novelty of the study mainly lies in a systematic, megacity-scale V2G potential evaluation framework. The proposed framework is capable of considering various V2G factors among hundreds of thousands of P-PEVs, including vehicle travelling-charging behaviors, user willingness and incentives, range anxiety, and battery degradation costs. The framework provides a user-centric viewpoint to depict travelling patterns and multi-dimensional energy dispatching effects of every individual. Equipped with different city-specific information, the framework further manages to evaluate V2G potential of other cities except Shenzhen. We believe the above contributions have not been made in other literature. In the context of rapidly growing P-PEV adoption worldwide, the insights of the study could be expected to enlighten megacity-level V2G potential evaluations and their policy-making.

According to your comment, we rewrite our title as:

“Unlocking vehicle-to-grid potential of load shifting in China’s megacities considering comprehensive real-world behaviors”

We rewrite our abstract as:

Global decarbonization necessitates large-scale electrification, yet urban grids face stability challenges from surging loads. Vehicle-to-grid (V2G) technology, leveraging private plug-in electric vehicles (P-PEVs) holds promise for balancing and shifting electric loads via bidirectional charging and discharging. However, megacity-level V2G potential evaluation has long been missing due to unresolved challenges in fine-grained modeling considering private user behavior patterns. Here, we propose a computationally feasible Mobility and V2G Coupled (MOVC) framework, which includes individual P-PEV travel-charge behavior, user compensation, user willingness to participate V2G, and battery degradation for unified user-centric V2G potential evaluation. Mobility and charging behavior data are systematically analyzed from 480,000 P-PEVs in Shenzhen, a megacity in China, revealing 2,300 MW peak-shaving capacity, reducing peak-valley ratios by 73% via a relaxed linear-programming approach. V2G strategies maintaining higher user satisfaction show insignificant sacrifice on the load shifting performance, while simultaneously lowering battery degradation costs and compensation costs by 30-40% and 5-13% respectively across different seasons. A forward-looking scenario considering quick charging (QC) availability in Shenzhen and other three China’s megacities showcases insignificant benefits (0%-12% lower peak-valley ratios) but exacerbates V2G scheduling inequity (4%-95% higher standard deviation). This work enlightens

megacity-level V2G potential evaluations and their policy-making considering real-world behaviors, inspiring a sustainable integration of electrified vehicles into many other critical energy infrastructures.

We rewrite our introduction in the revised manuscript, Introduction, pages 2-3, lines 32-95:

China's 2020 commitment to achieve carbon peak by 2030 and neutrality by 2060 aligns with its rapid electrification roadmap, targeting 64% industrial electrification by 2050 to transition from fossil fuel dependence toward sustainable energy¹⁻³. However, the foreseeable widespread adoption of electric equipment challenges urban distribution grids, requiring techniques *e.g.*, load shifting to stabilize voltage levels and improve efficiency⁴. While immobile batteries enable load shifting, their centralized deployment raises considerable safety concerns, shifting research focus to dispatch distributed private plug-in electric vehicles (P-PEVs) and their onboard batteries. This vision is supported by Vehicle-to-grid (V2G) technology, which is extensively explored to have noticeable benefits in load shifting⁵⁻⁷, energy savings⁸⁻¹¹, and economic gains¹²⁻¹⁵. China's rapidly expanding P-PEV market further broadens the operational potential of V2G, making it promising to utilize the onboard batteries for urban load shifting via V2G without sacrificing their mobility demand¹⁶⁻¹⁸.

However, making city-level decisions on V2G policies and incentives is difficult without reliable quantitative analysis.

Existing literature mainly considers V2G based on the mobility or charging behavior of PEVs. The mobility models are proposed to distinguish the spatial movement patterns of individuals. Those methods can be divided into Origin-Destination (OD) analysis¹⁹⁻²¹, Markov Chain-based models²²⁻²⁴, OD-based trip chain models²⁵⁻²⁷, graph-based models²⁸, *etc.* The charging models include more subdivisions, covering large-scale valley filling^{29,30}, load forecasting³¹, user-oriented charging behavior analysis and scheduling strategies³²⁻³⁵, V2G arbitrage³⁶⁻³⁷, user willingness of participating V2G and incentive mechanisms³⁸⁻⁴⁶, range anxiety⁴⁷, battery degradation effects⁴⁸⁻⁵², V2G optimization relaxation⁵³⁻⁵⁵, *etc.* Multiple methodologies are leveraged to address these problems with high accuracy or efficiency, including mixed-integer linear programming (MILP)^{37,47}, heuristic algorithms^{35,36,59}, stochastic programming^{40,41,51}, deep reinforcement learning⁵⁶⁻⁵⁸, user surveys^{38,42,43,45}, simulations^{48,52}, experimental measurements⁴⁹, *etc.* However, existing works have their scopes narrowed in single or several subareas by studying basically microgrids or small PEV fleets, with megacity-level user behaviors being considerably underinvestigated. This

noticeable gap is critically prohibitive for megacity policymakers to consider V2G promotion and their sustainable integration into many energy infrastructures.

A noteworthy research branch lies in integrating the mobility and charging patterns of PEVs to reveal spatial-temporal V2G insights from large-scale real-world data observations. The related works can be divided into local viewpoints and global ones. For the former, Bian et al. discuss the time-space distribution of PEVs within a given distribution grid ⁵⁹. Wu et al. propose a V2G scheduling method considering home, work and commercial microgrids ⁶⁰. Orfanoudakis et al. present a V2G simulation platform, which considers various user travelling behaviors in public, work and residential places, and evaluates user V2G satisfaction and battery degradation ⁶¹. However, these works mainly adopt a grid operator-centric perspective, where the user travelling features are simplified to entries and exits within predefined regions. For the latter, Zhang et al. analyze V2G implementation potential across Japan's prefectures, considering regional differences in power demand, temperature, user behavior, *etc* ⁶². Xu et al. integrate large-scale mobile phone data and PEV charging profiles to estimate fine-grained PEV mobility and energy demand ⁶³. Qian et al. develop a simulation platform to discuss V2G operations in coupled urban power and transportation networks ⁶⁴. Zhang et al. combine travel patterns and charging behaviors to predict EV charging demand in Beijing ⁶⁵. However, since the computational costs increase significantly with the growth of PEV scale, the scopes of these works are reduced in terms of either tasks (coordinated charging ⁶³, load forecasting ⁶⁵) or modelling details (regional transportation networks ⁶⁴, simplified charging costs ⁶⁴ or charging behaviors ⁶²). Unfortunately, systematic megacity-level V2G potential evaluation has long been missing due to unresolved challenges in fine-grained modeling considering user-centric private user behavior patterns.

In this paper, we develop a Mobility and V2G Coupled (MOVC) framework, which addresses the lack of systematic, megacity-scale V2G potential evaluation due to computational limitations or insufficient granularity in modeling user behaviors. Fig. 1 gives an overview of the research methodology. The framework comprehensively considers various V2G factors among hundreds of thousands of P-PEVs, including vehicle travelling-charging behaviors, user willingness and incentives, range anxiety, and battery degradation costs. The above considerations are derived from a user-centric perspective, capable of depicting travelling patterns and multi-dimensional energy dispatching effects of every individual. The framework illustrates V2G potential in Shenzhen and

other three China's megacities, considering seasonal load variations, growing P-PEV penetration rates, discrepancies of travelling characteristics in different cities, and quick charging (QC) availability. Specifically, we follow the settings of Timegeo²² to simulate travel patterns of P-PEVs in Shenzhen and Beijing based on 9,108,081 and 9,651,948 location records from 36,932 and 38,129 anonymous P-PEV users, respectively. The number of P-PEVs is further expanded to 480,000 or larger to match the holdings of the cities. To model charging actions, we modify the utility-based model³³ to consider possible inclination to range anxiety mitigation, larger charging power and lower charging costs among users. We implement large-scale V2G optimization via a linear-programming (LP)-based model with proper relaxations⁵³⁻⁵⁵ to flatten the city load curve as much as possible under different constraints. Our results indicate that fully utilizing the parking time of collective P-PEVs for V2G can release 2,332 MW and 2,320 MW of peak shaving and valley filling capacity during annual peak hours in Shenzhen. The weekly median peak-valley difference ratio can be accordingly reduced from 0.362 to 0.097. Adjusting V2G strategies to maintain higher user V2G satisfaction may reduce 30-40% battery degradation costs and 5-13% V2G compensation costs. In the assumed future with higher P-PEV penetration rates and upgraded residential distribution grids to support QC, our results indicate marginal additional V2G benefits but more unequal energy scheduling among users.

While our framework does not propose fundamentally new theories, it offers a feasible user-centric approach to analyzing megacity-scale V2G potential. The theoretical components employed in the work are customized to serve this specific purpose. Furthermore, the modular design allows the components to be substituted with other works as needed.

As we illustrate in the revised manuscript, Discussion, page 13, lines 315-328:

A key advantage of MOVC framework lies in its decoupled perspective for megacity-scale V2G analysis. Owing to its modular architecture, the framework allows integration of diverse V2G studies with varying focus areas. For instance, the mobility model is interchangeable with other trip-chain models⁶⁵, graph-based origin-destination models²⁸. The charging behavior model is compatible with analysis on charging station data⁶⁹, large language model (LLM)-based methods⁷⁰, or can incorporate user surveys^{38,42,45} to support realistic city-specific studies. The V2G optimization can also be substituted with deep reinforcement learning paradigms⁵⁶⁻⁵⁸. However, some drawbacks exist in the proposed framework. First, the mobility model is not accurate enough

in predicting 0-2km short distances, as shown in Fig. 3b. Second, the framework includes several economic metrics, but the primary objective is load shifting rather than detailed cost-effectiveness analysis (e.g., Pareto fronts between load shifting effects and economic costs). Exploring the latter problem with proper relaxations or other efficiency techniques would support practical applications of V2G strategies. Third, the framework focuses on individual V2G scheduling and omits behavioral similarity that probably exists in a large number of P-PEVs. Given that P-PEVs will continue to proliferate rapidly in China, incorporating carefully-designed clustering algorithms would further improve the computational efficiency.

Comment 2

The study's findings are specific to Shenzhen, and the manuscript does not adequately demonstrate the framework's transferability to other urban contexts. Broader applicability to cities with differing EV penetration rates, energy grids, or transportation systems should be explored.

Response to Comment 2

We appreciate your suggestions. We admit that the study presented in the first manuscript is limited to Shenzhen, which impairs the general applicability of our results and conclusions. In the revised manuscript, we analyze V2G effects in Beijing, Shanghai, Guangzhou and Shenzhen, which are the four recognized megacities in China. Your concerns about the transferability in terms of transportation systems, EV penetration rates, and energy grids are addressed as follows:

(a) Different transportation systems

We consider Beijing's case in the revised manuscript based on 2022 Beijing P-PEV Trajectories Dataset (2022 BPTD), which is consistently extracted with 2022 Shenzhen P-PEV Trajectories Dataset (2022 SPTD) analyzed in the first manuscript. Specifically, 2022 BPTD collects 9,651,948 location records from 38,129 P-PEV users active in Beijing from October 1, 2022 to December 31, 2022. We process Beijing's case following the collective paradigm used for Shenzhen's case. The modifications are listed as follows:

(a-1) Mobility results of Beijing

We add descriptions in the revised manuscript, Results, pages 3-4, lines 97-100:

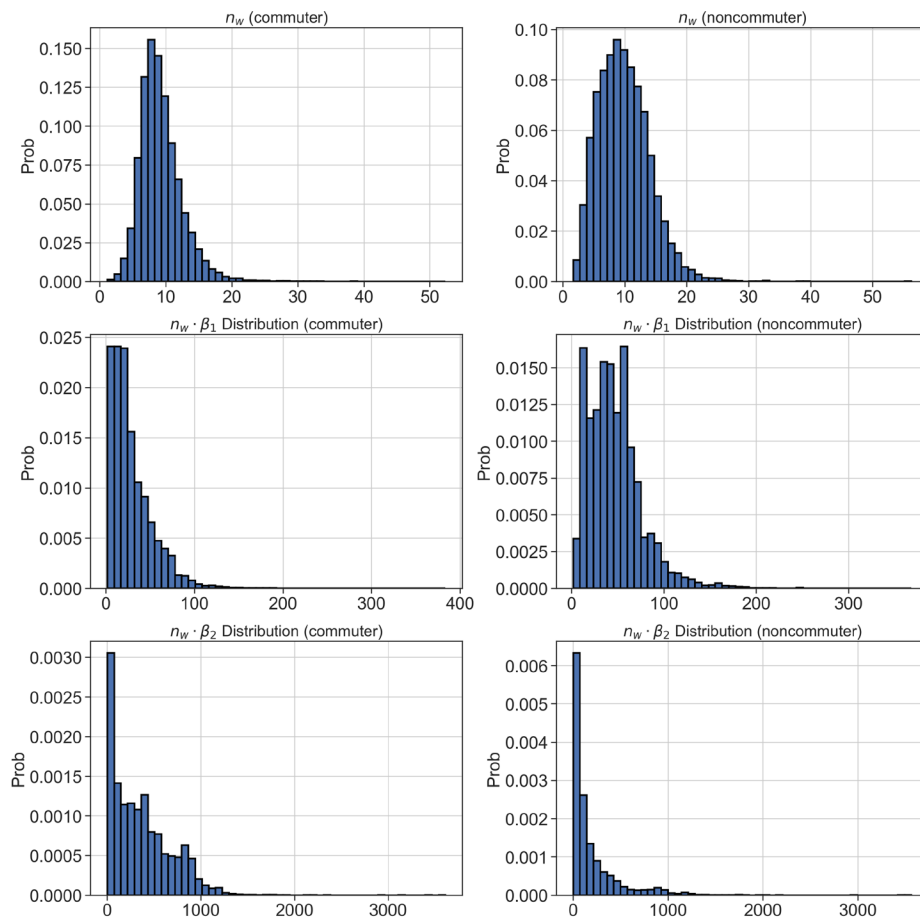
MOVC model. This work utilizes multiple real-world datasets, including (1) 2022 Shenzhen P-PEV Trajectories Dataset (2022 SPTD) and 2022 Beijing P-PEV Trajectories Dataset (2022 BPTD),

which collect three-month 9,108,081 and 9,651,948 location records of 36,932 and 38,129 anonymous P-PEV users in Shenzhen and Beijing, respectively.

In the revised manuscript, Results, page 12, lines 274-276:

We leverage 2022 BPTD to depict mobility and charging patterns of Beijing P-PEVs following the collective paradigm introduced in this work, as provided in Supplementary Fig. 2,4,6, 16-19.

We add results of key individual travelling parameters n_w, β_1, β_2 in Supplementary material, page 3, Supplementary Fig. 2:



Supplementary Fig. 2: Distributions of n_w, β_1, β_2 from 2022 BPTD.

We compare the median values of $n_w, n_w \cdot \beta_1, n_w \cdot \beta_2$ of commuters and noncommuters, respectively, between 2022 BPTD and 2022 SPTD, in Supplementary material, page 4, Supplementary Table. 1:

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

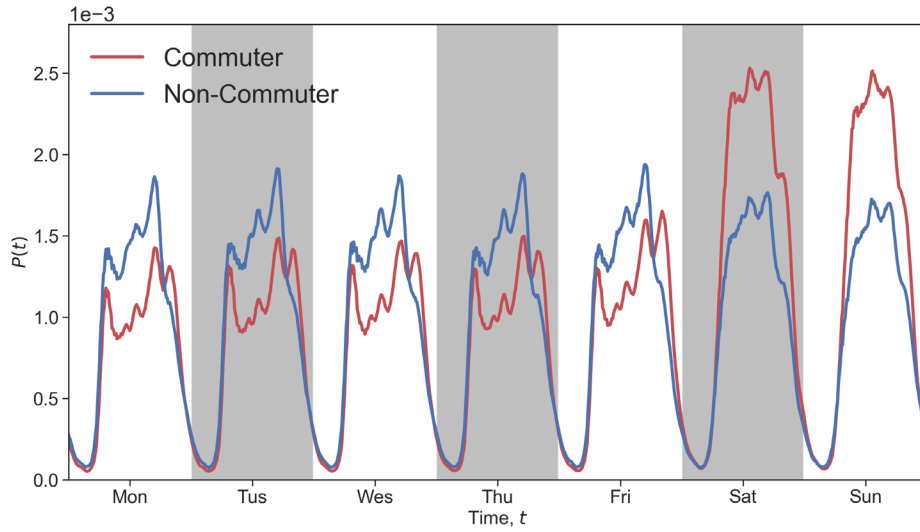
Median values

2022 BPTD

2022 SPTD

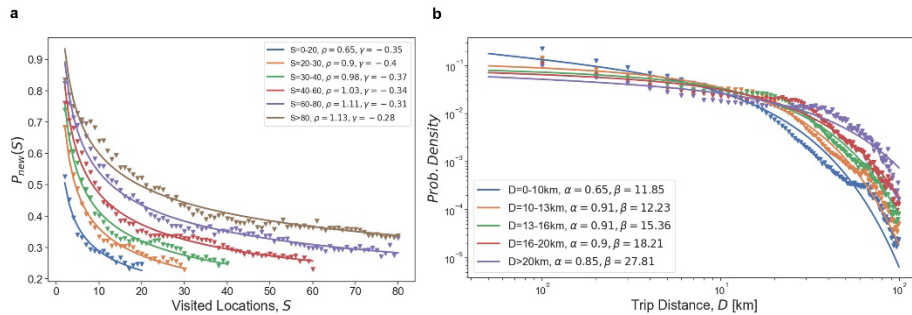
n_w (commuter)	8.77	9.45
n_w (noncommuter)	9.54	9.52
$n_w \cdot \beta_1$ (commuter)	22.18	20.78
$n_w \cdot \beta_1$ (noncommuter)	42.91	48.80
$n_w \cdot \beta_2$ (commuter)	302.00	408.36
$n_w \cdot \beta_2$ (noncommuter)	90.22	146.13

195 We provide the patterns of circadian travel rhythms ($P(t)$) in Supplementary material, page 5,
196 Supplementary Fig. 4:



197 **Supplementary Fig. 4: Circadian travel rhythms $P(t)$ for commuters and non-commuters of**
198 **2022 BPTD.**

199 We illustrate the grouped spatial parameters of P-PEVs in Supplementary material, page 6,
200 Supplementary Fig. 6:



201 **Supplementary Fig. 6: Grouped spatial parameters $\rho, \gamma, \alpha, \beta$ of Beijing P-PEV users.**

202 We compare the average travel distance between 2022 BPTD and 2022 SPTD in Supplementary
203 material, page 8, Supplementary Table. 3:

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

We use the number of district-level parking lots of Beijing for expanding the number of P-PEVs from 38,129 to larger values. As introduced in Supplementary material, page 8, Supplementary Note. 7, Table. 4:

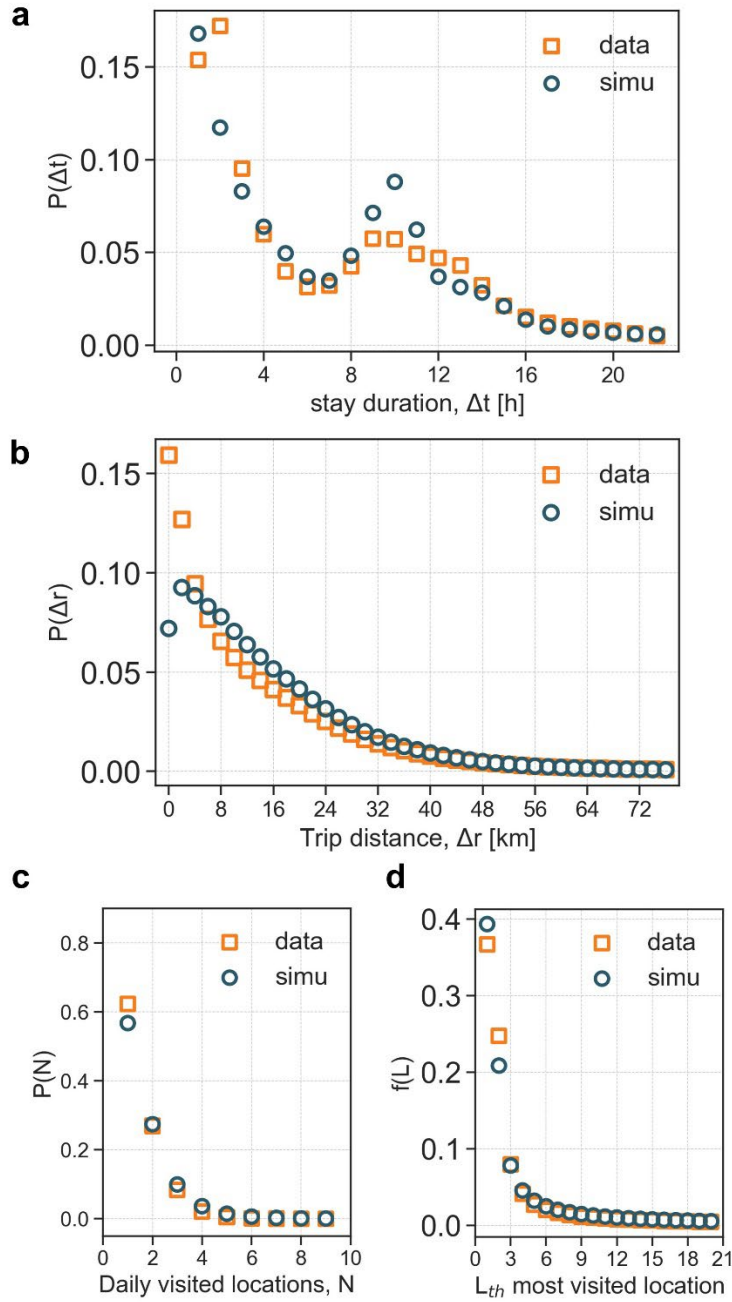
Supplementary Note. 7: Expanding the number of P-PEVs. Since 2022 BPTD and 2022 SPTD only contain 30,000-40,000 P-PEV users, which are not comparable to actual Beijing and Shenzhen P-PEV holdings, we consider to expand the number of P-PEVs by considering the parking spot distributions of the cities. 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 $N_e * 480000 * \frac{N_p^i}{\sum_i N_p^i}$. 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 Transport[2]. The district-level numbers of parking lots are provided in Supplementary Table. 4.

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

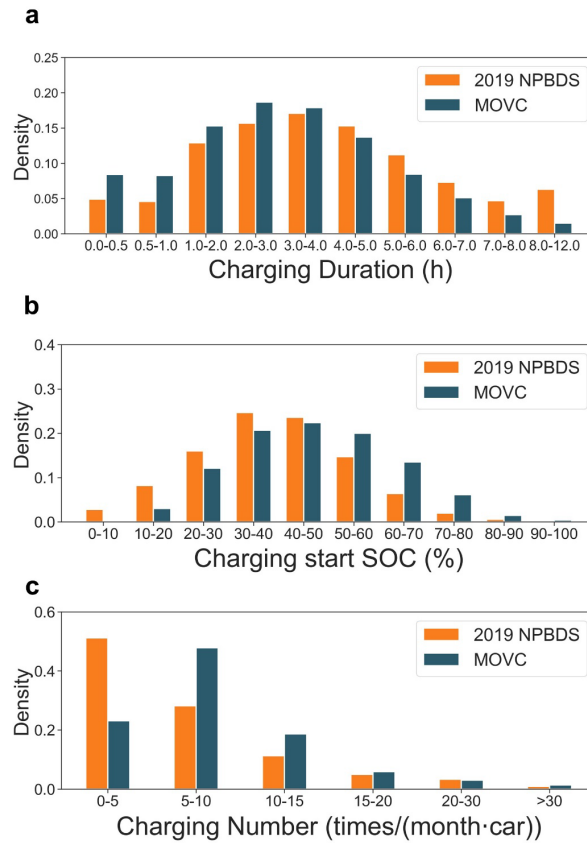
220 We provide key mobility characteristics of Beijing in Supplementary material, page 16,
221 Supplementary Fig. 16:



Supplementary Fig. 16: 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.

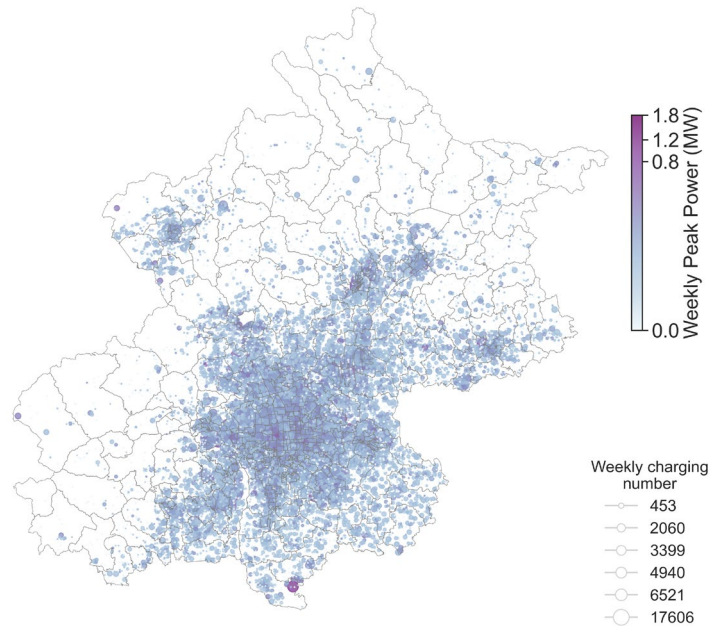
(a-2) Charging behaviors of Beijing P-PEVs under Charge Only

We validate the charging behavior simulation results of 2022 BPTD in Supplementary material, page 17, Supplementary Fig. 17:



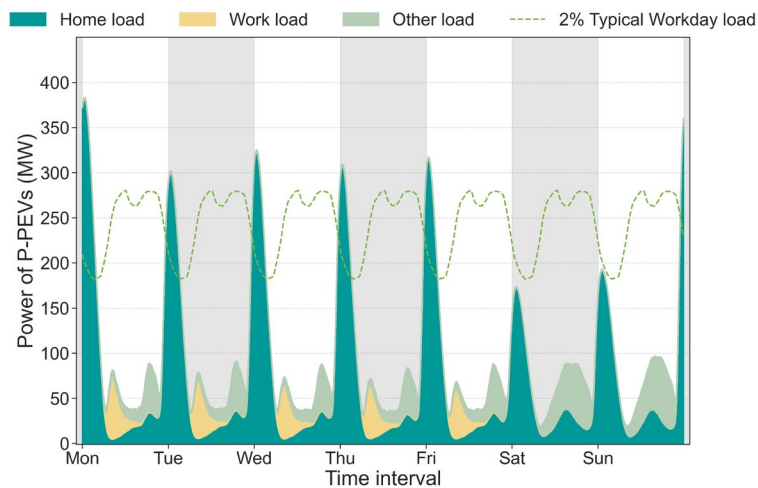
Supplementary Fig. 17: Validation of charging behavior simulation results of 2022 BPTD under Charge Only. a, Probability distributions of charging duration. b, Probability distributions of charging start SOC. c, Probability distributions of monthly charging number.

We illustrate the temporal-spatial distributions of Beijing P-PEV charging load in Supplementary material, page 18, Supplementary Fig. 18:



Supplementary Fig. 18: 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.

We decompose Beijing's aggregated P-PEV charging load according to the charging places, in Supplementary material, page 19, Supplementary Fig. 19:



Supplementary Fig. 19: The load curve of Beijing P-PEVs, divided into charging places.

(b) EV penetration rates

We consider deeper EV penetration rates by using P-PEV holdings in 2024. Specifically, the scale of P-PEVs is analyzed by comparing the V2G effects of 780,000 P-PEVs (Shenzhen), 640,000 P-PEVs (Beijing), 1,020,000 (Shanghai), and 650,000 (Guangzhou).

(c) Different energy grids

We perform V2G in the four megacities to flatten their own electricity load patterns. As introduced in the revised manuscript, Methods, page 14, lines 361-365:

We leverage the 2019 Shenzhen electricity load profile and 2019 Guangzhou electricity load profile to obtain the city's original overall load. The data contains hourly load values for Shenzhen and Guangzhou over the entire year. We select a week's values of load data and interpolate them to match the 10-minute interval resolution of the MOVC framework. For load profiles of Beijing and Shanghai, only typical workday load data is available. We thus replicate the single workday load pattern across several days for the following optimization.

(d) Integrated experiments considering different transportation systems, EV penetration rates and energy grids

We add experiments to discuss extreme V2G cases in the four cities. Our modifications also consider Reviewer 2, Comment 6 and Reviewer 3, Comment 6. In the revised manuscript, Results, page 12, Fig. 9:

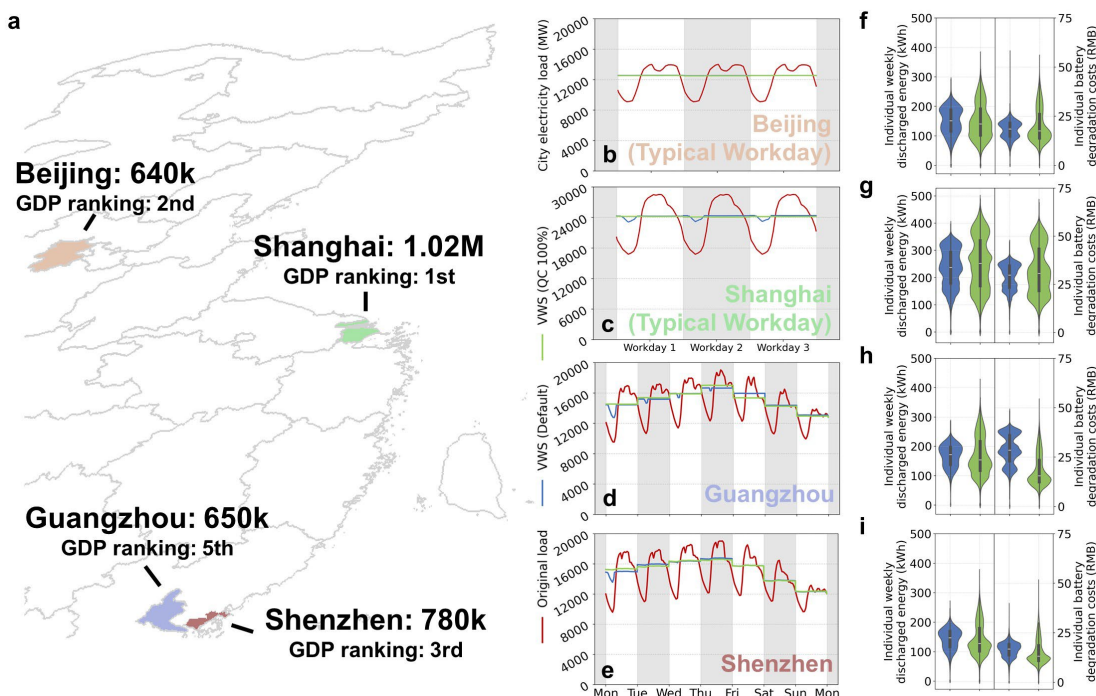


Fig. 9: V2G evaluation in Beijing, Shanghai, Guangzhou and Shenzhen based on MOVC. a,
P-PEV holdings of the four cities in 2024, and their GDP rankings in China's cities. b,c,d,e,
Load balancing results of the four cities under VWS (Default) and VWS (QC 100%), respectively.
Since we do not have the annual load data of Beijing and Shanghai, we illustrate their results in
typical workdays. For Guangzhou and Shenzhen we select a week with the annual peak. **f,g,h,i,**
Distributions of individual weekly discharged energy and battery degradation costs under the two
V2G scenarios.

In the revised manuscript, Results, pages 12-13, lines 269-306:

Extreme V2G imagination for China's megacities. MOVC has the transferability to other urban
contexts. The generalization ability is utilized to explore two substantial problems: (1) How V2G
effects evolve with larger P-PEV penetration rates, and (2) whether city-level upgrades in residential
distribution grids for accommodating comprehensive QC are cost-effective. To this end, we
illustrate the estimated P-PEV holdings of Beijing, Shanghai, Guangzhou and Shenzhen in 2024,
reaching 640,000, 1,020,000, 650,000 and 780,000, respectively. The selected cities represent the
most economically developed cities in China. We leverage 2022 BPTD to depict mobility and
charging patterns of Beijing P-PEVs following the collective paradigm introduced in this work, as
provided in Supplementary Fig. 2,4,6, 16-19. We use 2022 BPTD to generate vehicles for Beijing's
case, and use 2022 SPTD for Shanghai, Guangzhou and Shenzhen. We compare two V2G scenarios,
namely (1) VWS (Default), where the settings are identical to the ones in Fig. 6-Fig. 8 except for
the number of involved P-PEVs. (2) VWS (QC 100%), where P-PEVs are allowed to use QC even
in *Home* regions, and all the Park actions will perform VWS with QC power levels.

As shown in Fig. 9b-Fig. 9e, the discussed two scenarios exhibit similar load balancing effects
with larger P-PEV penetration rates. Specifically, the weekly median PVDR values for all cities
under VWS (Default) and VWS (QC 100%) are 0.0015, 0.0487, 0.0262, 0.0109 and 0.0015, 0.0015,
0.0016, 0.0092, respectively. Compared with PVDR of original loads (0.35, 0.41, 0.36, 0.36), extra
benefits only reach 0%, 12%, 7%, 5%, respectively. In Fig. 9f-Fig. 9i, after full QC upgrades, the
median values of weekly discharged energy in Beijing, Guangzhou and Shenzhen decrease from
151.25, 171.66, 147.04 kWh to 140.47, 154.21, 128.10 kWh, respectively. However, in Shanghai
where the overall city load is higher, the values increase from 237.30 to 252.20 kWh. Regarding
battery degradation costs, the median values of Beijing, Guangzhou and Shenzhen decrease from

18.55, 28.54, 16.78 RMB (around 2.56-2.31 USD) to 17.54, 15.57, 13.19 RMB (around 2.42-1.82 USD), respectively. In contrast, the metric of Shanghai slightly increases from 29.87 to 30.71 RMB (around 4.12 to 4.24 USD). Another noteworthy effect is observed in Shenzhen's case. With the deeper P-PEV penetration, the median weekly discharged energy decreases from 220.29 kWh (Fig. 7d, summer) to 147.04 kWh-128.10 kWh, and the median battery costs also decrease from 29.95 RMB (around 4.13 USD, Fig. 7e, summer) to 16.78-13.19 RMB (around 2.31-1.82 USD). Notably, the inequity of both metrics in all cities grows significantly. For the weekly discharged energy, the standard deviation of all individuals increases from 46.05, 70.17, 39.00, 35.43 kWh to 60.85, 97.37, 65.74, 54.59 kWh, respectively. For the battery degradation costs, the metric goes up from 4.42, 6.82, 7.71, 3.83 RMB (around 0.61-0.53 USD) to 8.61, 12.61, 7.98, 6.30 RMB (around 1.19-0.87 USD).

The results indicate that as the number of V2G-related vehicles increases, cities with mild load profiles (Beijing, Guangzhou, Shenzhen) exhibit lighter optimization burdens to each vehicle, leading to lower discharged energy and battery degradation costs. In such cases, QC upgrades yield marginal additional benefits. Conversely, in cities with higher load levels (*e.g.*, Shanghai), achieving comparable effects requires a higher penetration of P-PEVs. Additionally, expanding QC infrastructure significantly increases the variance of both discharged energy and degradation costs (Fig. 9f-Fig. 9i), enlarging the standard deviation by 32%-69% and 4%-95%, respectively. Customized V2G incentive policies may be required to mitigate such inequity. Given the substantial costs for residential grid upgrades, city policymakers should carefully weigh the trade-offs between investments and expected benefits.

In the revised manuscript, Discussion, page 13, lines 313-314:

Additionally, we discuss the impacts of both higher P-PEV penetration rates and fully-covered QC under extreme assumptions.

In the revised manuscript, Discussion, page 14, lines 341-343:

A generalization study across China's megacities indicates that the imaginary residential QC upgrades offer marginal benefits in load balancing (0%-12%) and battery lifespan, but may exacerbate individual energy scheduling inequity by 4%-95%.

Comment 3

The charging behavior model assumes a probabilistic approach based on stay duration and state of

charge (SOC), but it does not account for real-world behavioral factors, such as user cost sensitivity or range anxiety. Incorporating behavioral insights through surveys or literature-based models would enhance the realism of the framework.

Response to Comment 3

Thank you very much for the comment. We admit that the charging behavior model proposed in the first manuscript omits important real-world attributes. To improve the realism of the framework, we modify the utility-based charging behavior model of Liu et al⁴. We introduce the model in the revised manuscript, Methods, pages 15-16, lines 412-454:

User charging behavior modelling. We modify the utility-based charging behavior method³³ to establish our charging decision model. Available charging modes upon Stay arrivals are defined as $C_{i,s} = \{cm: L_{cm}^{i,s} = 1\}$ for the s -th Stay of the i -th P-PEV. We use $cm = 0,1,2$ to refer no-charging, slow charging and QC, respectively. Denote $\exp(V_{L_{cm}^{i,s}})$ as the indirect utility of charging mode $L_{cm}^{i,s}$. The choice probability of each charging mode is expressed as follows³³:

$$P_{\text{choice}_{i,s}}(L_{cm}^{i,s}) = \begin{cases} \frac{\exp(V_{L_{cm}^{i,s}})}{\sum_{j \in C_{i,s}} \exp(V_{L_j^{i,s}})} & \text{if } cm \in C_{i,s}, \\ 0 & \text{if } cm \notin C_{i,s}. \end{cases} \quad (3)$$

For the no-charging (Park) choice, we normalize the indirect utility to 0. Except for the extreme V2G case, we assume that slow charging is always available in all kinds of stay regions (*Home, Work, Other*), and QC is only available in public location (*Work, Other*).

Furthermore, the indirect utility of charging mode $L_{cm}^{i,s}$ is defined as

$$V_{L_{cm}^{i,s}} = \beta_0 + V_{\text{SOC}_{i,s}^e} + V_{R_{L_{cm}^{i,s}}} + V_{\Delta \text{SOC}_{L_{cm}^{i,s}}} + V_{\text{Cost}_{L_{cm}^{i,s}}} + V_{\text{Stay}_{L_{cm}^{i,s}}}, \quad (4)$$

$$cm \in C_{i,s} \setminus \{0\},$$

where β_0 is a constant representing the average effect of the non-listed attributes³³ set as 1 in the base case. Each component V_X refers to the utility of attribute X .

The first component $V_{\text{SOC}_{i,s}^e}$ describes the relationship between the current SOC and user i 's corresponding willingness to charge³³, which is modelled as

$$V_{\text{SOC}_{i,s}^e} = \begin{cases} +\infty & \text{if } \text{SOC}_{i,s}^e \in (0, \text{SOC}_{i,s}^{\text{dem}}] \\ \beta_{\text{SOC}} \cdot \ln \left(\frac{1 - \text{SOC}_{i,s}^e}{\left(\frac{1}{\text{SOC}_B} - 1\right) \text{SOC}_{i,s}^e} \right) & \text{if } \text{SOC}_{i,s}^e \in (\text{SOC}_{i,s}^{\text{dem}}, 1) \\ -\infty & \text{if } \text{SOC}_{i,s}^e = 1 \end{cases} \quad (5)$$

Where SOC_B is the range buffer set as 0.3, $SOC_{i,s}^e$ is the current SOC, β_{SOC} is set as 3 in the base case, and $SOC_{i,s}^{dem}$ refers to the moving consumption of the sequential three trips.

The second component $V_{R_{L_{cm}^{i,s}}}$ describes user i 's preference toward high charging power³³, defined as

$$V_{R_{L_{cm}^{i,s}}} = \beta_R \cdot (P_{L_{cm}^{i,s}} - P_{home}), \quad (6)$$

where $P_{L_{cm}^{i,s}}$ is the charging power of charging mode $L_{cm}^{i,s}$, P_{home} is the charging power at Home, and β_R is a coefficient. A larger β_R indicates that the user prefers faster charging to reduce the charging duration, and a smaller one shows his tendency to avoid battery degradation. In our base case, β_R is set as 0.

The third component $V_{\Delta SOC_{L_{cm}^{i,s}}}$ refers to user i 's evaluation of potential energy he can obtain during Stay s . Generally, the utility of this attribute is designed to increase when a larger amount of energy can be expected to charge during Stay s ³³. The formulation is

$$V_{\Delta SOC_{L_{cm}^{i,s}}} = \beta_{\Delta SOC} \cdot [1 - (\Delta SOC_{L_{cm}^{i,s}} - 1)^2], \quad (7)$$

where $\Delta SOC_{L_{cm}^{i,s}} = \min\left(\frac{P_{L_{cm}^{i,s}} \cdot T_{i,s}}{E_i}, 1 - SOC_{i,s}^e\right)$ represents the potential increased energy under charging mode $L_{cm}^{i,s}$, $P_{L_{cm}^{i,s}}$ is the corresponding charging power, $T_{i,s}$ is the duration of Stay s , E_i is the battery capacity, and $SOC_{i,s}^e$ is the SOC of P-PEV i upon arrival of Stay s . $\beta_{\Delta SOC}$ is the coefficient set as 3 in the base case.

The fourth component $V_{Cost_{L_{cm}^{i,s}}}$ describes users' sensitiveness of charging costs. Differed from the constant electricity price setting³³, we introduce time-of-use (ToU) price to represent a more realistic cost evaluation. Specifically, the charging cost $Cost_{L_{cm}^{i,s}}$ is designed as

$$Cost_{L_{cm}^{i,s}} = \sum_{t=Ta_{i,s}}^{Td_{i,s}} P_{L_{cm}^{i,s},t}^{ToU} \cdot \Delta SOC_{L_{cm}^{i,s}} \cdot E_i, \quad (8)$$

where $Ta_{i,s}, Td_{i,s}$ refer to the arrival and departure time slots of user i , Stay s , respectively. $P_{L_{cm}^{i,s},t}^{ToU}$ denotes the ToU price at time slot t . The utility $V_{Cost_{L_{cm}^{i,s}}}$ is computed following

$$V_{Cost_{L_{cm}^{i,s}}} = -\beta_{Cost} \cdot (Cost_{L_{cm}^{i,s}} - Cost_{home}), \quad (9)$$

where β_{Cost} is set as 0.35 in the base case, and $Cost_{home}$ can also be calculated following Eq. (8).

The only distinction is that $Cost_{L_{cm}^{i,s}}$ uses $P_{L_{cm},t}^{ToU, pub}$ and $Cost_{home}$ uses $P_{L_{cm},t}^{ToU, home}$, representing different ToU price of public and private charging piles, respectively, and the former is generally more expensive. The details of ToU price are obtained from charging service providers⁶⁸ and are given in Supplementary Fig. 8, 9.

The last component $V_{Stay, L_{cm}^{i,s}}$ is added to simulate users' patterns of slow charging decisions. We assume that either a too short or a too long stay duration will decrease the users' willingness to perform slow charging. $V_{Stay, L_{cm}^{i,s}}$ is thus given as

$$V_{Stay, L_{cm}^{i,s}} = \begin{cases} \max(\beta_{Stay} \cdot (T_{i,s} - 2) \cdot (T_{i,s} - 10), B_{Stay}) & \text{if } cm = 1, \\ 0 & \text{if } cm \in \{0, 2\}. \end{cases} \quad (10)$$

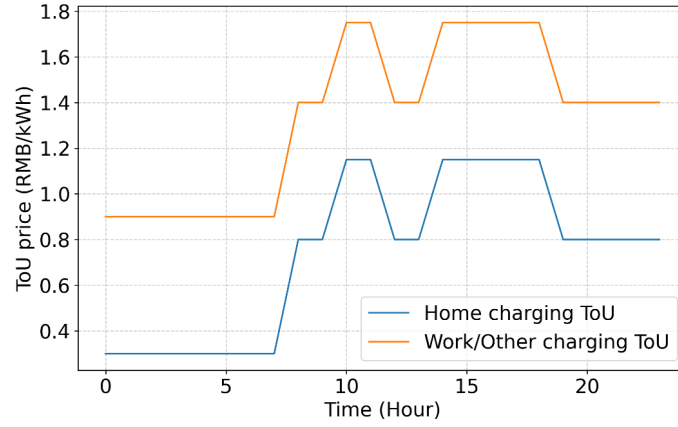
where the coefficient β_{Stay} and the buffer B_{Stay} are set as -0.2 and -3, respectively.

We provide settings of $\beta_0, \beta_{SOC}, \beta_R, \beta_{\Delta SOC}, \beta_{Cost}$ and β_{Stay} under different charging cases in Supplementary Table. 5.

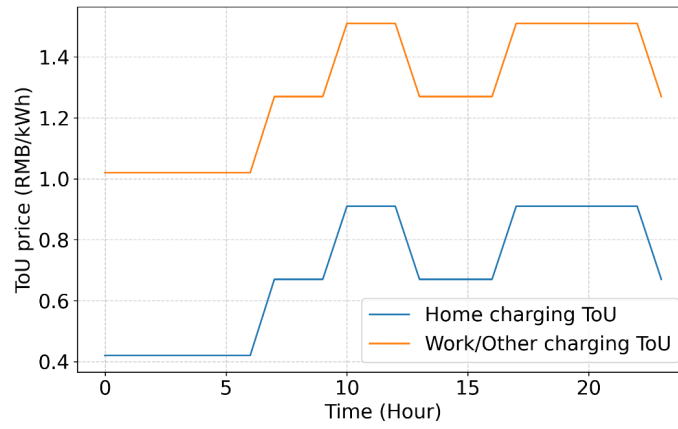
We add descriptions of the charging behavior model in Results, page 4, lines 114-119:

If a vehicle is in a Stay, MOVE uses a utility-based model³³ to specify its action of Park, Charge, or QC. The decisions are affected by multiple realistic factors, *e.g.*, the vehicle's current state of charge (SOC), the user's preference towards QC, the expected energy that can be obtained in the Stay depending on Stay duration along with available charging modes, and charging costs depending on the current time period and location (public or private). One can refer to Methods for details. Notably, we assume that all the decisions of Park, Charge and QC are made upon arrival of Stay-s.

Different from the work of Liu et al⁴, we introduce time-of-use (ToU) price to more precisely calculate users' utility of charging costs. We refer to the charging cost information in private/public places provided by charging service providers²⁴, as provided in Supplementary material, page 9, Supplementary Fig. 8,9:



Supplementary Fig. 8: ToU information of Shenzhen[3].



Supplementary Fig. 9: ToU information of Beijing[3].

We design a base case for the utility-based charging behavior model. Validation results of the model are provided in the revised manuscript, Results, page 5, Fig. 3h-Fig. 3j:

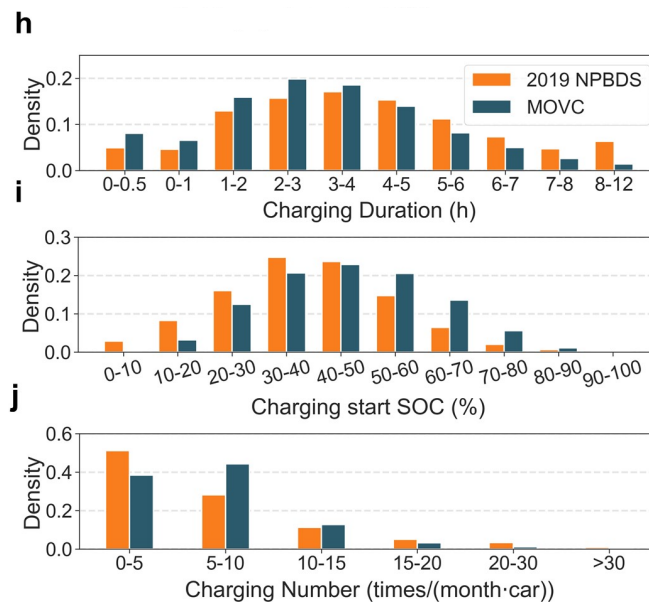


Fig. 3: Validation of the MOVC framework in mobility and charging behavior. h,i,j, Probability distributions of charging start SOC, charging duration, and monthly charging number, of MOVC results and 2019 NPBDS.

We update Shenzhen P-PEV charging profiles under the base case in the revised manuscript, Results, page 6, Fig. 4a-Fig. 4c:

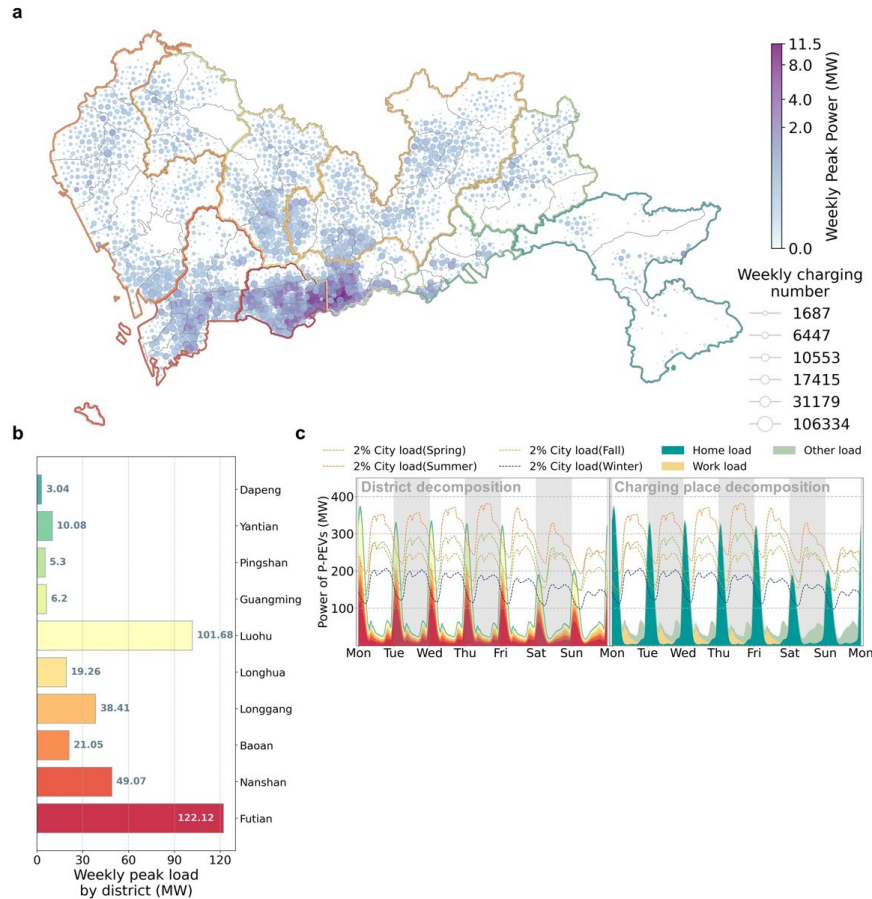


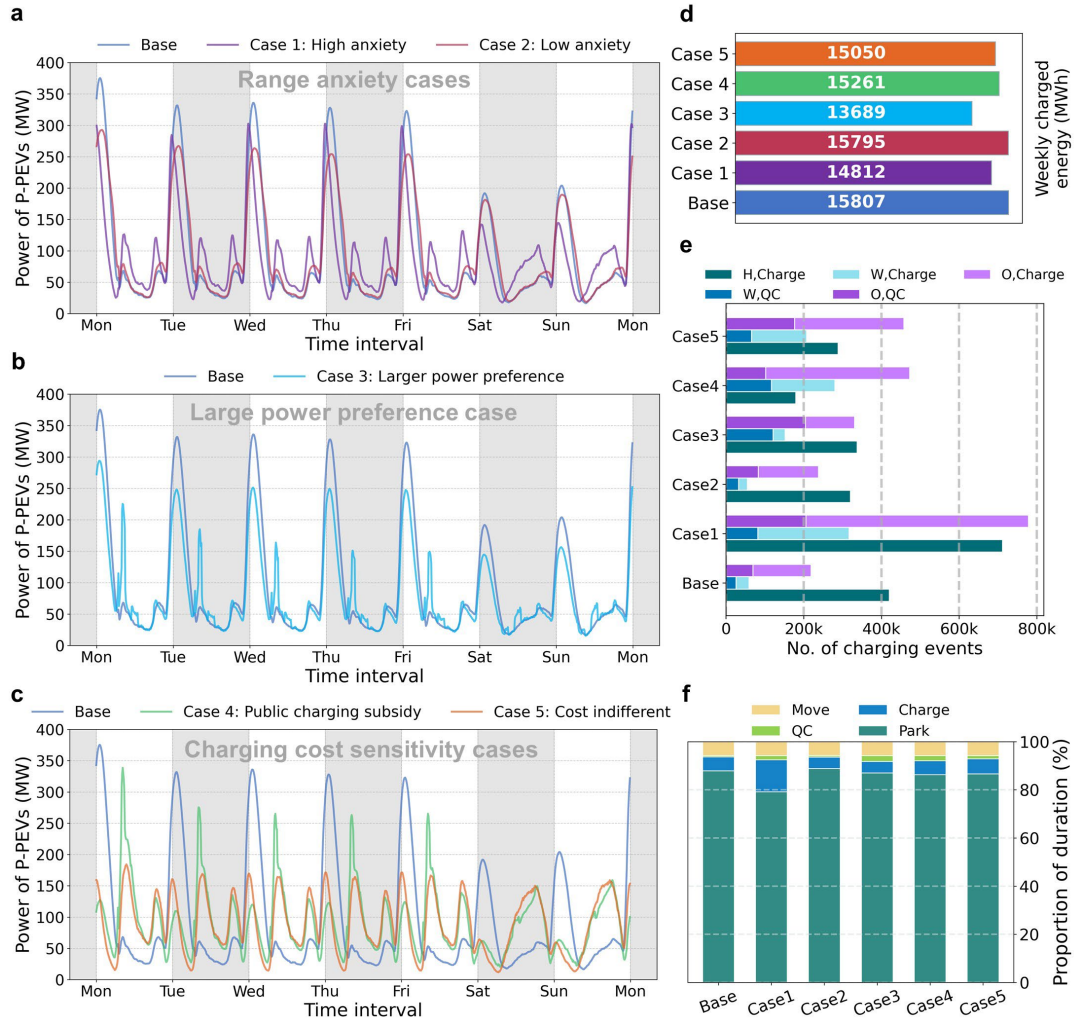
Fig. 4: Shenzhen P-PEV charging profiles and energy storage potential. a, Weekly peak load distribution of P-PEVs in Shenzhen, divided into stay regions. b, Weekly peak load distributions of P-PEVs, divided into districts. c, The load curve of Shenzhen P-PEVs, divided into charging places (left) and districts (right), respectively. The load magnitude of P-PEVs is comparable to 2% of Shenzhen's electricity peak load.

We add descriptions corresponding to Fig. 4a,4b in the revised manuscript, Results, pages 6-7, lines 149-164:

For spatial distributions, we illustrate the weekly peak load and the number of charging events of all stay regions in Fig. 4a, along with a district-level decomposition in Fig. 4b. The weekly peak mainly concentrates on Yuehai Street (Nanshan), Bantian Street (Longgang), and Futian Street

(Futian). The former two are the most economically active areas in Shenzhen, housing prominent companies, *e.g.*, Tencent, Huawei and China Resources, while the latter holds Futian and Huanggang ports as transportation hubs between Shenzhen and Hong Kong. Additionally, Luohu district has a noticeable aggregated peak load although its stay-region-level peak values are relatively intermediate.

We then decompose the P-PEV charging load into charging places (*Home, Work, Other*) and districts, as shown in Fig. 4c. Considering the time-of-use (ToU) electricity price, charging between 0 AM to 8 AM yields the lowest costs. Also, charging at private places (*Home*) is cheaper than at public places (*Work, Other*) by 0.6 RMB (around 0.083 USD)/kWh⁶⁸. Both factors make *Home* charging dominate especially at midnight, pushing the peak value to 375.18 MW. However, compared with Shenzhen's electricity load in different seasons, the load of P-PEVs accounts for only 2% in general, indicating that the given P-PEV penetration rate will not significantly burden the power grid. From the perspective of city load fluctuation, the grid may even benefit since the city valley load and the P-PEV charging peak load coincide at midnight under the guidance of ToU, a valley-filling effect can thus be expected.



To further simulate different cases considering user cost sensitivity and range anxiety, we design another five cases. As shown in the revised manuscript, Results, page 8, Fig. 5:

Fig. 5: P-PEV load profiles under different charging cases. a,b,c, Weekly load curves of P-PEVs under different charging cases and the corresponding values of weekly charged energy. **d,** The values of weekly charged energy under different cases. **e,** The numbers of Charge and QC events occurring at *Home*, *Work* and *Other* under different cases. **f,** Proportions of duration of all possible P-PEV actions (Move, Charge, QC, Park) under different cases.

We update the corresponding analysis in the revised manuscript, Results, pages 7-8, lines 177-202:

Case analysis of Shenzhen P-PEV load profile variations under Charge Only. The base case discussed above considers users' charging concerns relatively evenly. In this section, we design several cases to demonstrate the possible variations of P-PEV load profiles with different intensities of QC preference, range anxiety and sensitivity of charging costs. Specifically, we assume (1) Case 1, where users are very sensitive to a decrease in the SOC, and the charging probability quickly

increases when the SOC levels go lower³³. (2) Case 2, where users are not sensitive to SOC levels unless the SOC reaches a threshold³³. (3) Case 3, where users prefer larger charging power and tend to choose QC. (4) Case 4, where a public charging subsidy policy is introduced, and charging at *Home* yields identical costs compared with charging at *Work*, *Other*. (5) Case 5, where the factor of charging costs has no effect on users' charging decisions. The parameter settings of all cases are given in Supplementary Table. 5. Note that in all cases we assume that QC is only available at *Work*, *Other*, since most of the residential areas in China do not allow QC due to limited capacities of the distribution grid.

As shown in Fig. 5a, Base case, Case 1 and Case 2 have similar load patterns concentrated at midnight. In Fig. 5b, a daytime peak appears in Case 3, reflecting users' tendency towards QC at public places. In Fig. 5c, since public charging is subsidized in Case 4, the daytime peak further turns to a daily peak since *Home* charging is not attractive in terms of costs. The phenomenon will push the city's peak load higher since the latter also appears in daytime. Case 5 generates double-peak patterns at midnight and morning, the peak value becomes mild while the daytime load is generally higher than Base case, Case 1 and Case 2.

We compare the number of charging events under different cases in Fig. 5e. Case 1 generates significantly more charging events due to severe range anxiety. In contrast, Case 2 only keeps necessary charging activities. In Case 3 where users prefer larger power, the proportion of QC in public places is increased noticeably. In Case 4 and Case 5, *Work* charging occurs more frequently compared with Base case, and *Other* charging begins to dominate.

We illustrate the proportion of duration of all actions (Move, Park, Charge, QC) under different cases in Fig. 5f. In all cases, the proportion of Park accounts for more than 80% of all actions. Therefore, VWS is expected to utilize Stay duration much more than VWC for V2G scheduling. The V2G effects of VWC and VWS will be further discussed in the following section.

We provide parameter settings of different cases in Supplementary material, page 9, Supplementary Table. 5:

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

We add descriptions in the revised manuscript, Discussion, page 13, lines 331-335:

Results indicate that under Charge Only, P-PEV peak charging power only equals 2% of Shenzhen's peak load. Several variations of P-PEV charging patterns considering range anxiety, charging power, and charging costs are discussed, where the duration of Park actions consistently dominates. The observation motivates us to further discuss two V2G scenarios (VWC and VWS) and explore their combined effects under different compensation intensities.

We do not introduce insights from city-specific surveys to avoid introducing biases. However, prospective city-specific research can leverage such insights to modify the charging behavior model, as introduced in the revised manuscript, Discussion, page 13, lines 318-320:

The charging behavior model is compatible with analysis on charging station data⁶⁹, large language model (LLM)-based methods⁷⁰, or can incorporate user surveys^{38,42,45} to support realistic city-specific studies.

Comment 4

The code can be run as a "Reproducible Run" on codeocean.com. The data is also available. However, there is no README file to give instructions on how to use the code. Including such a file can greatly increase the usability of the code.

Response to Comment 4

Thank you very much for the comment. According to our collective modifications to the study, we update our codes and data on CodeOcean. We also add a README file in the "metadata" directory. We introduce package requirements, directory structure, and usage of every code in the README file. An excerpted preview is shown as follows:

Unlocking vehicle-to-grid potential of load shifting in China's megacities considering comprehensive real-world behaviors: Code instruction

This is a repository for reproducing the results of the article *Unlocking vehicle-to-grid potential of load shifting in China's megacities considering comprehensive real-world behaviors* in *Nature Communications*, paper ID: NCOMMS-24-79794A. Demo data for methodological reproduction is provided. Due to the related confidentiality agreement, full data is available on request.

Environments

The following packages are required:

```
jupyter
matplotlib
gurobipy>=9.5.1
numpy
pandas
scikit-learn
tqdm
joblib
pickle
```

Among which gurobipy requires an academic or commercial license. Please see <https://www.gurobi.com/> for details.

Data Structure

478 Figure.R1: Preview of the README file on CodeOcean.

Response to Reviewer #2

This paper aims to explore the potential of V2G for peak shaving by proposing a mobility and V2G coupled framework for charging and discharging strategies. The city of Shenzhen is taken as the case study. This study is valuable. The authors need to clarify the following issues.

Dear Respected Reviewer,

Thank you very much for taking valuable time to review our paper. We sincerely appreciate your recognition of the value of our work. We are also deeply grateful for the insightful suggestions provided, which have contributed to improving the quality of our manuscript. We have carefully addressed your concerns in a point-by-point manner, specifically regarding:

- (a) Study of the existing literature,
- (b) Probabilistic computation details in the P-PEV mobility model,
- (c) Clarifications in the overview figure,
- (d) Definition of V2G scenarios, and
- (e) Representativeness of urban load selection and the discussion of P-PEV scale.

In the following, we highlight the cited content from the manuscript and Supplementary files in blue and the corresponding revisions in the manuscript and Supplementary files in yellow. We hope this revised version will clarify your concerns and enhance the credibility of this work.

Comment 1

On Line 87-89, the statements are incorrect. Many studies have existed regarding the assessment of load balancing/leveling, peak shaving, and valley filling, such as: "High efficient valley-filling strategy for centralized coordinated charging of large-scale electric vehicles", "A systematic methodology for mid-and-long term electric vehicle charging load forecasting: The case study of Shenzhen, China", "The peak load shaving assessment of developing a user-oriented vehicle-to-grid scheme with multiple operation modes: The case study of Shenzhen, China". The authors are suggested to summarize the shortcomings of existing studies regarding peak-load shaving.

Response to Comment 1

Thank you very much for the comment. We admit that the research gap discussed in the first manuscript is not accurate. The works you mentioned are important in V2G-related research in terms of large-scale valley filling and user-oriented scheduling strategies. Compared with these works, our study mainly contributes to travelling behavior considerations and more detailed user-centric concerns. Therefore, we rewrite the part of the literature review in the revised manuscript, Introduction, page 2, lines 44-56:

Existing literature mainly considers V2G based on the mobility or charging behavior of PEVs. The mobility models are proposed to distinguish the spatial movement patterns of individuals. Those

methods can be divided into Origin-Destination (OD) analysis¹⁹⁻²¹, Markov Chain-based models²²⁻²⁴, OD-based trip chain models²⁵⁻²⁷, graph-based models²⁸, *etc.* The charging models include more subdivisions, covering large-scale valley filling^{29,30}, load forecasting³¹, user-oriented charging behavior analysis and scheduling strategies³²⁻³⁵, V2G arbitrage³⁶⁻³⁷, user willingness of participating V2G and incentive mechanisms³⁸⁻⁴⁶, range anxiety⁴⁷, battery degradation effects⁴⁸⁻⁵², V2G optimization relaxation⁵³⁻⁵⁵, *etc.* Multiple methodologies are leveraged to address these problems with high accuracy or efficiency, including mixed-integer linear programming (MILP)^{37,47}, heuristic algorithms^{35,36,59}, stochastic programming^{40,41,51}, deep reinforcement learning⁵⁶⁻⁵⁸, user surveys^{38,42,43,45}, simulations^{48,52}, experimental measurements⁴⁹, *etc.* However, existing works have their scopes narrowed in single or several subareas by studying basically microgrids or small PEV fleets, with megacity-level user behaviors being considerably underinvestigated. This noticeable gap is critically prohibitive for megacity policymakers to consider V2G promotion and their sustainable integration into many energy infrastructures.

Comment 2

The MOVC framework combines the Timegeo modeling framework in ref. (17) for urban mobility with the V2G optimization for charging and discharging strategies. It is a nice combination. My question is how to determine the probability for move or stay and for home or other places?

Response to Comment 2

Thank you very much for your recognition of MOVC's insights of integrating P-PEV mobility patterns with charging behaviors. The determination of temporal and spatial mobility choices mainly follows the paradigm provided by Timegeo¹. We introduce the details as follows:

(a) Identification of *Home* and *Work* places for a single P-PEV user

When labelling *Home* places, we consider P-PEV visits that start between 6 PM and 4 AM and end before 12 AM as potential home-type visits. A *Home* label will be identified from a user's historical travelling records with the most home-type visits. When labelling *Work* places, we assume that a longer trip has a higher probability for work purposes than a shorter one (e.g., shopping nearby).

We thus use a utility function $U(d_i, n_i) = n_i \times d_i$ to describe the likelihood that a stay region will be marked as work-type for a user i , where n_i denotes the visit frequency and d_i denotes the distance from user i 's home. Similarly, A *Work* label will be identified from a user's historical travelling records with the largest work-type utility.

We use some thresholds to avoid unreasonable labelling. For example, we assume that a valid *Home* label should account for at least 20% of the 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 explanations of identifying *Home* and *Work* places are added in the revised manuscript, Methods, page 14, lines 375-381:

Specifically, a visit that starts between 6 PM and 4 AM and ends before 12 AM will be labelled as a potential home-type visit. A *Home* label will be identified from a user's historical travelling records with the most home-type visits. When labelling *Work* places, we assume that a longer trip has a higher probability for work purposes than a shorter one (e.g., shopping nearby). We thus leverage a utility function $U(d_i, n_i) = n_i \times d_i$ to describe the likelihood that a stay region will be marked as work-type for a user i , where n_i denotes the visit frequency and d_i denotes the distance from user i 's home.

We add the description of thresholds in Supplementary material, page 2, Supplementary Note. 2:

Supplementary Note. 2: Identification of *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.

(b) Temporal choices of the mobility model

Temporal choices are about whether to move or to stay. The corresponding probability is determined by three P-PEV-specific parameters: a weekly home-based tour number (n_w), a dwell rate (β_1), and a burst rate (β_2). n_w is obtained by calculating the weekly average number of trips originating from *Home* for each user. β_1 and 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.

After the above parameters are determined for each user, the probability of temporal choices is given in the designed branches¹ as illustrated in Figure 2. Specifically, when a user is located at *Home*, the probability of moving is $n_w P(t)$, and the one of staying is $1 - n_w P(t)$. When a user is located at *Other*, the probability of moving is $\beta_1 n_w P(t)$, and the one of staying is $1 - \beta_1 n_w P(t)$. Here $P(t)$ denotes the weekly average travel circadian rhythm, which 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 *Work* labels). For the users located at *Other* places, they need to further determine whether to visit another *Other* place with the probability $\beta_2 n_w P(t)$ or to return *Home* with the probability $1 - \beta_2 n_w P(t)$.

(c) Spatial choices of the mobility model

To model P-PEV spatial decisions, we modify the rank-based exploration and preferential return (r-EPR) mechanism¹ to the gamma-distributed gravity-based EPR model. According to r-EPR, for a user who has made a temporal choice to move, he may explore a new location with the probability P_{new} or return to a previously visited location with the probability $1 - P_{new}$. The exploration probability P_{new} is defined as the function of the number of total previously visited locations S , presented as $P_{new} = \rho S^{-\gamma}$, where ρ is a scaling factor. Sequentially, the mobility model determines which place the user will move to. r-EPR assumes that the locations closer to the user's current location will have larger values of probability to be visited. We describe the mechanism using the gamma distribution, following:

$$f(x|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad (2)$$

where x denotes the distance between all possible locations to be visited and the current location, and parameters α, β are fitted from the ground truth dataset. Notably, we divide P-PEVs into

several groups to fit different parameters $\rho, \gamma, \alpha, \beta$ according to the various travelling patterns of each group.

The main modifications made according to your comment are as follows:

The revised manuscript, Methods, page 14, lines 376-377:

A *Home* label will be identified from a user's historical travelling records with the most home-type visits.

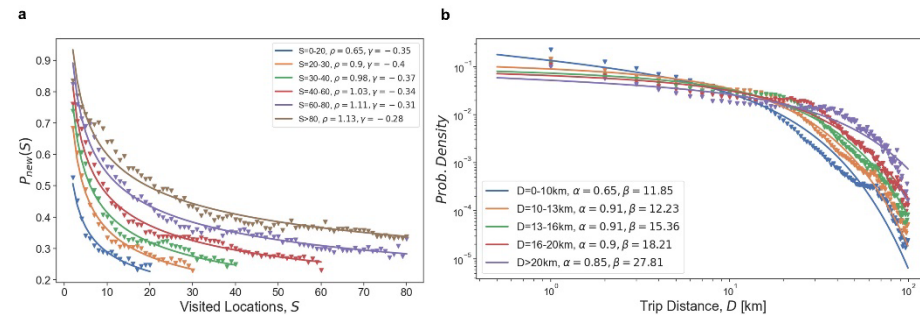
Supplementary material, page 2, Supplementary Note. 3:

Supplementary Note. 3: Determining 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 function[1]:

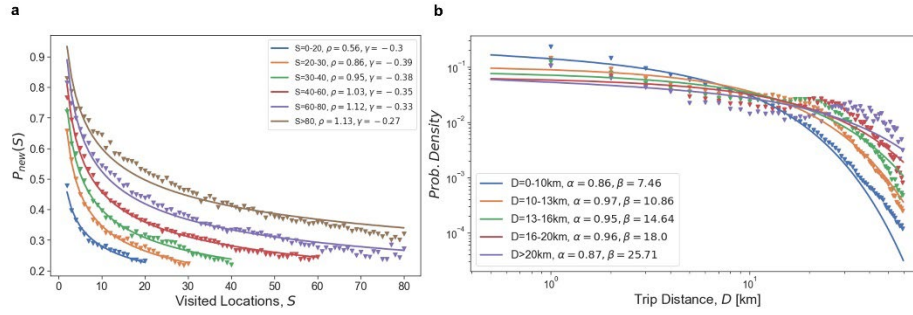
$$L(\beta_1, \beta_2) = \int |P_G(\Delta t) - P_S(\Delta t|\beta_1, \beta_2)| d\Delta t + \eta |\bar{N}_G - \bar{N}_S(\beta_1, \beta_2)|, \quad (1)$$

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.

Supplementary material, page 6, Supplementary Fig. 6,7:



Supplementary Fig. 6: Grouped spatial parameters $\rho, \gamma, \alpha, \beta$ of Beijing P-PEV users.



Supplementary Fig. 7: Grouped spatial parameters $\rho, \gamma, \alpha, \beta$ of Shenzhen P-PEV users.

Comment 3

In Figure 1, there are the statements “car charges the building and building charges the car”. I wonder if only the V2B is considered in this study.

Response to Comment 3

Thank you very much for your question. We apologize for causing your misunderstanding about the studied scheme in our work. In fact, we consider V2G in this study rather than V2B. We focus mainly on interactions between P-PEVs and the power grid, which we assume can take place in any stay region. We actually do not consider the vehicles’ energy operation details within specific building integrated energy systems. Although the topic is also noticeable, it falls outside the scope of this research.

According to your comment, we modify Figure 1 in the revised manuscript, Results, page 3, to correctly describe the scheme of our research, as follows:

Comment 4

In Figure 1 and Figure 2, the V2G action falls into three types: charge only, V2G when charging, and V2G when staying. In the V2G model, how do you distinguish these three scenarios for charging and discharging strategies?

Response to Comment 4

Thank you very much for the comment. In our work, the operation of V2G strategies is mainly constrained by two factors, namely the stay duration of P-PEVs and whether P-PEVs are connected to charging piles during a stay. We explain the considerations of V2G strategies according to the above factors as follows:

(a) Stay duration of P-PEVs

An important prerequisite of our V2G strategies is to keep the original travelling plans of users unchanged, which we believe will guarantee their autonomy of vehicle usage and avoid introducing non-economic interference in the V2G participation mechanism. Therefore, we use the mobility model to generate stay records of P-PEVs representing users' original travelling plans, which remain identical for all three V2G scenarios. In other words, time-related constraints (arrival time, departure time) in all V2G scenarios are the same.

(b) Whether P-PEVs are connected to charging piles during a stay

We realize that our inconsistent terminology may be confusing to readers. Hence, we modify the term "Stay" in the blue node in Figure 2 to "Park".

We use the charging behavior model to simulate P-PEV users' charging decisions, which include three cases, namely (1) Park, where there are no physical connections between the vehicle and the charging pile, (2) Charge, where users plug the charger in the slow charging socket of their vehicles, and (3) QC, where users plug the charger in the QC socket of their vehicles. We use the term "Stay" to denote a more general case where P-PEVs are immobile, opposite to the term "Move". The relationship among those activities is illustrated in Figure.R2:

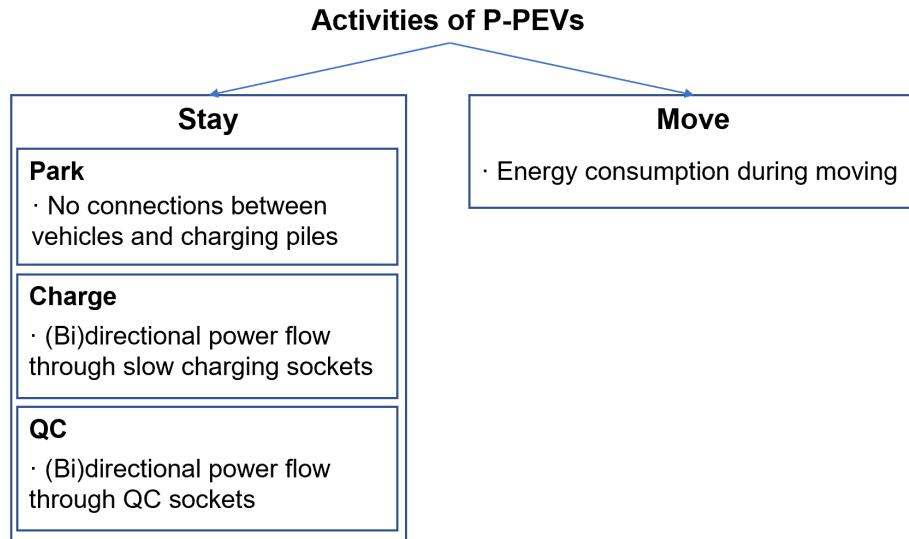


Figure.R2: Definition of all possible activities of P-PEVs.

Note that the above activities describe users' decisions when they do not participate in V2G.

Now we consider V2G activities based on users' decisions. The differences between Charge Only,

V2G When Charging, and V2G When Staying are illustrated in Figure.R3:

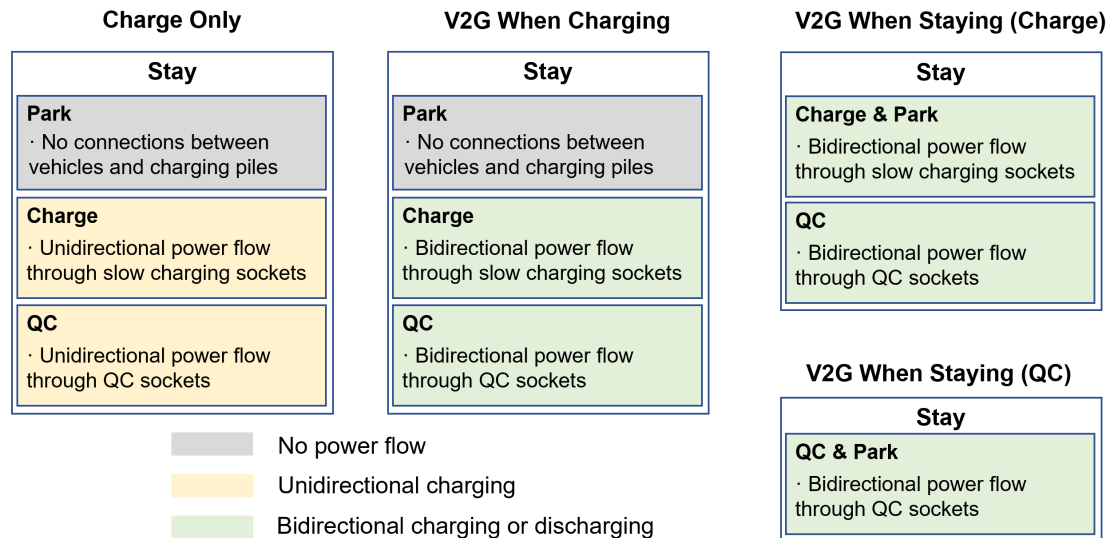


Figure.R3: Illustration of three V2G strategies: Charge Only, V2G When Charging, and V2G When Staying. Different situations of power flow are marked with corresponding colors.

Regarding real-world behaviors, we can see that Charge Only and V2G When Charging do not require users to change any of their original decisions. The only difference is the unidirectional or bidirectional power flow between the vehicles and the grid. However, V2G When Staying requires users to plug chargers into the charging sockets even when they do not intend to do so during a stay.

This scenario is designed to evaluate the load balancing potential of P-PEVs when they participate in V2G as much as possible. Accordingly, thanks to Reviewer 3's comment 6, we assume that users

participating in V2G When Staying will obtain higher economic subsidies compared with those who participate in V2G When Charging.

We adjust the constraints in our linear programming formulation to represent different V2G strategies. Specifically, in the revised manuscript, constraints related to V2G When Charging are Eq. (16-18) for Park, Eq. (19-24) for Charge, and Eq. (25-30) for QC. Constraints related to V2G When Staying (Default) are Eq. (19-24) for Charge and Park, Eq. (25-30) for QC. Constraints related to VWG When Staying (QC 100%) are Eq. (19-24) for Charge, Eq. (25-30) for Park and QC.

According to your comment, we add descriptions about V2G scenarios in the revised manuscript, Methods, page 18, lines 511-512:

Supplementary Table. 7 shows the difference of constraints used in optimizing V2G When Charging and V2G When Staying.

Supplementary material, page 19, Supplementary Table. 7:

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

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

We modify Figure 2 in the revised manuscript, Results, page 4:

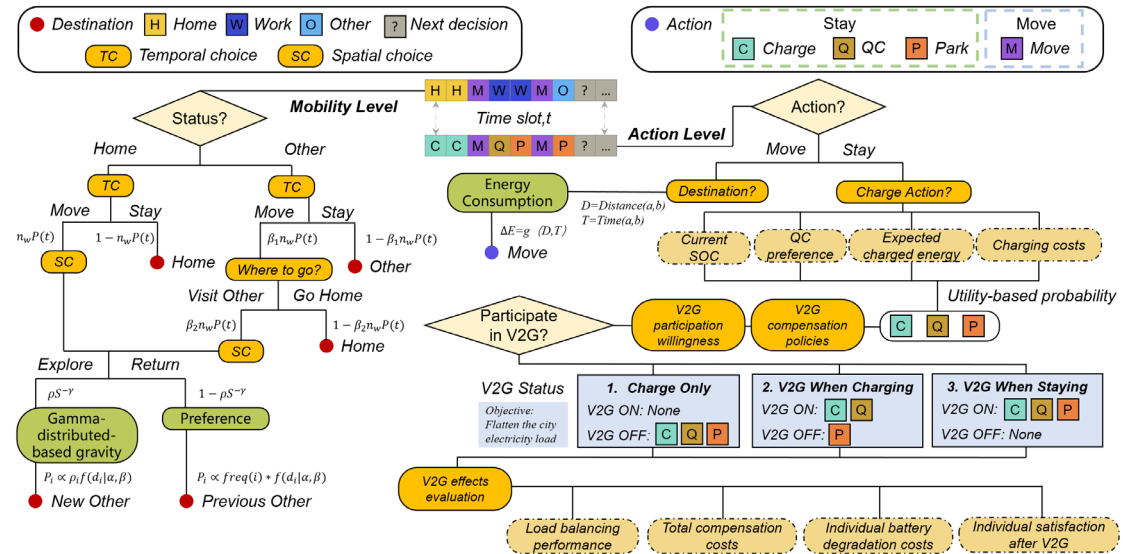


Fig. 2: An overview of the MOVC framework. The MOVC framework assigns mobility and charging information for each time slot. In the Timegeo²²-based mobility model, temporal

choices of PEVs are controlled by user-specific parameters, including home-based hours (n_w), dwell rate (β_1), and burst rate (β_2). Spatial choices are made by combining the gravity model⁶⁶ and the gamma-distributed-based EPR mechanisms. In the charging model, charging decisions are simulated via utility functions of multiple factors³³. A compensation mechanism⁴⁰ is utilized to simulate users' participation in three different V2G strategies. The objective is designed to flatten the city's electricity load as much as possible. The effects of V2G are evaluated considering load balancing performance, total compensation costs⁴⁰, battery degradation costs⁵⁰ and users' satisfaction after V2G⁶¹. Energy consumption during movement is estimated based on travel distance and travel time⁶⁷.

Comment 5

In Figure 2, at the node “stay”, with no charge, why is there V2G off or V2G on? When the charge occurs, there is also V2G off or V2G on. What is the difference between them?

Response to Comment 5

Thank you very much for the comment. As we illustrate in your Comment 4, the main difference of V2G actions stems from the setting of V2G When Staying, which requires users to change their no-plug-in actions in Park to plug-in. This setting is based on our observation that Park activities account for about 80% of all Stay activities of P-PEVs, while charging-related activities (spontaneous plug-in actions) account for only 10-20%. Our motivation for designing V2G When Staying is to illustrate how much load balancing potential may be released if the parking time of vehicles is fully engaged. Similar insights are proposed by the work of Fu et al.³, where they demonstrate the energy-management-related benefits brought by “plugging in upon parking and keeping connected until leaving”. You are kindly referred to Figure.R4 for more details about the difference of V2G scenarios:

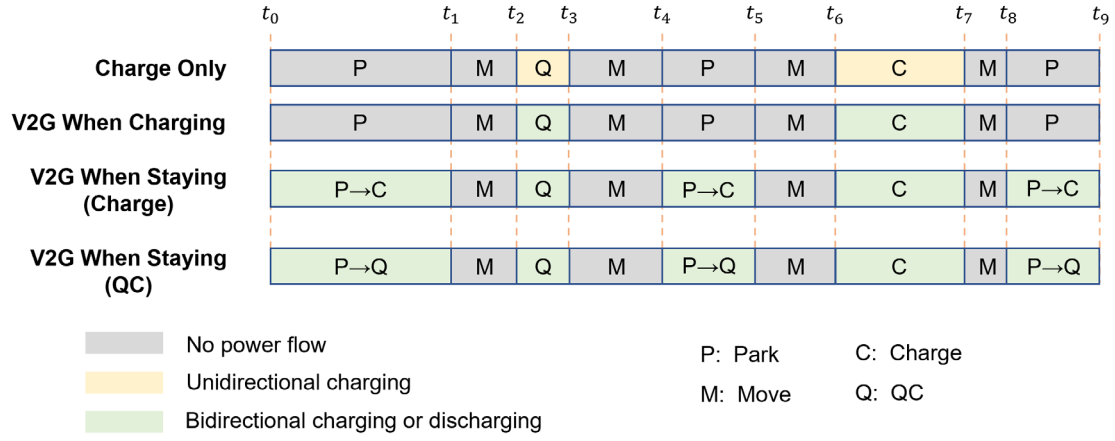


Figure.R4: Illustration of different V2G scenarios. t_i denotes the time slot of each arrival/departure. P, M, C, Q denote possible activities of P-PEV users, namely Park, Move, Charge and QC. Different situations of power flow are marked with corresponding colors. Right arrows \rightarrow mean the users' original decisions are changed from Park to Charge/QC in V2G When Staying.

According to your comment, we add descriptions in the revised manuscript, Results, page 4, lines 109-128:

For charging modelling, we consider all possible actions of P-PEVs as (1) Move, where the vehicles travel in the city. (2) Park, where users intend to park the vehicles without plugging chargers in them. (3) Charge, where users park the vehicles and plug the slow chargers into their slow charging sockets. (4) QC, where users park the vehicles and plug the QC chargers into their QC sockets. The latter three cases are summarized by the term “Stay” to represent a general situation where P-PEVs are immobile. Energy consumption during Move is calculated by the drivetrain losses⁶⁷ (see Methods). If a vehicle is in a Stay, MOVE uses a utility-based model³³ to specify its action of Park, Charge, or QC. The decisions are affected by multiple realistic factors, *e.g.*, the vehicle's current state of charge (SOC), the user's preference towards QC, the expected energy that can be obtained in the Stay depending on Stay duration along with available charging modes, and charging costs depending on the current time period and location (public or private). One can refer to Methods for details. Notably, we assume that all the decisions of Park, Charge and QC are made upon arrival of Stay-s. Next, MOVC simulates P-PEVs' participation in V2G by considering economic subsidies. The compensations required for obeying V2G scheduling are modelled as normal distributions⁴⁰. Given a specific compensation rate γ , users will decide which V2G scenario to join depending on whether their expected economic gains are met (see Methods, Supplementary Note. 9 and Supplementary Fig. 12). We design three scenarios, namely (1) Charge Only, where V2G is disabled and only unidirectional energy flows from the power grid to the vehicles, (2) V2G When Charging

(VWC), where V2G is enabled in Charge and QC actions, and (3) V2G When Staying (VWS), where users are required to plug the chargers in their vehicles to enable V2G once they are in Stay. In other words, V2S compulsorily switches the vehicles' states in Park from isolated to connected (to the grid). Accordingly, users participated in VWS will receive higher compensation compared to those in VWC.

We modify Figure 2 as shown in your comment 4, response letter.

Comment 6

The analysis of PEV load profiles is limited to two weeks representing summer and winter. How do these results account for the significant seasonal and temporal variations in electricity demand throughout the year? In addition, the number of PEVs would have a significant increase in the next years. The scale of PEVs should be considered for the result analysis.

Response to Comment 6

We appreciate your suggestions. We agree that the analysis of the selected two weeks may not represent all-year electricity demand. Our initial motivation is to select one week with all-year peak value and another week with a general lower demand. We label the selected weeks as “Winter” and “Summer” in Figure.R5. 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 values throughout the year. According to your comment, we add experiments related to “Spring” and “Fall” week as shown in Figure.R5. The peak value of “Fall” week is the median of the peak values of 53 weeks. The “Spring” week has similar patterns to the “Fall” week but has greater fluctuations. We combine your comment and Reviewer 3, Comment 6 to present new results in the revised manuscript.

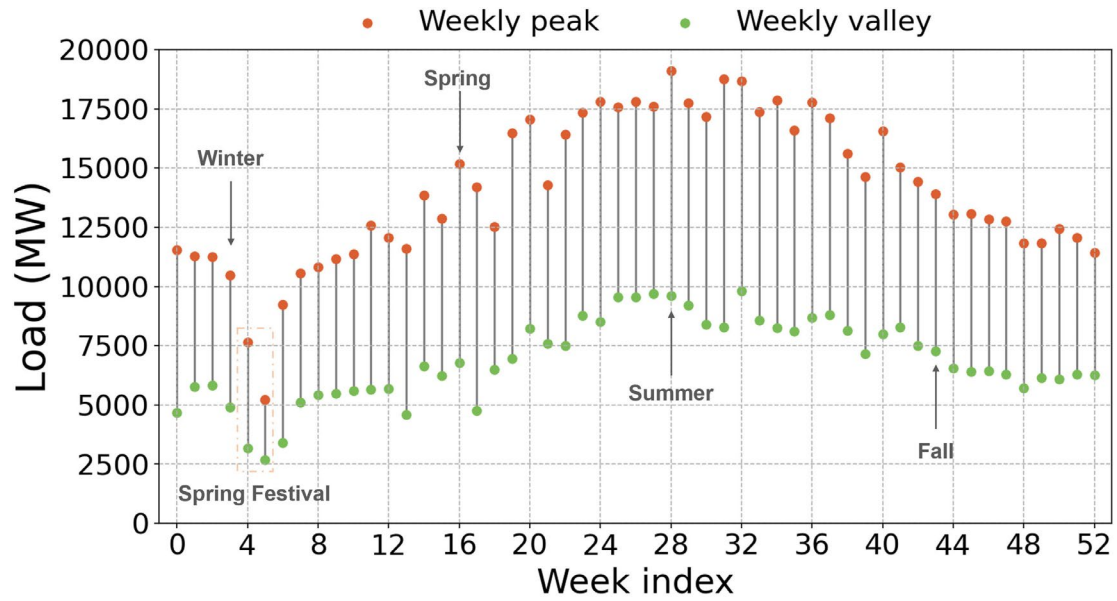


Figure.R5: 2019 Shenzhen weekly load profiles. Weeks “Winter” and “Summer” are analyzed in the original manuscript. Weeks “Spring” and “Fall” are additionally analyzed according to Reviewer 2, Comment 6. We avoid weeks related to Chinese public holidays.

Regarding your concern about the scale of P-PEVs, we agree that the increase of the number of P-PEVs should be discussed in our work. Therefore, we study the extreme case in our results, where we compare the power balancing performance when all P-PEVs are engaged in V2G When Staying. The scale of P-PEVs is analyzed by comparing the V2G effects of 780,000 P-PEVs (Shenzhen), 640,000 P-PEVs (Beijing), 1,020,000 (Shanghai), and 650,000 (Guangzhou), corresponding to their P-PEV holdings in 2024.

To summarize, according to your comment, we add results of V2G analysis across all seasons in the revised manuscript, Results, pages 8-10, lines 203-224:

V2G optimization considering variations of seasonal load and compensation intensities. In this section, we discuss how distinct compensation policies influence user participation in various V2G programs and how the proposed V2G strategies stabilize city electrical loads across different seasons with varying levels of fluctuation. For the former, we set up four compensation rates $\gamma = 0.5, 1.0, 1.5, 2.0$. Considering the discrepancy of participation duration between VWC and VWS, unit discharge compensation is set as 0.015, 0.03, 0.045, 0.06 RMB (around 0.002-0.008 USD)/kWh and 0.075, 0.15, 0.225, 0.3 RMB (around 0.01-0.04 USD)/kWh, respectively. It should be noted that these values represent only the policy-based V2G subsidies and do not include users’ revenue from discharging via feed-in tariffs. The specific participation mechanism is given in Methods and Supplementary Fig. 12. For the latter, we select four weeks from 2019 Shenzhen electricity load

profile to represent typical urban loads at all seasons (see Supplementary Note. 10 and Supplementary Fig. 13). The selected summer load contains an annual peak of 19,219 MW, and the fall load has the median weekly peak observed throughout all weeks in 2019. The spring load has a comparable peak to the fall one but exhibits a greater peak-valley difference. The winter load represents an overall lower power demand. The city load fluctuation is evaluated via peak- valley difference ratio (PVDR), defined as $PV = \frac{L_p - L_v}{L_v}$, where L_p and L_v represent peak load and valley load, respectively.

First, we demonstrate the city load curves under different compensation intensities in Fig. 6. The given curves reflect the combined effects of both VWC and VWS-involved vehicles. However, since the V2G effects of VWC are trivial, we illustrate the fitted relationship between PVDR and the number of P-PEVs in VWS in the right boxplots. With the increase of compensation intensity, the weekly median of the city load PVDR gradually decreases from 0.428, 0.362, 0.408, 0.396 to 0.118, 0.097, 0.058, 0.005 at all seasons, respectively. Additionally, the results indicate that every 100,000 P-PEVs participating in VWS scheduling are expected to mitigate the median city load PVDR by 0.06-0.08, accounting for 16.36%-20.20% of the metric of original loads.

In the revised manuscript, Results, page 9, Figure 6:

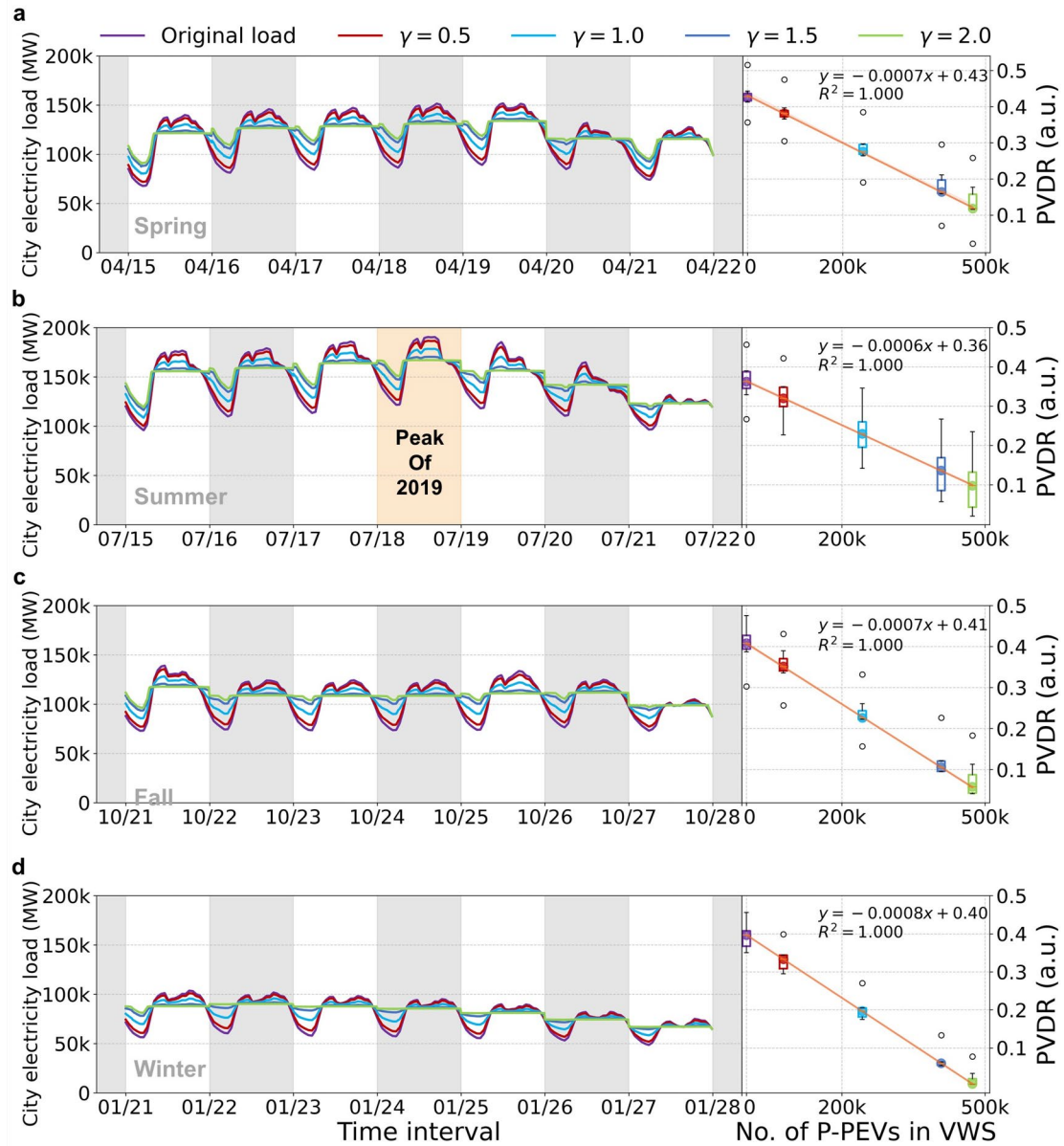


Fig. 6: Illustration of Shenzhen load balancing performance of P-PEVs under different compensation rates. a,b,c,d, The weekly curves of Shenzhen's original load and the load with V2G compensation ratio $\gamma = 0.5, 1.0, 1.5, 2.0$, respectively. The simulation is conducted in different overall load levels (spring, summer, fall, winter), where the annual peak occurs in summer. The box plots compare the daily peak-valley difference ratio (PVDR) of different curves. Note that for each γ , the corresponding load curve is affected by both P-PEVs involved in VWS and ones in VWC, and the numeric relationship is shown in Fig. 7a. However, since the contributions of VWC are trivial (see Fig. 7b, Fig. 7c), we directly fit the relationship between PVDR and the number of P-PEVs involved in VWS in the boxplots.

By combining your comment and Reviewer 3, Comment 6, we add results of V2G evaluation metrics across all seasons in the revised manuscript, Results, pages 10-11, lines 225-247:

Next, we further discuss the comparison of several important characteristics between VWC and VWS, including participation scale, peak-shaving and valley-filling effects, battery degradation costs, user satisfaction and total compensation costs. The numbers of P-PEVs involved in two V2G scenarios are provided in Fig. 7a. The differential impacts of varying participation scales in VWC and VWS scenarios on city weekly load peak shaving and valley filling effects are shown in Fig. 7b, Fig. 7c. As previously discussed, VWC exhibits very limited load balancing potential since it cannot schedule the substantial duration of Park actions. Typically, VWC achieves weekly peak shaving below 10 MW and valley filling under 80 MW. Considering that ToU inherently guides the Charge Only load to off-peak periods, the operational value of VWC is further impaired. In contrast, VWS enables full scheduling flexibility of Park duration and demonstrates pronounced load balancing efficacy, achieving peak shaving of 1330-2332 MW and valley filling of 1475-2320 MW under maximum compensation. Fig. 7e illustrates the distributions of individual weekly discharged energy, where the average and median values of VWC reach 6.70-8.98 kWh and 0 kWh, and the metrics of VWS reach 212.42, 213.19, 198.51, 131.66 kWh and 220.19, 220.29, 206.76, 133.84 kWh, respectively. Fig. 7f gives distributions of battery degradation costs⁵⁰, where the average and median costs of VWC locate between 2.95-3.22 RMB (around 0.41-0.44 USD)/week and 2.22-2.32 RMB (around 0.31-0.32 USD)/week, and the values of VWS reach 31.15, 29.41, 25.37, 15.72 RMB (around 4.30-2.17 USD)/week and 31.82, 29.95, 26.30, 16.20 RMB (around 4.39-2.23 USD)/week, respectively. Regarding users' satisfaction after V2G activities⁶¹ shown in Fig. 7g, the median values of VWC generally exceed 80%, while the metrics of VWS are inferior, reaching 50.73%, 47.81%, 67.03% and 87.30%, respectively. As important indicators of economic compromise, the values of total compensation costs⁴⁰ paid to P-PEV users are illustrated in Fig. 7d. Full investment ($\gamma = 2.0$) requires costs of 18.71M-30.29M RMB (around 2.58M-4.18M USD) per week. The compensation costs exhibit diminishing marginal returns, where approximately 25% of maximum investment yields peak shaving of 770-1196 MW and valley filling of 925-1195 MW, while 66% investment achieves 1172-2008 MW and 1475-1998 MW, respectively, approaching the performance plateau observed at full investment.

In the revised manuscript, Results, page 10, Figure 7:

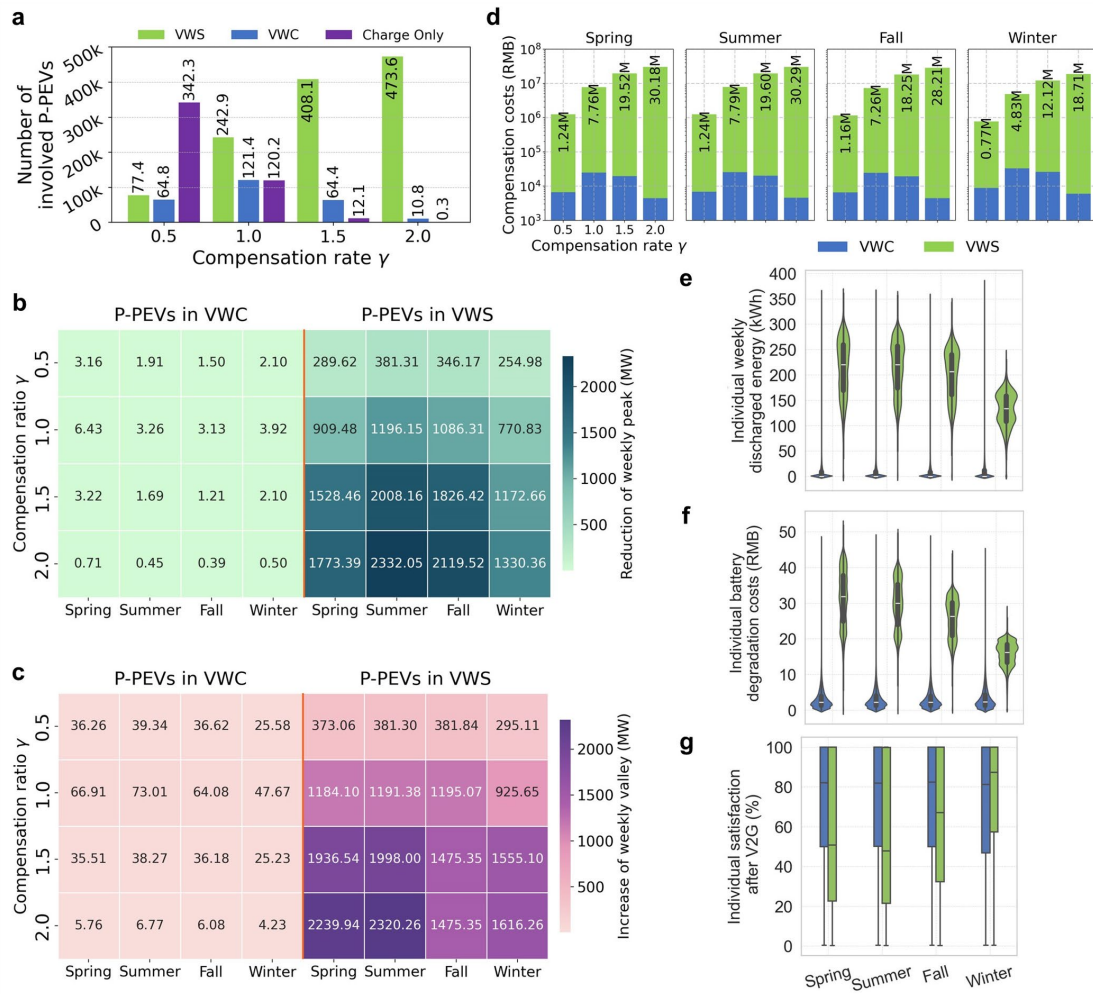


Fig. 7: Evaluation of V2G effects under V2G When Charging (VWC) and V2G When Staying (VWS) with different compensation rates. a, Numbers of P-PEVs participating in different V2G scenarios. **b,c**, Contributions of VWC and VWS in reducing the city weekly peak load and increasing the city's weekly valley load, respectively. **d**, Total compensation costs paid for users involved in VWC and VWS, respectively. **e**, Distributions of individual weekly discharged energy under VWC and VWS. **f**, Distributions of individual battery degradation costs under VWC and VWS. **g**, Boxplots of individual satisfaction after a V2G activity.

By combining your comment and Reviewer 3, Comment 6, we add results of V2G optimization considering higher user satisfaction across all seasons in the revised manuscript Results, page 11-12, line 248-268:

V2G scheduling considering higher user satisfaction. As discussed above, while VWS demonstrates superior load balancing performance, this comes at non-trivial economic costs. Furthermore, user satisfaction with V2G dispatching remains low across all seasons except winter. This section investigates how V2G evaluation metrics evolve when enforcing minimum user

satisfaction thresholds. Specifically, we denote user satisfaction after a V2G activity as ϵ^{usr} . We consider VWS, $\gamma = 2.0$ scenario with three thresholds $\epsilon^{usr} \geq 40\%$, $\epsilon^{usr} \geq 60\%$, $\epsilon^{usr} \geq 80\%$, respectively. Experiments are conducted for summer, spring and fall loads.

As shown in Fig. 8a, stricter satisfaction requirements differentially impact valley-filling performance across seasons. Notably, peak-shaving capability remains largely unaffected, except for two high-demand summer days when ϵ^{usr} is constrained to exceed 80%. In Fig. 8b, as ϵ^{usr} thresholds increase, the median PVDR values of three seasons are gradually lifted from 0.118, 0.098, 0.058 to 0.199, 0.184, 0.120, respectively. In Fig. 8c, the average weekly battery degradation costs across three seasons decline from 31.15, 29.41, 25.37 RMB (around 4.30-3.50 USD) to 17.47, 19.21, 17.20 RMB (around 2.41-2.37 USD), respectively. Similarly, the median values decrease from 31.82, 29.95, 26.30 RMB (around 4.39-3.63 USD) to 16.02, 17.94, 15.97 RMB (around 2.21-2.20 USD), respectively. In Fig. 8d, user satisfaction exhibits a significant improvement, with median values progressively increasing from 50.73%, 47.81%, 67.03% to 83.07%, 80.57%, 87.27%, respectively. In Fig. 8e, the total compensation costs for $\epsilon^{usr} \geq 40\%$ and $\epsilon^{usr} \geq 60\%$ exhibit both increases and decreases compared to the default scenario, whereas the total costs for $\epsilon^{usr} \geq 80\%$ consistently decrease relative to the baseline. Specifically, the reductions for $\epsilon^{usr} \geq 80\%$ reach 4.09M, 1.43M, and 2.24M RMB (around 0.56M-0.31M USD), respectively.

To briefly summarize, maintaining high user satisfaction leads to a 30-40% reduction in battery degradation costs and a 5-13% decrease in total compensation costs, though with modest trade-offs in valley filling efficiency and limited peak shaving performance under extreme high-load conditions.

In the revised manuscript, Results, page 11, Figure 8:

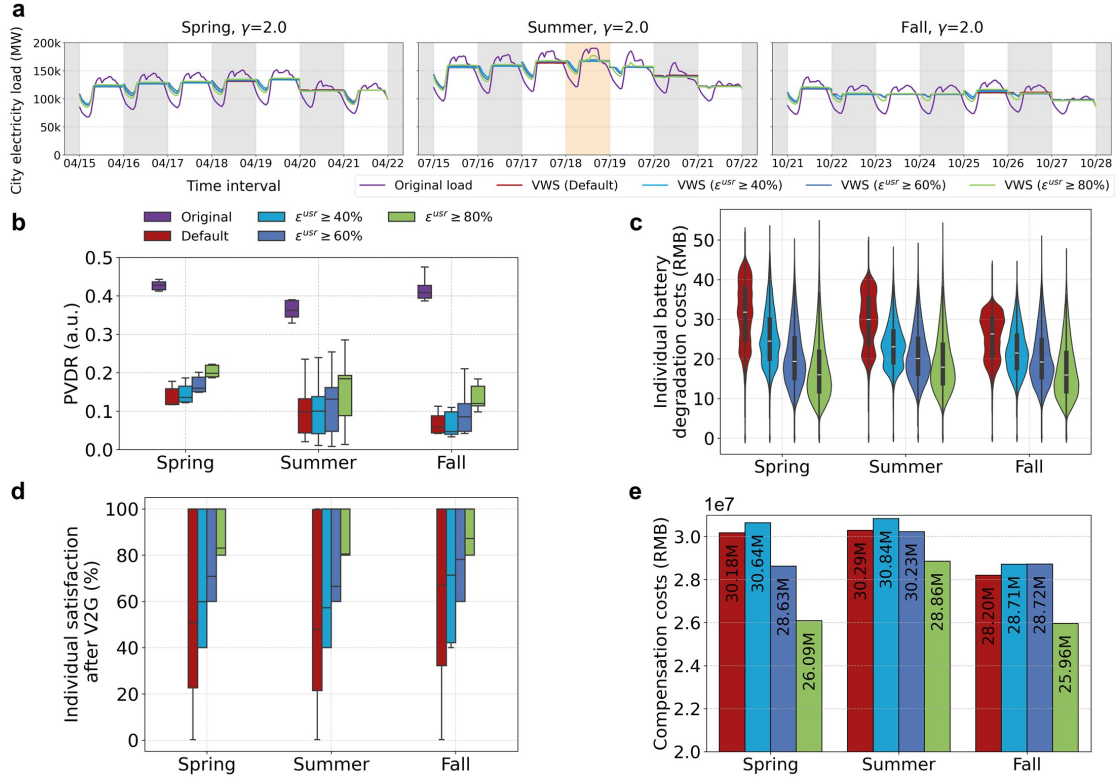


Fig. 8: Evaluation of V2G effects under VWS, $\gamma = 2.0$ considering higher minimum satisfaction after V2G activities. **a**, The weekly curves of Shenzhen's original load and the load under VWS, $\gamma = 2.0$, with minimum satisfaction $\epsilon^{usr} = 40\%, 60\%, 80\%$, respectively. The winter load is not considered since the median of the default VWS satisfaction in winter already exceeds 80% (see Fig. 7f). **b**, PVDR with different minimum values of ϵ^{usr} in three seasons. **c**, Distributions of individual battery degradation costs with different minimum values of ϵ^{usr} in three seasons. **d**, Boxplots of user satisfaction after V2G activities. The whiskers of some boxplots merge with their quartile markers (Q1/Q3) due to particular data distribution features. **e**, Total compensation costs paid with different minimum values of ϵ^{usr} .

By combining your comment with Reviewer 1, Comment 2 and Reviewer 3, Comment 6, we add results of V2G optimization considering higher user satisfaction across all seasons in the revised manuscript, Results, page 12, Figure 9:

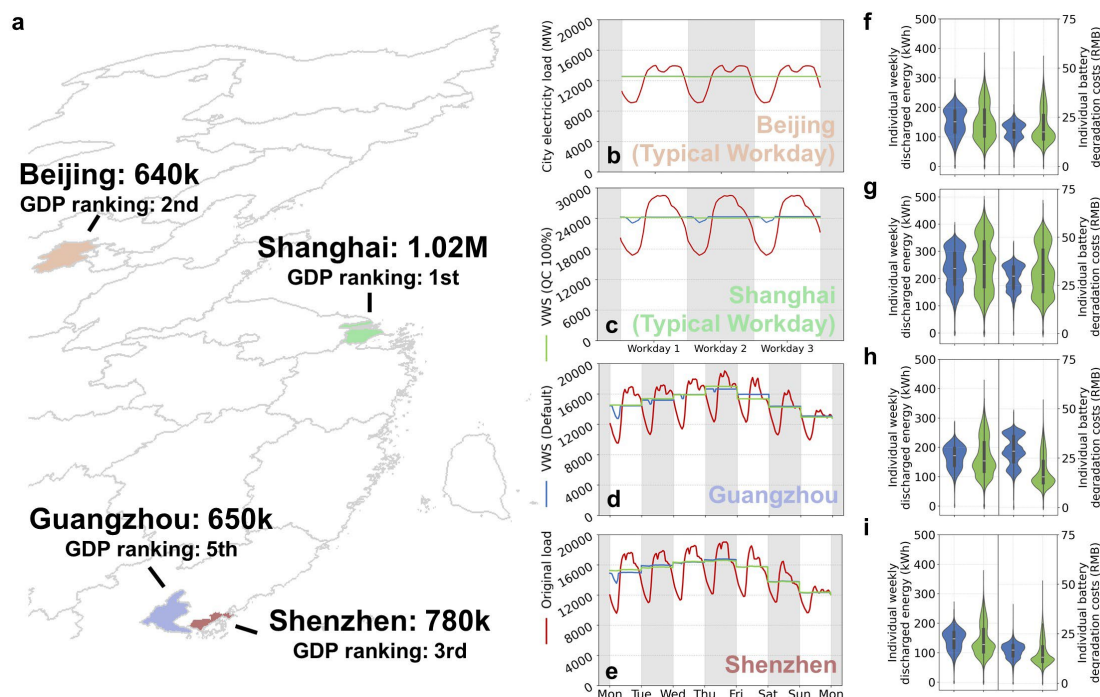


Fig. 9: V2G evaluation in Beijing, Shanghai, Guangzhou and Shenzhen based on MOVIC. a, P-PEV holdings of the four cities in 2024, and their GDP rankings in China's cities. b,c,d,e, Load balancing results of the four cities under VWS (Default) and VWS (QC 100%), respectively. Since we do not have the annual load data of Beijing and Shanghai, we illustrate their results in typical workdays. For Guangzhou and Shenzhen we select a week with the annual peak. f,g,h,i, Distributions of individual weekly discharged energy and battery degradation costs under the two V2G scenarios.

In the revised manuscript, Results, pages 12-13, lines 269-306:

Extreme V2G imagination for China's megacities. MOVIC has the transferability to other urban contexts. The generalization ability is utilized to explore two substantial problems: (1) How V2G effects evolve with larger P-PEV penetration rates, and (2) whether city-level upgrades in residential distribution grids for accommodating comprehensive QC are cost-effective. To this end, we illustrate the estimated P-PEV holdings of Beijing, Shanghai, Guangzhou and Shenzhen in 2024, reaching 640,000, 1,020,000, 650,000 and 780,000, respectively. The selected cities represent the most economically developed cities in China. We leverage 2022 BPTD to depict mobility and charging patterns of Beijing P-PEVs following the collective paradigm introduced in this work, as provided in Supplementary Fig. 2,4,6, 16-19. We use 2022 BPTD to generate vehicles for Beijing's

case, and use 2022 SPTD for Shanghai, Guangzhou and Shenzhen. We compare two V2G scenarios, namely (1) VWS (Default), where the settings are identical to the ones in Fig. 6-Fig. 8 except for the number of involved P-PEVs. (2) VWS (QC 100%), where P-PEVs are allowed to use QC even in *Home* regions, and all the Park actions will perform VWS with QC power levels.

As shown in Fig. 9b-Fig. 9e, the discussed two scenarios exhibit similar load balancing effects with larger P-PEV penetration rates. Specifically, the weekly median PVDR values for all cities under VWS (Default) and VWS (QC 100%) are 0.0015, 0.0487, 0.0262, 0.0109 and 0.0015, 0.0015, 0.0016, 0.0092, respectively. Compared with PVDR of original loads (0.35, 0.41, 0.36, 0.36), extra benefits only reach 0%, 12%, 7%, 5%, respectively. In Fig. 9f-Fig. 9i, after full QC upgrades, the median values of weekly discharged energy in Beijing, Guangzhou and Shenzhen decrease from 151.25, 171.66, 147.04 kWh to 140.47, 154.21, 128.10 kWh, respectively. However, in Shanghai where the overall city load is higher, the values increase from 237.30 to 252.20 kWh. Regarding battery degradation costs, the median values of Beijing, Guangzhou and Shenzhen decrease from 18.55, 28.54, 16.78 RMB (around 2.56-2.31 USD) to 17.54, 15.57, 13.19 RMB (around 2.42-1.82 USD), respectively. In contrast, the metric of Shanghai slightly increases from 29.87 to 30.71 RMB (around 4.12 to 4.24 USD). Another noteworthy effect is observed in Shenzhen's case. With the deeper P-PEV penetration, the median weekly discharged energy decreases from 220.29 kWh (Fig. 7d, summer) to 147.04 kWh-128.10 kWh, and the median battery costs also decrease from 29.95 RMB (around 4.13 USD, Fig. 7e, summer) to 16.78-13.19 RMB (around 2.31-1.82 USD). Notably, the inequity of both metrics in all cities grows significantly. For the weekly discharged energy, the standard deviation of all individuals increases from 46.05, 70.17, 39.00, 35.43 kWh to 60.85, 97.37, 65.74, 54.59 kWh, respectively. For the battery degradation costs, the metric goes up from 4.42, 6.82, 7.71, 3.83 RMB (around 0.61-0.53 USD) to 8.61, 12.61, 7.98, 6.30 RMB (around 1.19-0.87 USD).

The results indicate that as the number of V2G-related vehicles increases, cities with mild load profiles (Beijing, Guangzhou, Shenzhen) exhibit lighter optimization burdens to each vehicle, leading to lower discharged energy and battery degradation costs. In such cases, QC upgrades yield marginal additional benefits. Conversely, in cities with higher load levels (*e.g.*, Shanghai), achieving comparable effects requires a higher penetration of P-PEVs. Additionally, expanding QC infrastructure significantly increases the variance of both discharged energy and degradation costs

(Fig. 9f-Fig. 9i), enlarging the standard deviation by 32%-69% and 4%-95%, respectively.

Customized V2G incentive policies may be required to mitigate such inequity. Given the substantial

costs for residential grid upgrades, city policymakers should carefully weigh the trade-offs between

investments and expected benefits.

We add the description in the revised manuscript, Discussion, pages 13-14, lines 335-338:

The weekly median PVDR of the original city load can be decreased up to 72.4%-98.7% seasonally,

where VWC and VWS-related vehicles release tens of MW and 2,000 MW load balancing capacities,

respectively, at compensation costs of 18.71M-30.29M RMB (around 2.58M-4.18M USD).

Response to Reviewer #3

The manuscript "Unlocking the Potential of Private EVs for Shaving the Electricity Peak Load of the City" presents a study on mobility and V2G coordination to optimize the charging and discharging patterns of EVs in a mega-city in China. While the data analysis is interesting, the V2G strategy has some drawbacks, the manuscript is difficult to read, and several concepts are unclear. Based on these concerns, I do not recommend accepting the paper in its current form.

Dear Respected Reviewer,

Thank you very much for taking valuable time to review our paper. We sincerely appreciate your recognition of the merits of our data analysis methodology, and we are equally grateful for your constructive critique regarding the limitations in our V2G strategy design, paper readability, and conceptual exposition.

In the following, we address your comments point by point to further clarify related concepts and to illustrate our modification in experimental design and proofreading. We highlight the cited content from the manuscript and Supplementary files in blue and the corresponding revisions in the manuscript and Supplementary files in yellow. We hope this revised version will meet your expectations of a qualified research paper as a decent publication for Nature Communications.

Comment 1

Introduction: The identified research gap related to private V2G does not seem entirely accurate. Previous studies have already addressed this issue in various ways (e.g., 10.1109/OJVT.2023.3323087, 10.1109/ACCESS.2022.3185236, 10.1016/j.scs.2021.102744). Please refine the discussion on research gaps and clarify the contributions of your study.

Response to Comment 1

Thank you very much for the comment. We admit that the research gap analyzed in the first manuscript is not accurate. The works you specified are important in V2G-related research, with contributions in heuristic V2G algorithms and user-oriented V2G scheduling strategies. Compared with these works, our research mainly contributes to travelling behavior considerations and more detailed user-centric concerns. Therefore, we rewrite the part of the literature review in the revised manuscript, Introduction, page 2, line 44-56:

Existing literature mainly considers V2G based on the mobility or charging behavior of PEVs. The mobility models are proposed to distinguish the spatial movement patterns of individuals. Those methods can be divided into Origin-Destination (OD) analysis¹⁹⁻²¹, Markov Chain-based models²²⁻²⁴, OD-based trip chain models²⁵⁻²⁷, graph-based models²⁸, etc. The charging models include more

subdivisions, covering large-scale valley filling ^{29,30}, load forecasting ³¹, user-oriented charging behavior analysis and scheduling strategies ³²⁻³⁵, V2G arbitrage ³⁶⁻³⁷, user willingness of participating V2G and incentive mechanisms ³⁸⁻⁴⁶, range anxiety ⁴⁷, battery degradation effects ⁴⁸⁻⁵², V2G optimization relaxation ⁵³⁻⁵⁵, *etc.* Multiple methodologies are leveraged to address these problems with high accuracy or efficiency, including mixed-integer linear programming (MILP) ^{37,47}, heuristic algorithms ^{35,36,59}, stochastic programming ^{40,41,51}, deep reinforcement learning ⁵⁶⁻⁵⁸, user surveys ^{38,42,43,45}, simulations ^{48,52}, experimental measurements ⁴⁹, *etc.* However, existing works have their scopes narrowed in single or several subareas by studying basically microgrids or small PEV fleets, with megacity-level user behaviors being considerably underinvestigated. This noticeable gap is critically prohibitive for megacity policymakers to consider V2G promotion and their sustainable integration into many energy infrastructures.

Comment 2

Quick Charging (QC): The manuscript considers the use of quick charging, but the assumed power level is not specified. Please provide this information. Additionally, you state: "Assumed that all vehicles can perform QC on-site when needed, the number of QC activities takes possession of 10.34% of all charging events." This statement is unclear—please clarify the reasoning behind this assumption.

Response to Comment 2

Thank you very much for proposing your concern about our Quick Charging (QC) settings. In fact, we provided information about slow charging and QC power levels in Supplementary Table 1, Supplementary Information page 2:

Supplementary Table 1: 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

The statement "Assumed that all vehicles can perform QC on-site when needed, the number of QC activities takes possession of 10.34% of all charging events" is based on the simulation of our charging model. Specifically, in our first manuscript, Methods, page 15, lines 388-389:

A vehicle will start QC if its planned charging time T_p and its stay duration ΔD satisfies $T_p * 0.8 \geq \Delta D$.

In our charging-related experiments, the probability that a user decides to charge is described in our first manuscript, Methods, page 14-15, Eq. (3-6):

$$P(\text{charge}|SOC_{in}, \Delta D, C_u) = \frac{P(SOC_{in}, \Delta D, C_u|\text{charge})P(\text{charge})}{P(SOC_{in}, \Delta D, C_u)} \quad (3)$$

$$= \frac{P(SOC_{in}|\text{charge})P(\Delta D, C_u|\text{charge})P(\text{charge})}{P(SOC_{in})P(\Delta D, C_u)} \quad (4)$$

$$= \frac{P(SOC_{in}|\text{charge})P(\text{charge}|\Delta D, C_u)}{P(SOC_{in})} \quad (5)$$

$$P(\text{charge}|\Delta D, C_u) = f(\Delta D, C_u) = \frac{C_u * \Delta D}{30 * 144} \quad (6)$$

In our implementation, a random variable p uniformly sampled from $[0, 1]$ is used for simulating charging decisions. If $p \leq P(\text{charge}|SOC_{in}, \Delta D, C_u)$, a user will decide to charge. Here, the term "charge" means the action of connecting a vehicle to a charging pile, and it contains the situations of both slow charging and QC. Therefore, when a user decides to charge, we use $T_p * 0.8 \geq \Delta D$ to further determine whether he will choose QC. In our original experiments, we assume a base case where QC actions completely depend on the condition $T_p * 0.8 \geq \Delta D$. Additionally, we assume several cases QC 20%, QC 40%, QC 60%, QC 80% and QC 100%. In those cases, QC piles are constructed to first cover the stay regions with the highest number of visits, and QC actions will occur either $T_p * 0.8 \geq \Delta D$ or (the current stay region is covered with QC piles) is satisfied. In the base case, we observe in our simulation that 10.34% of the charging actions further enter the QC branch since these actions satisfy $T_p * 0.8 \geq \Delta D$. In other cases, the satisfaction of the two conditions becomes overlapped. We thus state in our first manuscript, Methods, page 15, lines 412-416:

The proportion of QC activities reaches 23.29%, 42.66%, 63.65%, 83.09%, and 100% of all charging behaviours, respectively, which reflects the combination of users' spontaneous QC actions in non-labelled stay regions and mandated QC actions in labelled stay regions.

Hopefully, these explanations resolve the concerns you raised.

However, thanks to your comment and Reviewer 1, Comment 3, we realize that the design of our original charging model has some drawbacks. Specifically, the model does not consider user cost factors and range anxiety impacts when simulating charging behaviors. Also, the statement "Assumed that all vehicles can perform QC on-site when needed" does not correspond to the real-world situation, where users can hardly perform QC in their homes due to the limited capacities of residential distribution grids. Therefore, we modify the research of Liu et al⁴. and enhance the reality of our charging model following the utility-based paradigm.

We introduce the modified charging behavior model in the revised manuscript, Methods, pages 15-16, lines 412-454:

User charging behavior modelling. We modify the utility-based charging behavior method³³ to establish our charging decision model. Available charging modes upon Stay arrivals are defined as $C_{i,s} = \{cm: L_{cm}^{i,s} = 1\}$ for the s -th Stay of the i -th P-PEV. We use $cm = 0,1,2$ to refer no-charging, slow charging and QC, respectively. Denote $\exp(V_{L_{cm}^{i,s}})$ as the indirect utility of charging mode $L_{cm}^{i,s}$. The choice probability of each charging mode is expressed as follows³³:

$$P_{\text{choice}_{i,s}}(L_{cm}^{i,s}) = \begin{cases} \frac{\exp(V_{L_{cm}^{i,s}})}{\sum_{j \in C_{i,s}} \exp(V_{L_j^{i,s}})} & \text{if } cm \in C_{i,s}, \\ 0 & \text{if } cm \notin C_{i,s}. \end{cases} \quad (3)$$

For the no-charging (Park) choice, we normalize the indirect utility to 0. Except for the extreme V2G case, we assume that slow charging is always available in all kinds of stay regions (*Home*, *Work*, *Other*), and QC is only available in public location (*Work*, *Other*).

Furthermore, the indirect utility of charging mode $L_{cm}^{i,s}$ is defined as

$$V_{L_{cm}^{i,s}} = \beta_0 + V_{\text{SOC}_{i,s}^e} + V_{R_{L_{cm}^{i,s}}} + V_{\Delta \text{SOC}_{L_{cm}^{i,s}}} + V_{\text{Cost}_{L_{cm}^{i,s}}} + V_{\text{stay}, L_{cm}^{i,s}}, \quad (4)$$

$$cm \in C_{i,s} \setminus \{0\},$$

where β_0 is a constant representing the average effect of the non-listed attributes³³ set as 1 in the base case. Each component V_X refers to the utility of attribute X .

The first component $V_{SOC_{i,s}^e}$ describes the relationship between the current SOC and user i 's corresponding willingness to charge³³, which is modelled as

$$V_{SOC_{i,s}^e} = \begin{cases} +\infty & \text{if } SOC_{i,s}^e \in (0, SOC_{i,s}^{dem}] \\ \beta_{SOC} \cdot \ln \left(\frac{1 - SOC_{i,s}^e}{\left(\frac{1}{SOC_B} - 1\right) SOC_{i,s}^e} \right) & \text{if } SOC_{i,s}^e \in (SOC_{i,s}^{dem}, 1) \\ -\infty & \text{if } SOC_{i,s}^e = 1 \end{cases} \quad (5)$$

Where SOC_B is the range buffer set as 0.3, $SOC_{i,s}^e$ is the current SOC, β_{SOC} is set as 3 in the base case, and $SOC_{i,s}^{dem}$ refers to the moving consumption of the sequential three trips.

The second component $V_{R_{L_{cm}^{i,s}}}$ describes user i 's preference toward high charging power³³, defined as

$$V_{R_{L_{cm}^{i,s}}} = \beta_R \cdot (P_{L_{cm}^{i,s}} - P_{home}), \quad (6)$$

where $P_{L_{cm}^{i,s}}$ is the charging power of charging mode $L_{cm}^{i,s}$, P_{home} is the charging power at *Home*, and β_R is a coefficient. A larger β_R indicates that the user prefers faster charging to reduce the charging duration, and a smaller one shows his tendency to avoid battery degradation. In our base case, β_R is set as 0.

The third component $V_{\Delta SOC_{L_{cm}^{i,s}}}$ refers to user i 's evaluation of potential energy he can obtain during Stay s . Generally, the utility of this attribute is designed to increase when a larger amount of energy can be expected to charge during Stay s ³³. The formulation is

$$V_{\Delta SOC_{L_{cm}^{i,s}}} = \beta_{\Delta SOC} \cdot [1 - (\Delta SOC_{L_{cm}^{i,s}} - 1)^2], \quad (7)$$

where $\Delta SOC_{L_{cm}^{i,s}} = \min\left(\frac{P_{L_{cm}^{i,s}} \cdot T_{i,s}}{E_i}, 1 - SOC_{i,s}^e\right)$ represents the potential increased energy under charging mode $L_{cm}^{i,s}$, $P_{L_{cm}^{i,s}}$ is the corresponding charging power, $T_{i,s}$ is the duration of Stay s , E_i is the battery capacity, and $SOC_{i,s}^e$ is the SOC of P-PEV i upon arrival of Stay s . $\beta_{\Delta SOC}$ is the coefficient set as 3 in the base case.

The fourth component $V_{Cost_{L_{cm}^{i,s}}}$ describes users' sensitiveness of charging costs. Differed from the constant electricity price setting³³, we introduce time-of-use (ToU) price to represent a more realistic cost evaluation. Specifically, the charging cost $Cost_{L_{cm}^{i,s}}$ is designed as

$$Cost_{L_{cm}^{i,s}} = \sum_{t=Ta_{i,s}}^{Td_{i,s}} P_{L_{cm},t}^{ToU} \cdot \Delta SOC_{L_{cm},t}^{i,s} \cdot E_i, \quad (8)$$

where $Ta_{i,s}, Td_{i,s}$ refer to the arrival and departure time slots of user i , Stay s , respectively.

$P_{L_{cm},t}^{ToU}$ denotes the ToU price at time slot t . The utility $V_{Cost_{L_{cm}^{i,s}}}$ is computed following

$$V_{Cost_{L_{cm}^{i,s}}} = -\beta_{Cost} \cdot (Cost_{L_{cm}^{i,s}} - Cost_{home}), \quad (9)$$

where β_{Cost} is set as 0.35 in the base case, and $Cost_{home}$ can also be calculated following Eq. (8).

The only distinction is that $Cost_{L_{cm}^{i,s}}$ uses $P_{L_{cm},t}^{ToU, pub}$ and $Cost_{home}$ uses $P_{L_{cm},t}^{ToU, home}$, representing

different ToU price of public and private charging piles, respectively, and the former is generally

more expensive. The details of ToU price are obtained from charging service providers⁶⁸ and are

given in Supplementary Fig. 8, 9.

The last component $V_{Stay, L_{cm}^{i,s}}$ is added to simulate users' patterns of slow charging decisions. We

assume that either a too short or a too long stay duration will decrease the users' willingness to

perform slow charging. $V_{Stay, L_{cm}^{i,s}}$ is thus given as

$$V_{Stay, L_{cm}^{i,s}} = \begin{cases} \max(\beta_{Stay} \cdot (T_{i,s} - 2) \cdot (T_{i,s} - 10), B_{Stay}) & \text{if } cm = 1, \\ 0 & \text{if } cm \in \{0, 2\}. \end{cases} \quad (10)$$

where the coefficient β_{Stay} and the buffer B_{Stay} are set as -0.2 and -3, respectively.

We provide settings of $\beta_0, \beta_{SOC}, \beta_R, \beta_{\Delta SOC}, \beta_{Cost}$ and β_{Stay} under different charging cases in

Supplementary Table. 5.

Comment 3

Variable Definitions & Notation: Several variables are not explicitly defined, including SOC in, ΔD , and Cuafter. Additionally, after Figure 3, you mention the L-th most visited location, but its meaning is unclear. Does depot refer to charging stations? If so, please use consistent terminology. Moreover, the values of α (alpha) and β (beta) are not justified—please explain their origin.

Response to Comment 3

Thank you very much for pointing out the confusing parts of our variables and terminology. Our explanations are as follows:

(a) Definition of $SOC_{in}, \Delta D$ and C_u

Our Timegeo-based mobility model simulates the travelling patterns of P-PEVs by generating

Move-s and Stay-s probabilistically. Suppose for a P-PEV i , the mobility model gives M Stay-s

in time slots $t = 1, 2, \dots, T$, and there will be a Move between two consecutive Stay-s. We use the subscript to denote the index of the time slot and use the superscript to denote the index of Stay. The SOC information of P-PEV i is thus denoted as $SOC_1, SOC_2, \dots, SOC_T$, and each Stay is denoted as S^1, S^2, \dots, S^M . For a single Stay m covering k time slots, we denote the related time slots as $t_1^m, t_2^m, \dots, t_k^m$. Therefore, SOC_{in} can be interpreted as a set of arrival SOC levels $\{SOC_{in}^1, SOC_{in}^2, \dots, SOC_{in}^M\}$. SOC_{in}^m is defined as follows:

$$SOC_{in}^m = SOC_{t_1^m}. \quad (7)$$

Similarly, ΔD can be interpreted as a set of Stay duration $\{\Delta D^1, \Delta D^2, \dots, \Delta D^M\}$, and ΔD^m is defined as follows:

$$\Delta D^m = t_k^m - t_1^m. \quad (8)$$

The variable C_u means the average number of monthly charging times of each P-PEV. We assume that C_u follows an exponential distribution with the form $P = \lambda e^{-\lambda}$, where we set $\lambda = 1/6.63$. To allocate C_u for N P-PEVs, we first generate travelling patterns via the mobility model. The total energy consumption of each P-PEV depends on driving distance and driving time of its Move-s, which can be calculated according to our first manuscript, Methods, page 15, lines 393-396:

We model the energy consumption during travelling using the drivetrain losses[63], defined as:

$$P_{dr} = \alpha V^3 + \beta V^2 + \gamma V + P_{stop},$$

where V denotes the average velocity of the trip, and we set $\alpha = 9.494 * 10^{-6}, \beta = 1.93 * 10^{-4}, \gamma = 0.0506, P_{stop} = 2.375$.

In our implementation, we randomly generate $C_{u,n}$ for P-PEV $1, \dots, n, \dots, N$ according to $P = \lambda e^{-\lambda}$, and we allocate $C_{u,n}$ to P-PEV n proportionally to its individual energy consumption $E_{dr,n}$. In other words, a P-PEV n with the largest energy consumption $E_{dr,n}$ will be allocated with the largest $C_{u,n}$, and the rest can be done in the same manner.

However, as we explained in your Comment 2, we modify our charging model in the revised manuscript. The new model does not require $SOC_{in}, \Delta D$ and C_u to make charging decisions.

(b) Explanation of the L-th most visited location

For a P-PEV i , the mobility model generates M Stay-s in T time slots as introduced above. Each Stay has an index $I_{S^1}, I_{S^2}, \dots, I_{S^M}$ 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 $|\{I_{S^m}\}|$ satisfies $|\{I_{S^m}\}| \leq$

M . We sort the number of visits of each stay region $V_{I_{st}}, I_{st} \in \{I_{sm}\}$ in a descending order. In our first manuscript, we use the L -th most visited location to describe the visit frequency of sorted $V_{I_{st}}$ of P-PEV i in T time slots. We give an example of L -th most visited location of a single user in Table.R1:

Table.R1: 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

The range of L may vary for different users depending on their travelling patterns. In our first manuscript, Results, page 6, Fig. 3(d), we calculate the summed normalized visit frequency of all P-PEV users.

We add the description in the revised manuscript, Results, page 5, lines 136-137:

(4) the L -th most visited locations, denoting the frequency of all P-PEV users visiting their 1-st, 2-nd, 3-rd, ..., L -th most visited locations, as explained in details in Supplementary Note. 6 and Supplementary Table. 2.

In Supplementary material, page 7, Supplementary Note. 6:

Supplementary Note. 6: Explanations of the L -th most visited location. For a P-PEV i , the mobility model generates M Stay-s in T time slots as introduced above. Each Stay has an index

$I_{S^1}, I_{S^2}, \dots, I_{S^M}$ 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 $|\{I_{S^m}\}|$ satisfies $|\{I_{S^m}\}| \leq M$. We sort the number of visits of each stay region $V_{I_{S^l}}, I_{S^l} \in \{I_{S^m}\}$ in a descending order. The L-th most visited location is introduced to describe the visit frequency of sorted $V_{I_{S^l}}$ 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.

In Supplementary material, page 7, Supplementary Table. 2:

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

(c) Definition of depots

We apologize for causing your misunderstanding. We admit that the term “depot” may be imprecise in this context. In the revised manuscript, we replace “depot” with “parking spot” to denote individual parking spaces within a “parking lot” (the latter referring to the collective facility). In the following, we briefly introduce the reasoning behind introducing parking spots into our research. Since the 2022 Shenzhen P-PEV Trajectories Dataset (2022 SPTD) only contains 36,932 P-PEV users, which is not comparable to Shenzhen P-PEV holding of 480,000, we consider reasonably expanding the number of P-PEVs by considering the parking spot distribution of Shenzhen. In our first manuscript, a depot refers to a parking spot enabling one vehicle to park. The 2019 Shenzhen Parking Census Dataset (2019 SPCD) provides information on 9,748 parking lots in Shenzhen, regarding their longitude, latitude, and the number of parking spots within. Our results in the first manuscript, Results, page 6, Fig. 3(e-g) are proposed for validating the spatial relationship between (1) the parking spots distribution in Fig. 3(e), given by 2019 SPCD, and (2) the visit number distribution of each street in Fig. 3(f), given by 2022 SPTD). The results indicate that streets with a higher number of parking spots generally experience more frequent P-PEV visits. This supports our approach of scaling the number of P-PEVs according to parking spot distribution density. Specifically, we allocate the total P-PEV number (480,000) to all streets based on their parking spot capacity. 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 $N_e * 480000 * \frac{N_p^i}{\sum_i N_p^i}$. As explained above, introducing depot (parking spot) is mainly for expanding the number of P-PEVs. The concept is independent of charging-related ones (*e.g.*, charging stations). According to your comment, we modify Fig. 3e, 3g in the revised manuscript, Results, page 5:

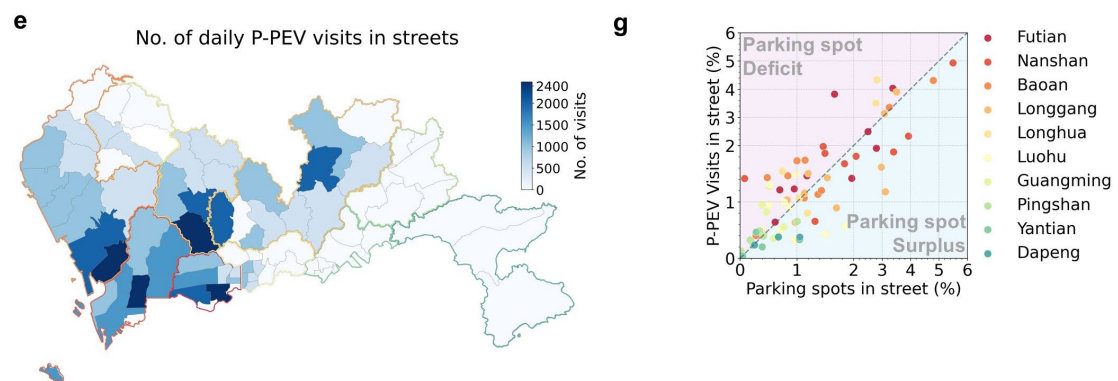


Fig. 3: Validation of the MOVC framework in mobility and charging behavior. e,f, Spatial distributions of parking spot numbers (from 2019 SPCD) and visit numbers (from 2022 SPTD) in Shenzhen. g, Comparison between the number of parking spots and the one of visits, decomposed into each district in Shenzhen.

We add descriptions in the revised manuscript, Results, page 5, lines 138-141:

We compare the identified *Home* labels of P-PEVs with the distribution of parking spots in Shenzhen, according to 2019 SPCD (Fig. 3e-Fig. 3g). The similar distributions of parking spots and *Home* labels support us to expand the number of P-PEVs from 36,932 to larger values according to the proportions of street-level parking spots, as explained in Supplementary Note. 7.

In Supplementary material, page 8, Supplementary Note. 7:

Supplementary Note. 7: Expanding the number of P-PEVs. 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 $N_e * 480000 * \frac{N_p^i}{\sum_i N_p^i}$. 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 Transport[2]. The district-level numbers of parking lots are provided in Supplementary Table. 4.

Considering Reviewer 1, Comment 2, we add experiments of Beijing P-PEVs. The expanding method is similar to Shenzhen. In Supplementary material, page 8, Supplementary Table. 4:

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

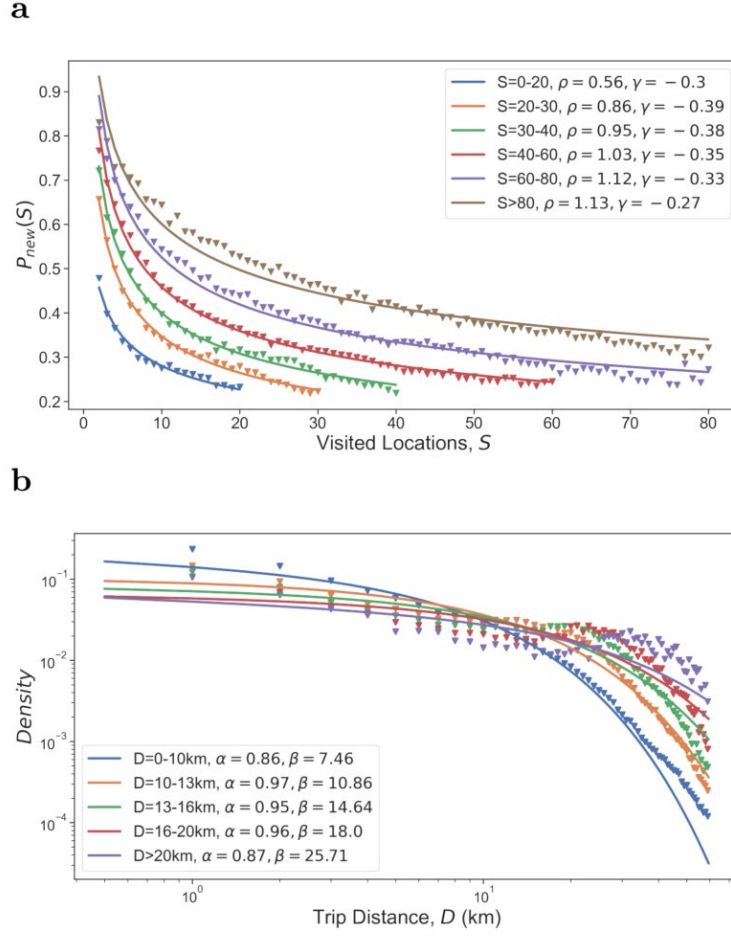
(d) Definition of alpha (α) and beta (β)

The two variables are proposed to describe the spatial decisions of P-PEVs. According to the rank-based exploration and preferential return (r-EPR) mechanism¹, a user will first decide whether to explore a new location or return to a previously visited one (you are kindly referred to Responses, Reviewer 2 Comment 2, page 32-33, line 630-643 for details). Sequentially, the mobility model determines which place the user will move to. r-EPR assumes that the locations closer to the user's current location will have larger probability of being visited. We describe the mechanism using the gamma distribution⁵, following:

$$f(x|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad (9)$$

where x denotes the distance between all possible locations to be visited and the current location, and parameters α, β are the shape parameter and the rate parameter⁵, respectively, which are fitted from the ground truth dataset. Notably, we divide P-PEVs into several groups to fit different parameters $\rho, \gamma, \alpha, \beta$ according to the various travelling patterns of each group.

In our first manuscript, Supplementary Fig. 1(b), page 2:



Supplementary Fig 1: Grouped travel patterns of P-PEVs. **a**, group behaviour of P-PEVs in spatial exploration preference, following $P_{new} = \rho S^{-\gamma}$ with different values of ρ and γ . **b**, Distributions of P-PEVs trip distance with grouped gamma distribution parameters.

According to your comment, we modify the descriptions in the revised manuscript, Methods, page 15, lines 398-405:

If the decision is to explore, probabilities of visiting different locations should be specified. We denote the current location as SR^{ori} and the next potential location as SR^{des} . We couple the gravity model⁶⁶ and the gamma distribution to describe the probability of visiting SR^{des} , following

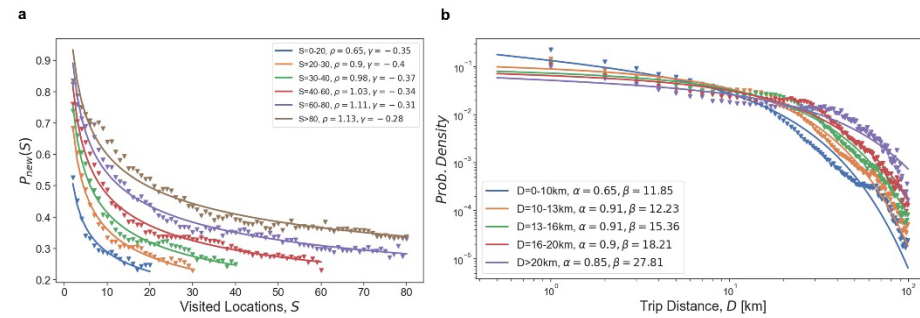
$$P(SR^{des}|SR^{ori}, d) = \frac{\rho(SR^{des})}{\text{gamma}(d)}, \quad (2)$$

where $\rho(SR^{des})$ refers to the visiting popularity index of the stay region SR^{des} , $\text{gamma}(d) = \frac{\beta^\alpha}{\Gamma(\alpha)} d^{\alpha-1} e^{-\beta d}$ is a gamma distribution, and $d = |SR^{des} - SR^{ori}|$ denotes the distance between SR^{des} and SR^{ori} . α, β are the shape parameter and the rate parameter[80] fitted from the trajectories datasets.

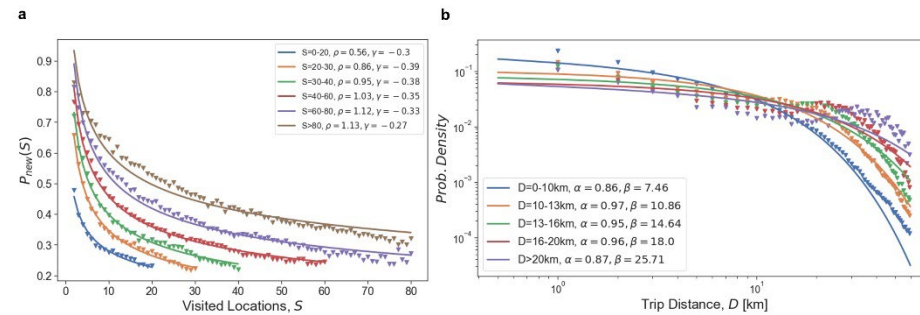
We separate PEVs into several S groups, and different groups possess different spatial exploration and preference parameters $\alpha, \beta, \rho, \gamma$, as shown in Supplementary Fig. 6, 7.

In Supplementary material, page 6, Supplementary Note. 5, Supplementary Fig. 6,7:

Supplementary Note. 5: Grouped spatial parameters of P-PEVs. 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 $gamma(d) = \frac{\beta^\alpha}{\Gamma(\alpha)} d^{\alpha-1} e^{-\beta d}$, as shown in Supplementary Fig. 6 and Supplementary Fig. 7.



Supplementary Fig. 6: Grouped spatial parameters $\rho, \gamma, \alpha, \beta$ of Beijing P-PEV users.



Supplementary Fig. 7: Grouped spatial parameters $\rho, \gamma, \alpha, \beta$ of Shenzhen P-PEV users.

Comment 4

Optimization Formulation: Why is the battery capacity not constrained? Additionally, in function f2, which is a sum of differences, please explain how this prevents simultaneous charging and discharging.

Response to Comment 4

Thank you very much for your question about the V2G optimization. Our explanations are as follows:

(a) Battery capacity constraints

The battery capacity is actually constrained in our formulation. Specifically, in our first manuscript, Methods, page 16, Eq. (10-11) and Eq. (16-17):

$$\begin{aligned} E_{i,t}/E_i^{max} &\geq 0, \\ E_{i,t}/E_i^{max} &\leq 1, \end{aligned}$$

where E_i^{max} denotes the maximum capacity of P-PEV i . In our implementation, we randomly allocate EV models to each P-PEV according to the market share information provided in Supplementary Table 1, Supplementary Information page 2:

Supplementary Table 1: 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

After the EV model is allocated, the battery capacity E_i^{max} is correspondingly constrained to specific values, *e.g.*, 78 kWh for Tesla Model 3.

(b) Simultaneous charging and discharging (SCD)

The utilization of function f_2 is empirical based on our observation. To enhance the credibility of our method, we modify f_2 to a sum of additions $\frac{1}{T} \sum_{i=1}^N \sum_{t=1}^T (P_{i,t}^{ch} + P_{i,t}^{dis})$ in our revised manuscript, the principles of which are discussed or practically used in multiple previous works^{6,7,8}. However, SCD can be avoided in both cases should the weight λ be properly given. To illustrate this, we take 3,000 P-PEVs to optimize under V2G When Staying with $f_1 + \frac{\lambda_{diff}}{T} \sum_{i=1}^N \sum_{t=1}^T (P_{i,t}^{ch} - P_{i,t}^{dis})$ and $f_1 + \frac{\lambda_{add}}{T} \sum_{i=1}^N \sum_{t=1}^T (P_{i,t}^{ch} + P_{i,t}^{dis})$, respectively, while keeping the constraints identical. We set $\lambda_{diff} = 2$ and $\lambda_{add} = 0.015$.

According to your comment, we modify our optimization formulation in the revised manuscript, Methods, page 17, Eq. (15):

$$\begin{aligned} \min \quad & f_1 + \lambda \cdot f_2 \\ & = \frac{1}{T} \sum_{t=1}^T Dev_t + \lambda \cdot \frac{1}{T} \sum_{i=1}^N \sum_{t=1}^T (P_{i,t}^{ch} + P_{i,t}^{dis}) \end{aligned} \quad (15)$$

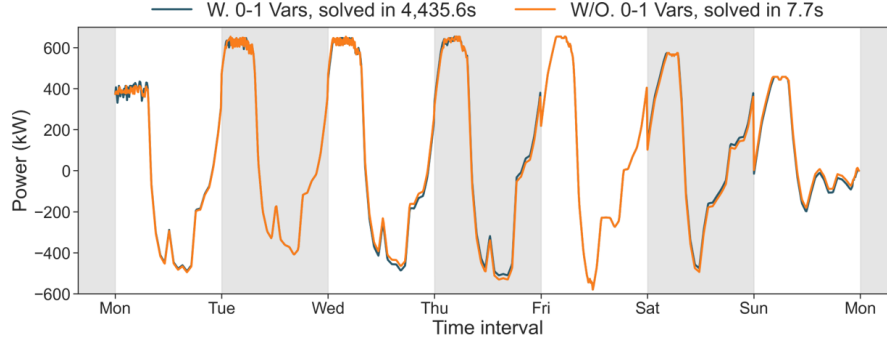
We modify the descriptions in the revised manuscript, Methods, page 18, line 500:

The second weighted term f_2 in Eq. (15) is used to avoid simultaneous charging and discharging (SCD) ⁵³⁻⁵⁵.

In the revised manuscript, Methods, page 18, lines 504-508:

We provide numerical results in Supplementary Fig. 14 to show the significant difference in computation time between our solution and the one with 0-1 variables, and we illustrate different designs of f_2 and the respective weights to avoid SCD in Supplementary Fig. 15 and Supplementary Note. 11. We solve the optimization using the Gurobi solver, setting the value of λ to 0.15.

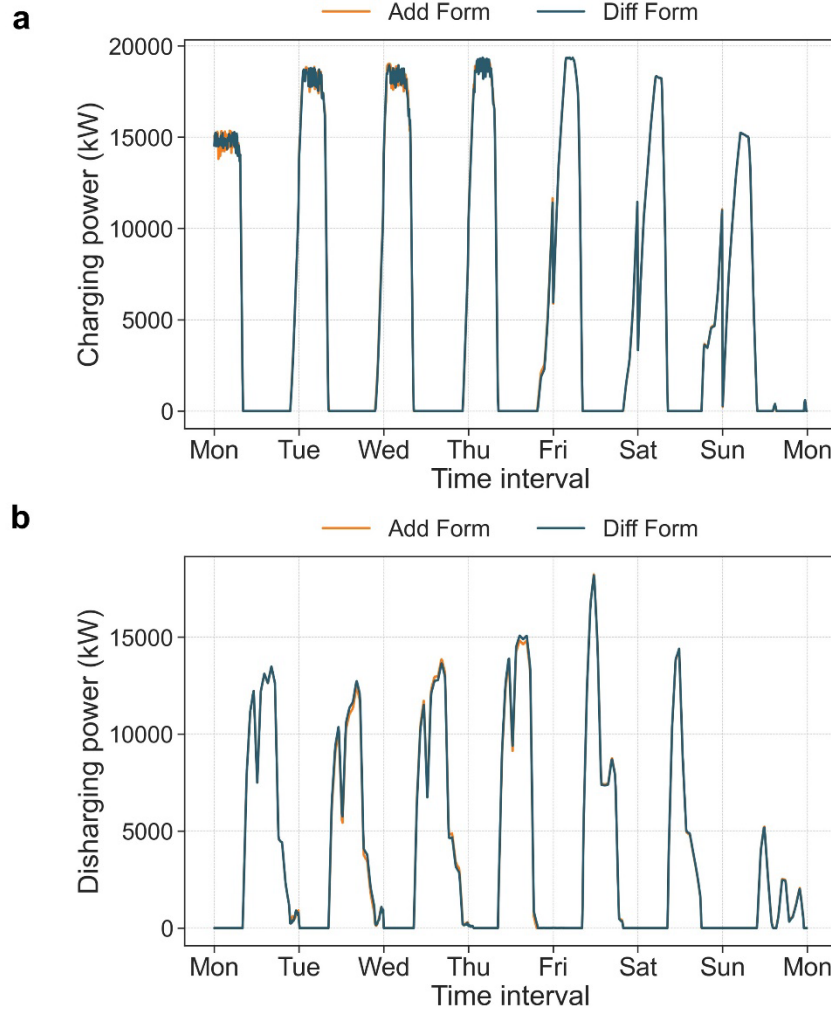
In Supplementary material, page 14, Supplementary Fig. 14, Supplementary Note. 11:



Supplementary Fig. 14: Comparison of results and computational costs between the MILP optimization and LP optimization. As introduced in Method 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.

Supplementary Note 11: Comparison of different forms of λf_2 to eliminate simultaneous charging and discharging (SCD). The second objective term f_2 presented in the manuscript, Eq. (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 manuscript[4-6], and we set $\lambda = 0.15$. Notably, another form of difference $\lambda' f_2' = \lambda' \frac{1}{T} \sum_{i=1}^N \sum_{t=1}^T (P_{i,t}^{ch} - P_{i,t}^{dis})$ can also empirically yield almost identical results. To illustrate this, we consider a batch of 3,000 P-PEVs, and optimize with $f_1 + \lambda f_2$ and $f_1 + \lambda' f_2'$, respectively. We set $\lambda' = 2$ in the experiments. The results are provided in Supplementary Fig. 15. Both $f_1 + \lambda f_2$ and $f_1 + \lambda' f_2'$ produce 0 records of SCD.

In Supplementary material, page 15, Supplementary Fig. 15:



Supplementary Fig. 15: Comparison of different forms of f_2 to eliminate SCD. ‘Add Form’ refers to $\lambda f_2 = \lambda \frac{1}{T} \sum_{i=1}^N \sum_{t=1}^T (P_{i,t}^{ch} + P_{i,t}^{dis})$ and ‘Diff Form’ refers to $\lambda' f_2' =$

$\lambda' \frac{1}{T} \sum_{i=1}^N \sum_{t=1}^T (P_{i,t}^{ch} - P_{i,t}^{dis})$. **a**, Aggregated charging power of 3,000 P-PEVs. **b**, Aggregated discharging power of 3,000 P-PEVs.

Comment 5

Figures & Data Presentation: Provide numerical error values for Figures 3a, 3b, 3c, and 3d. The computation method for Figure 4c is unclear. How is the charging potential considered? Does the charger not inject full power? Please provide a detailed explanation. The scenarios are not clearly defined. Please introduce them explicitly before presenting the results.

Response to Comment 5

Thank you very much for your suggestions about data presentation, the concept of charging potential, and the exposition of V2G scenarios. Our explanations are as follows:

(a) Numerical error values of Figure 3a, 3b, 3c, 3d

Numerical values of Figure 3a, 3b, 3c, 3d are listed in Table.R2, Table.R3, Table.R4, Table.R5.

Notably, all raw data presented in the figures of the manuscript will be provided as Source Data.

Table.R2: Numerical values of single stay duration.

Single stay duration (h)	Density of 2022 SPTD	Density of MOVC
[0, 1]	0.17540	0.19562
[1, 2]	0.17890	0.12730
[2, 3]	0.09423	0.08705
[3, 4]	0.05852	0.06623
[4, 5]	0.04005	0.04928
[5, 6]	0.03217	0.03434
[6, 7]	0.03135	0.03170
[7, 8]	0.03761	0.04272
[8, 9]	0.05084	0.06124
[9, 10]	0.05815	0.07786
[10, 11]	0.04946	0.06067
[11, 12]	0.04496	0.03539
[12, 13]	0.04020	0.02983
[13, 14]	0.02968	0.02472
[14, 15]	0.01922	0.01802
[15, 16]	0.01375	0.01259

[16, 17]	0.01082	0.00977
[17, 18]	0.00920	0.00825
[18, 19]	0.00811	0.00757
[19, 20]	0.00707	0.00692
[20, 21]	0.00583	0.00650
[21, 22]	0.00448	0.00641

1296 Table.R3: Numerical values of single trip distance.

Single trip distance (km)	Density of 2022 SPTD	Density of MOVC
[0, 2]	0.22242	0.11368
[2, 4]	0.16680	0.14851
[4, 6]	0.10514	0.12445
[6, 8]	0.07511	0.09992
[8, 10]	0.05940	0.08152
[10, 12]	0.05031	0.06936
[12, 14]	0.04506	0.06064
[14, 16]	0.03955	0.05160
[16, 18]	0.03457	0.04307
[18, 20]	0.02993	0.03565
[20, 22]	0.02614	0.02995
[22, 24]	0.02293	0.02559
[24, 26]	0.02051	0.02206
[26, 28]	0.01829	0.01875
[28, 30]	0.01586	0.01580
[32, 34]	0.01365	0.01323
[34, 36]	0.01153	0.01069
[36, 38]	0.00974	0.00873
[38, 40]	0.00814	0.00710
[40, 42]	0.00684	0.00572
[42, 44]	0.00577	0.00460
[44, 46]	0.00482	0.00376
[46, 48]	0.00407	0.00310
[48, 50]	0.00341	0.00253

1297 Table.R4: Numerical values of daily visited locations.

Daily visited locations	Density of 2022 SPTD	Density of MOVC
1	0.605590	0.520840
2	0.272701	0.294127
3	0.091062	0.116687
4	0.024175	0.042644
5	0.005298	0.015884
6	0.000992	0.006180
7	0.000151	0.002373
8	0.000025	0.0000853
9	0.000005	0.000302
10	0.000001	0.000109

1298 Table.R5: Numerical values of L-th visited locations.

L-th most visited location	Density of 2022 SPTD	Density of MOVC
1	0.36806	0.38494
2	0.24661	0.19536
3	0.08086	0.08343
4	0.04291	0.04916
5	0.02836	0.03456
6	0.02120	0.02650
7	0.01695	0.02153
8	0.01411	0.01805
9	0.01207	0.01549
10	0.01054	0.01357
11	0.00936	0.01205
12	0.00842	0.01083
13	0.00764	0.00980
14	0.00699	0.00893
15	0.00643	0.00818
16	0.00596	0.00755
17	0.00555	0.00699
18	0.00520	0.00651
19	0.00489	0.00607
20	0.00460	0.00570

We add descriptions in the revised manuscript, Discussion, page 13, lines 320-322:

However, some drawbacks exist in the proposed framework. First, the mobility model is not accurate enough in predicting 0-2km short distances, as shown in Fig. 3b.

(b) Computation of charging potential

We add descriptions of charging potential calculation in the revised manuscript, Methods, page 18, lines 513-519, which are identical to the procedure used in our original manuscript:

Computation of charging potential and discharging potential. Charging potential $CP_{i,s}$ is the maximum capacity of a P-PEV i to absorb electricity in a Stay s . $CP_{i,s}$ is computed following

$$CP_{i,s} = \min(100, SOC_{i,s}^e + P_i * T_{i,s}), \quad (38)$$

where $SOC_{i,s}^e$ denotes the arrival SOC of Stay s , P_i denotes the slow charging power of P-PEV i , and $T_{i,s}$ refers to the duration of Stay s . Similarly, discharging potential $DP_{i,s}$ is defined as follows:

$$DP_{i,s} = \max(SOC_{i,s}^{need}, SOC_{i,s}^e - P_i * T_{i,s}), \quad (39)$$

where $SOC_{i,s}^{need}$ is the SOC format of E^{need} as introduced above. Capacity limit and SOC demand are set as 100 and $SOC_{i,s}^{need}$, respectively. Since we only consider Stay vehicles in time slot t , the values of both limits will vary with time.

Therefore, the values of charging/discharging potential are mainly constrained by stay duration. In other words, if $T_{i,s}$ is not long enough, $CP_{i,s}$ and $DP_{i,s}$ will not reach their thresholds 100 and $SOC_{i,s}^{need}$, respectively, which are also the definitions of Capacity limit and SOC demand, respectively. Thus, there will be a gap either between charging potential and capacity limit or between discharging potential and SOC demand. It does not mean the charger do not inject full power.

In our revised manuscript, we add a yellow dashed line in Fig. 4d in Results, page 6, to illustrate a possible current capacity after V2G. Notably, since we modify the charging behavior model (as explained in your Comment 2), the patterns of charging potential and discharging potential will differ from the results given in our first manuscript.

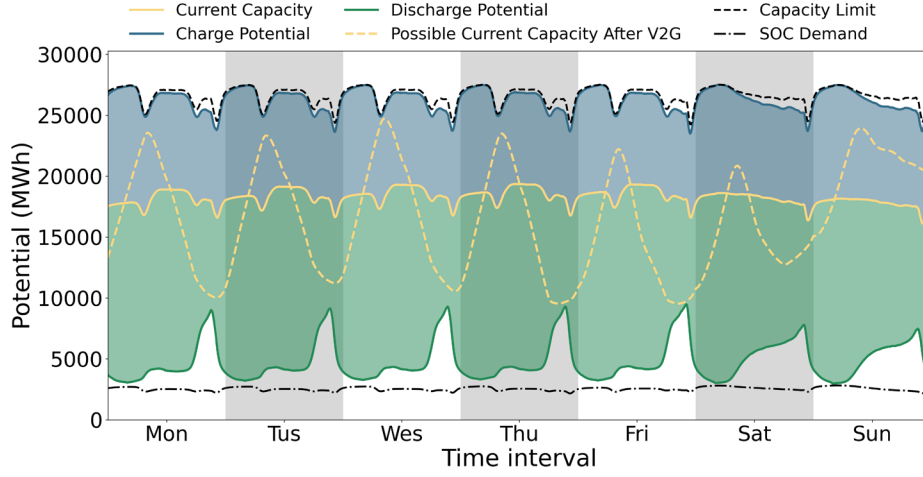


Fig. 4: Shenzhen P-PEV charging profiles and energy storage potential. d, the weekly curve of the current capacity of all PEVs, and the maximum (Charge Potential, blue area) or minimum (Discharge Potential, green area) energy that can be achieved by all PEVs charging or discharging as much as possible. The two levels are constrained by the total battery capacity (Capacity Limit) and the SOC required for travelling (SOC Demand). The yellow dotted line represents a possible variation of the current capacity curve after V2G scheduling.

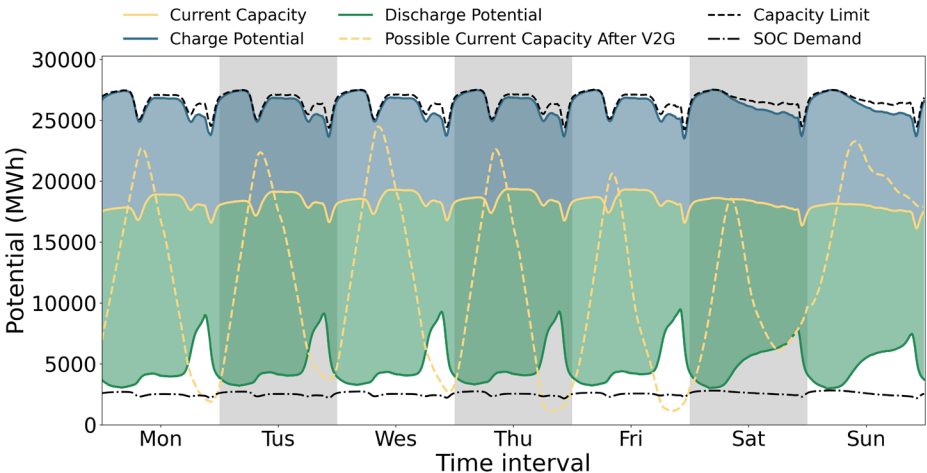
The main purpose of providing Fig. 4d is to qualitatively illustrate that a large portion of P-PEV energy storage resources is underexploited. We emphasize the qualitative characteristic of Fig. 4d in the revised manuscript, Results, page 7, lines 175-176:

Note that Fig. 4d only gives a schematic diagram, which may differ from actual constraints considered in the V2G optimization. The reasons are given in Supplementary Note. 8 and Supplementary Fig. 10,11.

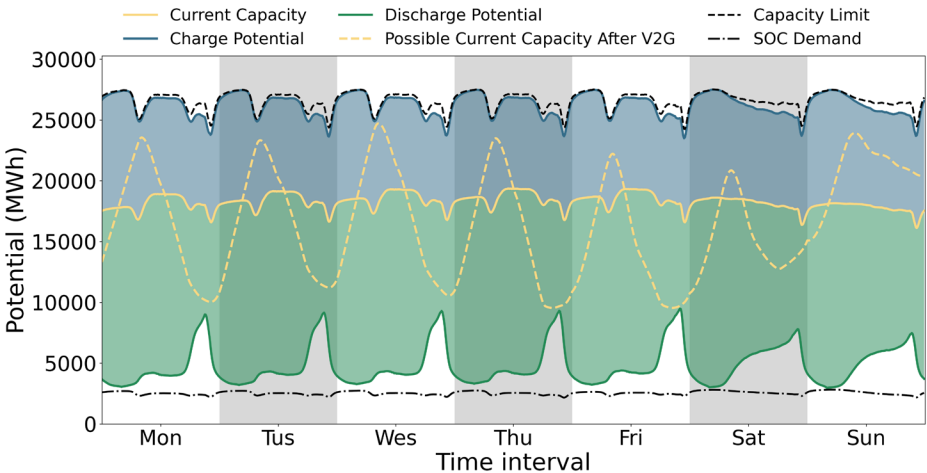
We add explanations in Supplementary material, page 10, Supplementary Note. 8, Supplementary Fig. 10,11:

Supplementary Note. 8: Explanations of charging potential. We illustrate an actual variation of current capacity in Supplementary Fig. 10, which is given by the V2G optimization under V2G When Staying. The yellow dotted line may exceed the range of discharge potential and SOC demand. The reasons are twofold. First, SOC demand $DP_{i,s}$ defined in Eq. (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 Eq. (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 (Eq. (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. 4c (Supplementary Fig. 11). We emphasize that these results serve solely as a schematic illustration and may not precisely reflect the actual optimization results.



Supplementary Fig. 10: An actual result of possible current capacity after V2G.



Supplementary Fig. 11: A modified result of possible current capacity after V2G, as shown in the manuscript, Fig. 4c.

(c) Introduction of V2G scenarios

Based on our explanations for Reviewer 2, Comment 4 (you are kindly referred to Figure.R2, Figure.R3 in Response, Reviewer 2 Comment 4, page 37 for detailed explanations of V2G scenarios), the concepts of different V2G scenarios are as follows:

Charge Only: This scenario is taken as a base case where P-PEV users charge their vehicles upon their arrival, reflecting a general real-world situation when V2G planning is not involved. In Charge Only, users only plug the chargers into the slow charging or QC sockets of their vehicles when they intend to do so. Additionally, a plugged-in charger only operates for unidirectional power flow, *i.e.*, energy flows from the power grid to the vehicle.

V2G When Charging: This scenario is designed to represent mild V2G activities. In V2G When Charging, users only plug the chargers into the slow charging or QC sockets of their vehicles when they intend to do so. Differed from Charge Only, a plugged-in charger under this scenario will operate for bidirectional power flow, *i.e.*, energy may flow either from the power grid to the vehicle, or the opposite.

V2G When Staying: This scenario is designed to represent full V2G activities. In V2G When Staying, users are required to plug the charges into the slow charging or QC sockets of their vehicles in every Stay, even if they do not intend to do so. Identical to V2G When Charging, a plugged-in charger under this scenario will operate for bidirectional power flow.

According to your comment, we add descriptions about V2G scenarios in the revised manuscript, Results, page 4, line 109-128:

For charging modelling, we consider all possible actions of P-PEVs as (1) Move, where the vehicles travel in the city. (2) Park, where users intend to park the vehicles without plugging chargers in them. (3) Charge, where users park the vehicles and plug the slow chargers into their slow charging sockets. (4) QC, where users park the vehicles and plug the QC chargers into their QC sockets. The latter three cases are summarized by the term “Stay” to represent a general situation where P-PEVs are immobile. Energy consumption during Move is calculated by the drivetrain losses ⁶⁷ (see Methods). If a vehicle is in a Stay, MOVE uses a utility-based model ³³ to specify its action of Park, Charge, or QC. The decisions are affected by multiple realistic factors, *e.g.*, the vehicle’s current state of charge (SOC), the user’s preference towards QC, the expected energy that can be obtained in the Stay depending on Stay duration along with available charging modes, and charging costs depending on the current time period and location (public or private). One can refer to Methods for details. Notably, we assume that all the decisions of Park, Charge and QC are made upon arrival of Stay-s. Next, MOVC simulates P-PEVs’ participation in V2G by considering economic subsidies. The compensations required for obeying V2G scheduling are modelled as normal distributions⁴⁰.

Given a specific compensation rate γ , users will decide which V2G scenario to join depending on whether their expected economic gains are met (see Methods, Supplementary Note. 9 and Supplementary Fig. 12). We design three scenarios, namely (1) Charge Only, where V2G is disabled and only unidirectional energy flows from the power grid to the vehicles, (2) V2G When Charging (VWC), where V2G is enabled in Charge and QC actions, and (3) V2G When Staying (VWS), where users are required to plug the chargers in their vehicles to enable V2G once they are in Stay. In other words, V2S compulsorily switches the vehicles' states in Park from isolated to connected (to the grid). Accordingly, users participated in VWS will receive higher compensation compared to those in VWC.

Comment 6

V2G Considerations: The paper assumes that V2G increases energy, but how is the discharged energy accounted for? Are EVs allowed to discharge at any moment? Additionally, how is V2G implemented in the model? Please clarify aspects such as scheduling, user willingness, and battery degradation effects.

Response to Comment 6

Thank you very much for highlighting the issues related to V2G implementation. We acknowledge that the original V2G strategies did not account for critical real-world factors such as user willingness and battery degradation costs, which limited the robustness of our findings. Your insightful comment has prompted us to refine the V2G strategies with more comprehensive considerations. The responses to your comment are as follows:

(a) Whether V2G increases energy, and how the discharged energy is accounted for

We apologize for causing your misunderstanding. Your judgement that V2G increases energy might come from our first manuscript, Results, page 8, Figure 5a:

a

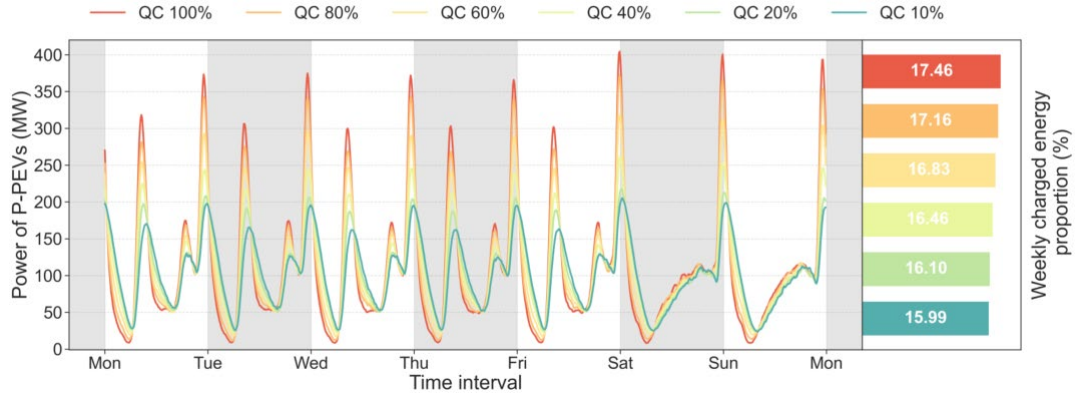


Fig. 5: P-PEV statistics under different QC piles coverages. **a**, Weekly load curves of P-PEVs under different QC pile deployment ratios and the corresponding normalized proportions of weekly charged energy.

The results of Figure 5a are given under Charge Only. As we illustrate in Figure.R3, this scenario does not consider the discharging process and it represents the situation of disordered charging prevalent in the real-world. In Figure 5a, we illustrate the impacts of covering QC piles on the shapes of the charging load curves of P-PEVs. The right subfigure compares the weekly charged energy under different QC pile coverages, the results of which are equivalent to the areas enclosed by the charging curves and the x-axis. The main motivation for providing the results in the right subfigure is to verify whether our simulation of higher QC ratios simply raises the peak load or increases the overall injected energy noticeably. Our results indicate that the former dominates – when the QC ratio rises from 10% to 100%, the total injected energy increases by less than 10%. Considering the inherent limitations of P-PEV battery capacities, we believe the given results are reasonable. However, these results do not imply V2G increases energy, since Charge Only exclusively models unidirectional charging.

The other two strategies V2G When Charging and V2G When Staying also do not simply increase the energy of P-PEVs. Instead, the strategies may either increase or decrease the energy of P-PEVs according to the patterns of the city's electricity load profile. As an example, we consider the case of V2G When Staying under QC 10%, which we illustrated in our first manuscript, Results, page 10, Figure 6a:

a

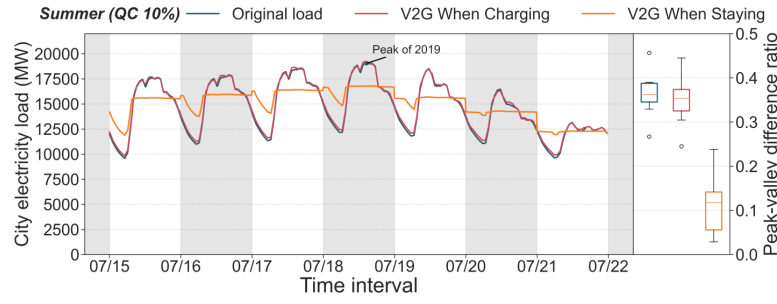


Figure 6: Illustration of the urban load balancing performance of P-PEVs under different scenarios.

a, The weekly curves of Shenzhen's original load, the load under V2G When Charging, and the load under V2G When Staying.

For the V2G When Staying load marked in orange, we illustrate the corresponding patterns of P-PEV total energy variation in Figure.R6:

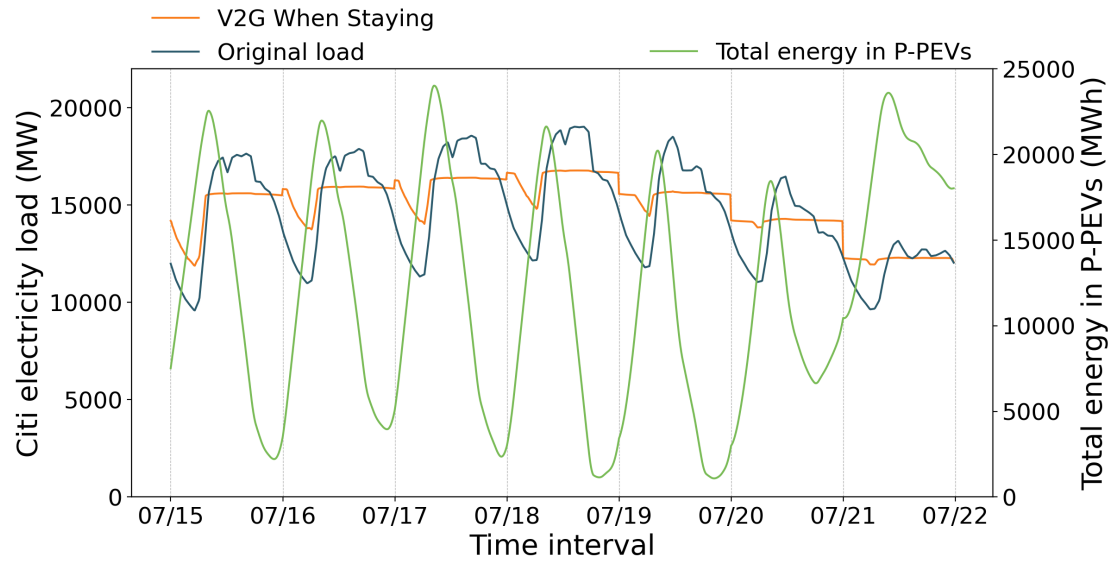


Figure.R6: The weekly curves of Shenzhen's original load, the load under V2G When Staying, and the total energy variation of P-PEVs under V2G When Staying.

As shown in Figure.R6, the total energy of P-PEVs increases during periods of low city electricity demand, indicating a net charging behavior that contributes to valley filling. Conversely, during peak demand periods, the total energy of P-PEVs decreases, reflecting a net discharging behavior that supports peak shaving. Regarding our implementation, both charging and discharging power are constrained in our first manuscript, Methods, page 16, Eq. (12-13) and Eq. (18-19):

$$\begin{aligned} 0 &\leq p_{i,t}^{ch} \leq C_i^{ch}, \\ 0 &\leq p_{i,t}^{dis} \leq C_i^{ch}, \end{aligned}$$

$$\begin{aligned} 0 \leq P_{i,t}^{ch} &\leq C_i^{QC}, \\ 0 \leq P_{i,t}^{dis} &\leq C_i^{QC}, \end{aligned}$$

where C_i^{ch} , C_i^{QC} are constants depending on the EV model of P-PEV i , given in Supplementary Table 1 in the original supplementary material (you are kindly referred to your Comment 2, page 53 for details). The variation of battery energy from time slot t to $t + 1$ is given in our first manuscript, Methods, page 16, Eq. (14, 20):

$$E_{i,t+1} = E_{i,t} + \eta P_{i,t}^{ch}/6 - P_{i,t}^{dis}/6/\eta,$$

where $\eta = 0.9$ denotes the efficiency of charging and discharging. We can extract the values of $E_{i,t}$, $P_{i,t}^{ch}$ and $P_{i,t}^{dis}$ after optimization to look into the details of charged/discharged energy of each P-PEV i .

(b) Whether EVs are allowed to discharge at any moment

EVs are only allowed to discharge at their Stay-s, as we illustrated in Responses, Reviewer 2, Comment 4-5, Figure.R2-Figure.R4. Specifically, in Charge Only, EVs are not allowed to discharge. In V2G When Charging, EVs are allowed to discharge when users intend to plug the chargers into the charging sockets of EVs. In V2G When Staying, EVs are allowed to discharge at every Stay, since users are required to plug the chargers into the charging sockets even if they do not intend to do so.

(c) V2G implementation, and real-world related considerations

Thank you again for pointing out the drawbacks of our V2G strategies. In our first manuscript, the V2G implementation assumes that all P-PEVs will participate in the V2G strategies, and it does not consider real-world aspects such as user willingness and battery degradation, which we admit yields less convincing conclusions of our research. Our modifications are listed as follows:

(c-1) New considerations of user willingness to participate in V2G

We introduce customer damage cost (CDC) to reflect users' concerns about convenience and battery degradation, following the research of Wang et al⁹. CDC can be taken as the extra amount of compensation claimed by EV users for obeying V2G control in addition to the revenue gained from flexible pricing⁹. Assume the average and standard deviation of CDC for P-PEV users are denoted as μ_c and δ_c . For P-PEV i , the CDC is ξ^i following the normal distribution¹⁰ as shown below:

$$\xi^i \text{ obeys } N(\mu_c, \delta_c) \quad (10)$$

In practice, parameters μ_c and δ_c are assumed to be known to the policy makers. They need to determine the unit V2G compensation rate γ , and the P-PEV users fulfilling Eq. (11) potentially consider accepting V2G control⁹:

$$\xi^i \leq \gamma \cdot \mu_c \quad (11)$$

Economic costs f^{cc} to compensate CDC are calculated as follows:

$$f^{cc} = \gamma \cdot \mu_c \sum_{i=1}^N \sum_{t=1}^T P_{i,t}^{dis} \quad (12)$$

Since our research contains two V2G strategies, we set $\mu_c^{VWS} = 0.15 \text{ RMB/kWh}$, $\delta_c^{VWS} = 0.075 \text{ RMB/kWh}$ to compensate users involved in V2G When Staying, and set $\mu_c^{VWC} = 0.03 \text{ RMB/kWh}$, $\delta_c^{VWC} = 0.015 \text{ RMB/kWh}$ to compensate users involved in V2G When Charging. We set four levels of compensation rate γ as $\gamma = 0.5, 1, 1.5, 2$, respectively. The mechanism of accepting a specific V2G strategy is illustrated in Figure.R7:

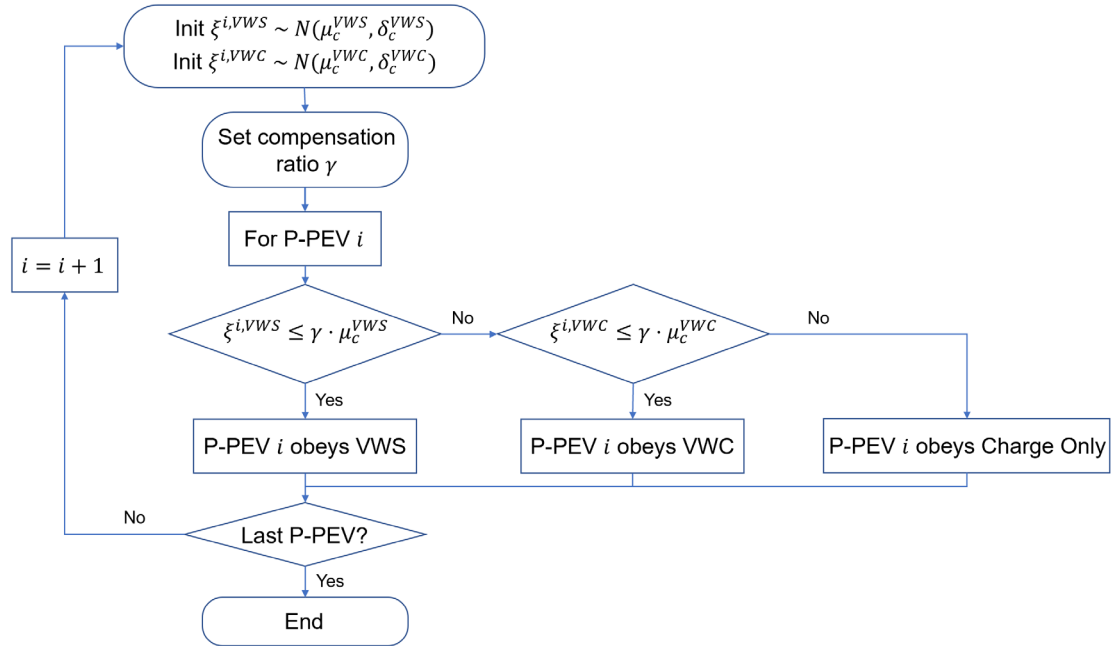


Figure.R7: Flowchart of the V2G participation mechanism.

Under different compensation rates γ , the number of P-PEVs involved in three scenarios are illustrated in Figure.R8:

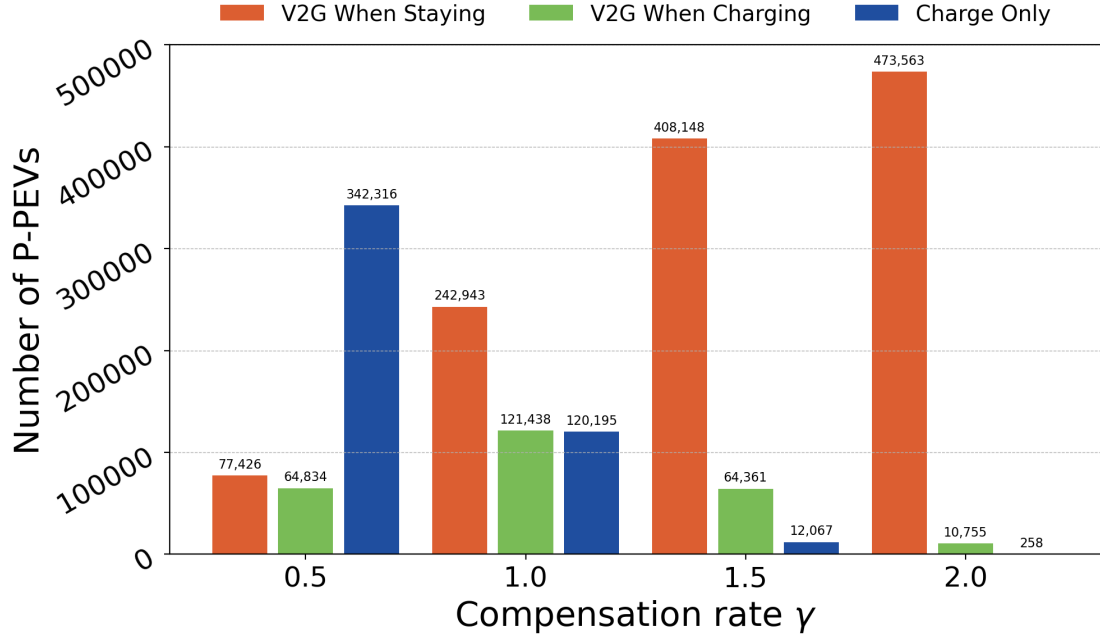


Figure.R8: The number of P-PEVs involved in three scenarios under different compensation rates γ .

We add descriptions in the revised manuscript, Results, page 4, lines 118-123:

Next, MOVC simulates P-PEVs' participation in V2G by considering economic subsidies. The compensations required for obeying V2G scheduling are modelled as normal distributions⁴⁰. Given a specific compensation rate γ , users will decide which V2G scenario to join depending on whether their expected economic gains are met (see Methods, Supplementary Note. 9 and Supplementary Fig. 12).

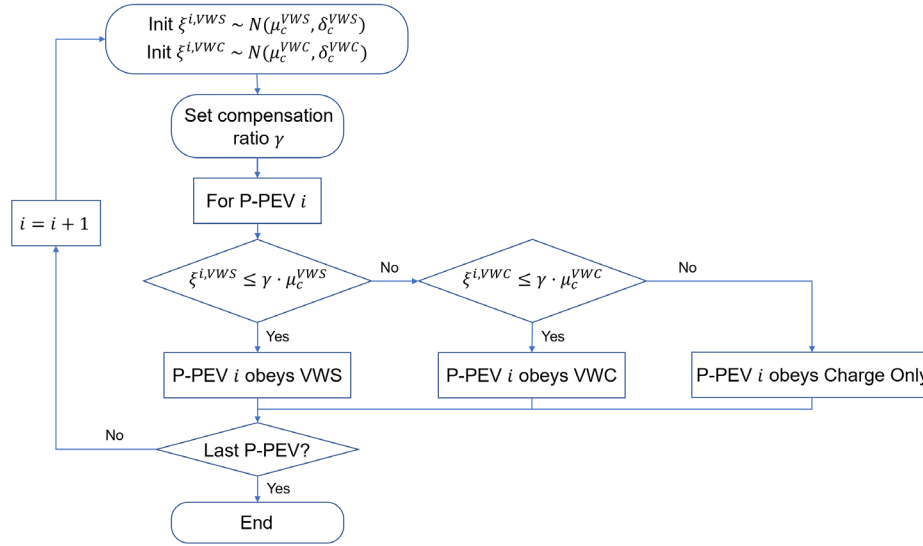
In the revised manuscript, Results, page 9, lines 206-211:

For the former, we set up four compensation rates $\gamma = 0.5, 1.0, 1.5, 2.0$. Considering the discrepancy of participation duration between VWC and VWS, unit discharge compensation is set as 0.015, 0.03, 0.045, 0.06 RMB (around 0.002-0.008 USD)/kWh and 0.075, 0.15, 0.225, 0.3 RMB (around 0.01-0.04 USD)/kWh, respectively. It should be noted that these values represent only the policy-based V2G subsidies and do not include users' revenue from discharging via feed-in tariffs. The specific participation mechanism is given in Methods and Supplementary Fig. 12.

In Supplementary material, page 11, Supplementary Note. 9, Supplementary Fig. 12:

Supplementary Note 9: Details 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, Eq. (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. 12 gives a flowchart of the participation mechanism.



Supplementary Fig. 12: Flowchart of the V2G participation mechanism.

In the revised manuscript, Results, page 10, Fig. 7a:

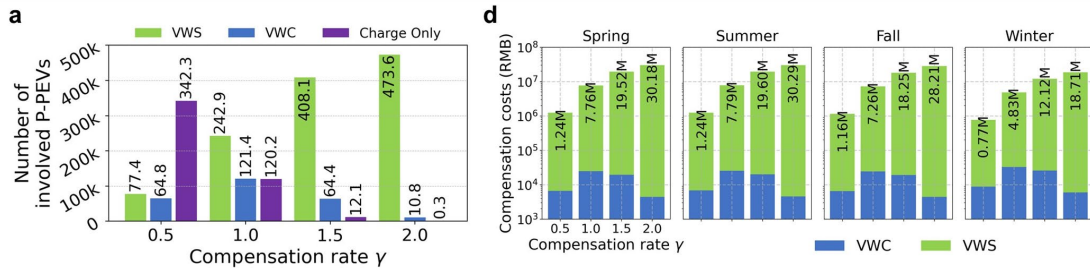


Fig. 7: Evaluation of V2G effects under V2G When Charging (VWC) and V2G When Staying (VWS) with different compensation rates. a, Numbers of P-PEVs participating in different V2G scenarios. d, Total compensation costs paid for users involved in VWC and VWS, respectively.

Corresponding analysis in the revised manuscript, Results, page 11, lines 242-247:

As important indicators of economic compromise, the values of total compensation costs⁴⁰ paid to P-PEV users are illustrated in Fig. 7d. Full investment ($\gamma = 2.0$) requires costs of 18.71M-30.29M RMB (around 2.58M-4.18M USD) per week. The compensation costs exhibit diminishing marginal returns, where approximately 25% of maximum investment yields peak shaving of 770-1196 MW and valley filling of 925-1195 MW, while 66% investment achieves 1172-2008 MW and 1475-1998 MW, respectively, approaching the performance plateau observed at full investment.

(c-2) New considerations of user satisfaction after V2G activities

Similar to the settings of Orfanoudakis et al.¹¹, we evaluate user satisfaction after each V2G activity through Eq. (13):

$$\epsilon^{usr} = \frac{SOC_k}{SOC_k^*} \quad (13)$$

where SOC_k denotes the actual SOC level after a V2G activity in Stay k , and SOC_k^* denotes the desired SOC level of the user upon departure, which can be calculated under Charge Only. The only difference from the literature¹¹ is that we do not use the average values.

We add descriptions in the revised manuscript, Methods, page 18, lines 520-524:

User satisfaction after V2G activities. We evaluate user satisfaction after each V2G activity through Eq. (40)⁶¹:

$$\epsilon_{i,s}^{usr} = \frac{SOC_{k,i,s}}{SOC_{k,i,s}^*} * 100\%, \quad (40)$$

where $SOC_{k,i,s}$ denotes the actual SOC level after a V2G activity in Stay s , and $SOC_{k,i,s}^*$ denotes the desired SOC level of the user i upon departure, which can be calculated following Eq. (3)(11).

$\epsilon_{i,s}^{usr}$ with values larger than 100% are set as 100%.

In the revised manuscript, Results, page 10, Fig. 7g:

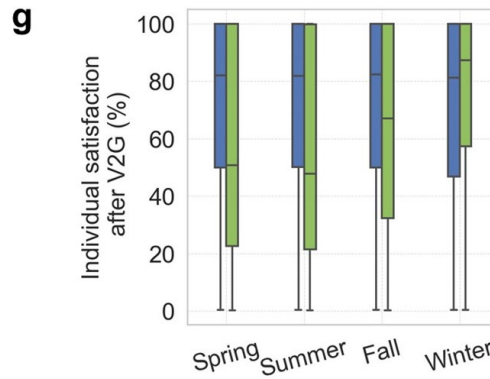


Fig. 7: Evaluation of V2G effects under V2G When Charging (VWC) and V2G When Staying (VWS) with different compensation rates. g, Boxplots of individual satisfaction after a V2G activity.

In the revised manuscript, Results, page 11, lines 240-242:

Regarding users' satisfaction after V2G activities⁶¹ shown in Fig. 7g, the median values of VWC generally exceed 80%, while the metrics of VWS are inferior, reaching 50.73%, 47.81%, 67.03% and 87.30%, respectively.

More experiments related to user satisfaction are conducted, as introduced in the following (c-3).

(c-3) New considerations of V2G scheduling

The added experiments of V2G scheduling are associated with (c-2). Specifically, we evaluate user satisfaction level with no extra constraints. The energy upon departure is constrained in the same way as our first manuscript, Methods, page 16, Eq.(15, 21):

$$E_{i,t_i^{dep}} \geq E_{i,t_i^{dep}}^{need},$$

where $E_{i,t_i^{dep}}^{need}$ is calculated through our first manuscript, Methods, page 16, lines 433-438:

Eq.(15,21) claim that the P-PEV's energy at the departure time must not be less than the user's required energy E_{need} , which is a necessary condition to ensure that the user's travel plans are not affected. We take the minimum value from two candidates as the value of E_{need} . The former is the total energy consumption due to mobility over the next 24 hours, and the latter is the same consumption until the next charging session.

We observe that the user satisfaction level is relatively high when optimizing winter loads. For the summer case where user satisfaction is not good, we modify the optimization formulation by adding the following constraint:

$$E_{i,t_i^{dep}} \geq E_{i,t_i^{dep}}^* \cdot \epsilon^{th}, \quad (14)$$

where $E_{i,t_i^{dep}}^*$ is the desired energy level of the user i upon departure time t_i^{dep} , and ϵ^{th} denotes a minimum satisfaction threshold that must be reached at t_i^{dep} .

We add descriptions in the revised manuscript, Methods, page 19, lines 525-529:

Eq. (40) is regarded as an evaluation metric after V2G optimization. For scenarios with inferior satisfaction (e.g., spring, summer, fall), we illustrate optimization results with a stricter constraint to the SOC level upon departure in Fig. 8. The constraints are added as follows:

$$E_{i,t_i^{dep}} \geq E_{i,t_i^{dep}}^* \cdot \epsilon^{usr,th}, \quad (41)$$

where $E_{i,t_i^{dep}}^*$ is the desired energy level of user i upon departure time t_i^{dep} , and $\epsilon^{usr,th}$ denotes a minimum satisfaction threshold that must be reached at t_i^{dep} . We set $\epsilon^{usr,th}$ as 40%, 60% and 80% in our experiments.

In the revised manuscript, Results, page 11, Fig. 8a,b,d,e:

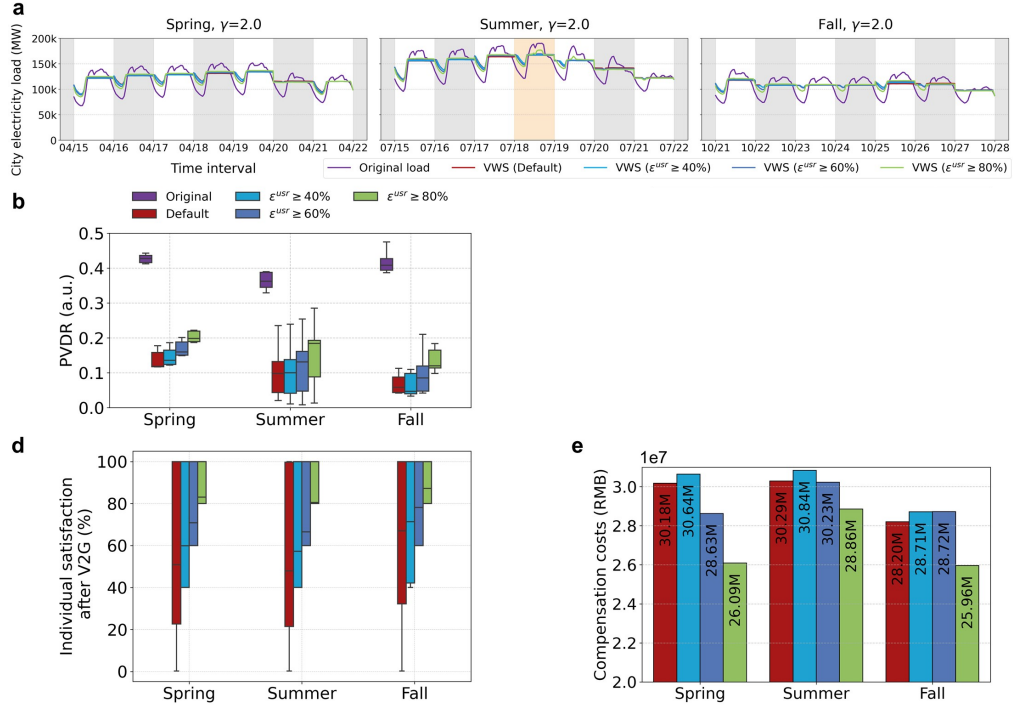


Fig. 8: Evaluation of V2G effects under VWS, $\gamma = 2.0$ considering higher minimum satisfaction after V2G activities. a, The weekly curves of Shenzhen's original load and the load under VWS, $\gamma = 2.0$, with minimum satisfaction $\epsilon^{usr} = 40\%, 60\%, 80\%$, respectively. The winter load is not considered since the median of the default VWS satisfaction in winter already exceeds 80% (see Fig. 7g). b, PVDR with different minimum values of ϵ^{usr} in three seasons. c, Distributions of individual battery degradation costs with different minimum values of ϵ^{usr} in three seasons. d, Boxplots of user satisfaction after V2G activities. The whiskers of some boxplots merge with their quartile markers (Q1/Q3) due to particular data distribution features. e, Total compensation costs paid with different minimum values of ϵ^{usr} .

In the revised manuscript, Results, page 11, lines 248-257:

V2G scheduling considering higher user satisfaction. As discussed above, while VWS demonstrates superior load balancing performance, this comes at non-trivial economic costs. Furthermore, user satisfaction with V2G dispatching remains low across all seasons except winter. This section investigates how V2G evaluation metrics evolve when enforcing minimum user satisfaction thresholds. Specifically, we denote user satisfaction after a V2G activity as ϵ^{usr} . We consider VWS, $\gamma = 2.0$ scenario with three thresholds $\epsilon^{usr} \geq 40\%, \epsilon^{usr} \geq 60\%, \epsilon^{usr} \geq 80\%$, respectively. Experiments are conducted for summer, spring and fall loads.

As shown in Fig. 8a, stricter satisfaction requirements differentially impact valley-filling performance across seasons. Notably, peak-shaving capability remains largely unaffected, except for two high-demand summer days when ϵ^{usr} is constrained to exceed 80%. In Fig. 8b, as ϵ^{usr} thresholds increase, the median PVDR values of three seasons are gradually lifted from 0.118, 0.098, 0.058 to 0.199, 0.184, 0.120, respectively.

In the revised manuscript, Results, page 12, lines 261-265:

In Fig. 8d, user satisfaction exhibits a significant improvement, with median values progressively increasing from 50.73%, 47.81%, 67.03% to 83.07%, 80.57%, 87.27%, respectively. In Fig. 8e, the total compensation costs for $\epsilon^{usr} \geq 40\%$ and $\epsilon^{usr} \geq 60\%$ exhibit both increases and decreases compared to the default scenario, whereas the total costs for $\epsilon^{usr} \geq 80\%$ consistently decrease relative to the baseline. Specifically, the reductions for $\epsilon^{usr} \geq 80\%$ reach 4.09M, 1.43M, and 2.24M RMB (around 0.56M-0.31M USD), respectively.

(c-4) New considerations of battery degradation costs

General quantification of battery degradation requires information of operating temperature, operating current, internal resistance, and calendar aging effects^{12,13,14,15}. Since these features can hardly be obtained when considering large-scale P-PEVs, we refer to the works of Le et al.¹⁶, Antoniadou-Plytaria et al.¹⁷, and Nunna et al.¹⁸ to leverage a feasible degradation model that focuses on cycling aging costs. Specifically, we denote the depth of discharge (DoD) of P-PEV i as $D_{i,t} = 1 - \frac{E_{i,t}}{E_i^{max}}$, where E_i^{max} is the maximum battery capacity of P-PEV i according to Supplementary Table 1 in the first manuscript. Then the DoD-number of cycles (NC) curve can be fitted as Eq. (15) through measurement¹⁶:

$$\phi(D_{i,t}) = k_1 D_{i,t}^{k_2} + k_3, \quad (15)$$

where k_1, k_2, k_3 are coefficients describing the relationship between the NC and DoD. We set $k_1 = 1.4 * 10^5, k_2 = -5.01 * 10^{-1}, k_3 = -1.23 * 10^5$ following Le et al.¹⁶. The degradation cost function of P-PEV i in time slot t with respect to DoD can be expressed as follows^{16,17,18}:

$$C_{i,t}^D = \left| \frac{1}{\phi(D_{i,t})} - \frac{1}{\phi(D_{i,t-1})} \right| B_i^{cost}, \quad (16)$$

where $\phi(D_{i,t})$ and $\phi(D_{i,t-1})$ are the NC at DoD = $D_{i,t}$ and $D_{i,t-1}$, respectively. The battery investment cost B_i^{cost} can be computed as follows:

$$B_i^{cost} = b_i^{cost} \cdot E_i^{max}, \quad (17)$$

where b_i^{cost} is the coefficient set as 957 RMB/kWh.

We add descriptions of battery degradation costs in the revised manuscript, Methods, page 19, lines 530-541:

Battery degradation modelling. General quantification of battery degradation requires information on operating temperature, operating current, internal resistance, and calendar aging effects, which can hardly be obtained for large-scale P-PEVs. Therefore, we focus on cycling aging costs based on previous works^{50,71,72}. The depth of discharge (DoD) of P-PEV i is denoted as $D_{i,t} = 1 - \frac{E_{i,t}}{E_i^{max}}$, where $E_{i,t}$ can be determined following either Eq. (3)(11) or V2G formulations in Eq. (15)(33), and E_i^{max} is the maximum battery capacity of P-PEV i . Then the DoD-number of cycles (NC) curve can be fitted via measurement⁵⁰:

$$\phi(D_{i,t}) = k_1 D_{i,t}^{k_2} + k_3, \quad (42)$$

where k_1, k_2, k_3 are coefficients describing the relationship between NC and DoD. We set $k_1 = 1.4 \times 10^5, k_2 = -5.01 \times 10^{-1}, k_3 = -1.23 \times 10^5$, respectively⁵⁰. The degradation costs of P-PEV i in time slot t with respect to DoD can be calculated as follows^{50,71,72}:

$$C_{i,t}^D = \left| \frac{1}{\phi(D_{i,t})} - \frac{1}{\phi(D_{i,t-1})} \right| B_i^{cost}, \quad (43)$$

where $\phi(D_{i,t})$ and $\phi(D_{i,t-1})$ are the NC at DoD= $D_{i,t}$ and $D_{i,t-1}$, respectively. The battery investment cost B_i^{cost} is given as:

$$B_i^{cost} = b_i^{cost} \cdot E_i^{max}, \quad (44)$$

where b_i^{cost} is the coefficient set as 957 RMB (132 USD)/kWh⁵⁰.

New results in the revised manuscript, Page 10, Fig. 7f:

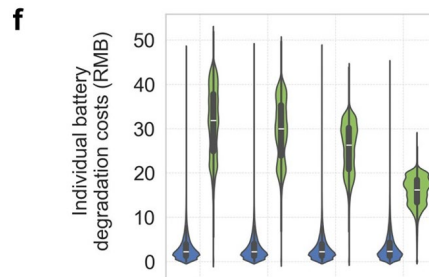


Fig. 7: Evaluation of V2G effects under V2G When Charging (VWC) and V2G When Staying (VWS) with different compensation rates. f, Distributions of individual battery degradation costs under VWC and VWS.

Corresponding analysis in the revised manuscript, Results, page 11, lines 237-240:

Fig. 7f gives distributions of battery degradation costs⁵⁰, where the average and median costs of VWC locate between 2.95-3.22 RMB (around 0.41-0.44 USD)/week and 2.22-2.32 RMB (around 0.31-0.32 USD)/week, and the values of VWS reach 31.15, 29.41, 25.37, 15.72 RMB (around 4.30-2.17 USD)/week and 31.82, 29.95, 26.30, 16.20 RMB (around 4.39-2.23 USD)/week, respectively.

New results in the revised manuscript, Page 11, Fig. 8c:

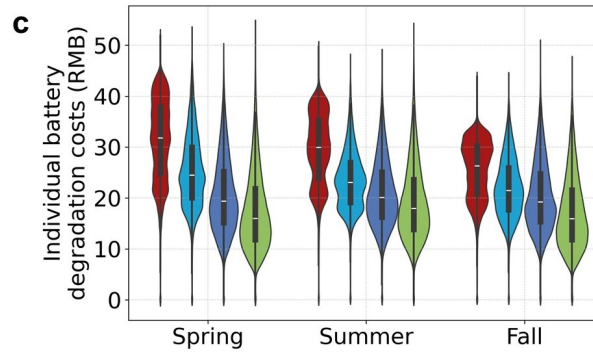


Fig. 8: Evaluation of V2G effects under VWS, $\gamma = 2.0$ considering higher minimum satisfaction after V2G activities. c, Distributions of individual battery degradation costs with different minimum values of ϵ^{usr} in three seasons.

Corresponding analysis in the revised manuscript, Results, pages 11-12, lines 257-259:

In Fig. 8c, the average weekly battery degradation costs across three seasons decline from 31.15, 29.41, 25.37 RMB (around 4.30-3.50 USD) to 17.47, 19.21, 17.20 RMB (around 2.41-2.37 USD), respectively.

In the revised manuscript, Results, page 12, lines 266-268:

To briefly summarize, maintaining high user satisfaction leads to a 30-40% reduction in battery degradation costs and a 5-13% decrease in total compensation costs, though with modest trade-offs in valley filling efficiency and limited peak shaving performance under extreme high-load conditions.

Considering your comment and Reviewer 1, Comment 2, Reviewer 2, Comment 6, we add results in the revised manuscript, Results, page 12, Fig. 9f-Fig. 9i:

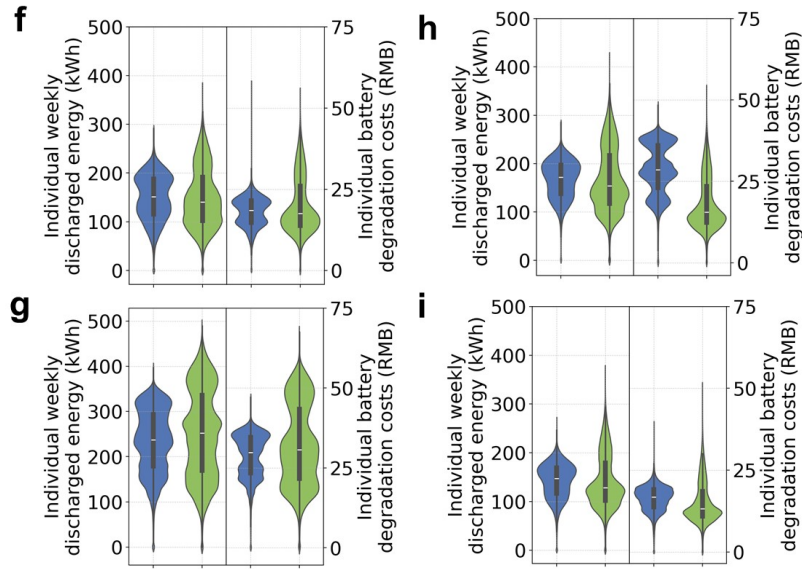


Fig. 9: V2G evaluation in Beijing, Shanghai, Guangzhou and Shenzhen based on MOVC.

f,g,h,i, Distributions of individual weekly discharged energy and battery degradation costs under the two V2G scenarios.

Corresponding analysis in the revised manuscript, Results, page 13, lines 284-306:

In Fig. 9f-Fig. 9i, after full QC upgrades, the median values of weekly discharged energy in Beijing, Guangzhou and Shenzhen decrease from 151.25, 171.66, 147.04 kWh to 140.47, 154.21, 128.10 kWh, respectively. However, in Shanghai where the overall city load is higher, the values increase from 237.30 to 252.20 kWh. Regarding battery degradation costs, the median values of Beijing, Guangzhou and Shenzhen decrease from 18.55, 28.54, 16.78 RMB (around 2.56-2.31 USD) to 17.54, 15.57, 13.19 RMB (around 2.42-1.82 USD), respectively. In contrast, the metric of Shanghai slightly increases from 29.87 to 30.71 RMB (around 4.12 to 4.24 USD). Another noteworthy effect is observed in Shenzhen's case. With the deeper P-PEV penetration, the median weekly discharged energy decreases from 220.29 kWh (Fig. 7e, summer) to 147.04 kWh-128.10 kWh, and the median battery costs also decrease from 29.95 RMB (around 4.13 USD, Fig. 7f, summer) to 16.78-13.19 RMB (around 2.31-1.82 USD). Notably, the inequity of both metrics in all cities grows significantly. For the weekly discharged energy, the standard deviation of all individuals increases from 46.05, 70.17, 39.00, 35.43 kWh to 60.85, 97.37, 65.74, 54.59 kWh, respectively. For the battery degradation costs, the metric goes up from 4.42, 6.82, 7.71, 3.83 RMB (around 0.61-0.53 USD) to 8.61, 12.61, 7.98, 6.30 RMB (around 1.19-0.87 USD).

The results indicate that as the number of V2G-related vehicles increases, cities with mild load profiles (Beijing, Guangzhou, Shenzhen) exhibit lighter optimization burdens to each vehicle, leading to lower discharged energy and battery degradation costs. In such cases, QC upgrades yield marginal additional benefits. Conversely, in cities with higher load levels (*e.g.*, Shanghai), achieving comparable effects requires a higher penetration of P-PEVs. Additionally, expanding QC infrastructure significantly increases the variance of both discharged energy and degradation costs (Fig. 9f-Fig. 9i), enlarging the standard deviation by 32%-69% and 4%-95%, respectively. Customized V2G incentive policies may be required to mitigate such inequity. Given the substantial costs for residential grid upgrades, city policymakers should carefully weigh the trade-offs between investments and expected benefits.

We summarize the contents of (c-1)-(c-4) in the revised manuscript, Discussion, pages 13-14, lines 335-343:

The weekly median PVDR of the original city load can be decreased up to 72.4%-98.7% seasonally, where VWC and VWS-related vehicles release tens of MW and 2,000 MW load balancing capacities, respectively, at compensation costs of 18.71M-30.29M RMB (around 2.58M-4.18M USD). However, lower costs can also attract a sufficient number of users to participate and yield decent load shifting outcomes. Furthermore, if considering higher user satisfaction after V2G activities, individual battery degradation costs and total compensation costs can be reduced by 30-40% and 5-13%, respectively, with an acceptable sacrifice on load balancing performance. A generalization study across China's megacities indicates that the imaginary residential QC upgrades offer marginal benefits in load balancing (0%-12%) and battery lifespan, but may exacerbate individual energy scheduling inequity by 4%-95%.

Your comment is summarized as a participation mechanism and several V2G effects evaluation factors, as introduced in the revised manuscript, Results, page 4, Figure 2:

We modify Figure 2 in the revised manuscript, Results, page 4:

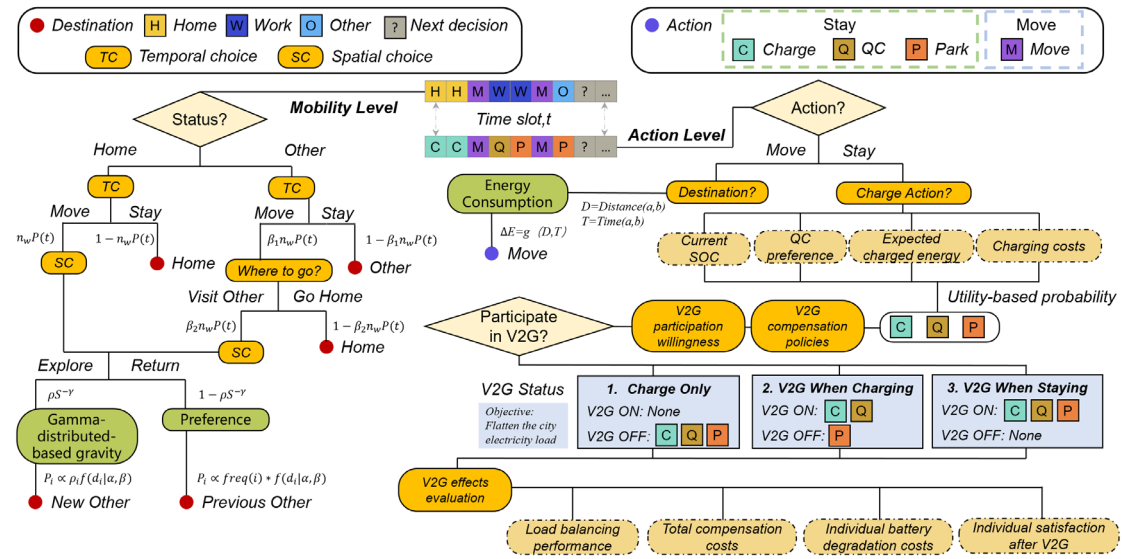


Fig. 2: An overview of the MOVC framework. The MOVC framework assigns mobility and charging information for each time slot. In the Timegeo²²-based mobility model, temporal choices of PEVs are controlled by user-specific parameters, including home-based hours (n_w), dwell rate (β_1), and burst rate (β_2). Spatial choices are made by combining the gravity model⁶⁶ and the gamma-distributed-based EPR mechanisms. In the charging model, charging decisions are simulated via utility functions of multiple factors³³. A compensation mechanism⁴⁰ is utilized to simulate users' participation in three different V2G strategies. The objective is designed to flatten the city's electricity load as much as possible. The effects of V2G are evaluated considering load balancing performance, total compensation costs⁴⁰, battery degradation costs⁵⁰ and users' satisfaction after V2G⁶¹. Energy consumption during movement is estimated based on travel distance and travel time⁶⁷.

Corresponding explanations in the revised manuscript, Results, page 4, lines 119-123:

Next, MOVC simulates P-PEVs' participation in V2G by considering economic subsidies. The compensations required for obeying V2G scheduling are modelled as normal distributions⁴⁰. Given a specific compensation rate γ , users will decide which V2G scenario to join depending on whether their expected economic gains are met (see Methods, Supplementary Note. 9 and Supplementary Fig. 12).

In the revised manuscript, Results, pages 4-5, lines 128-132:

Finally, the V2G effects are evaluated in four aspects, namely (1) load balancing performance, (2) total compensation costs paid to users involved in V2G, (3) individual battery degradation costs,

and (4) individual satisfaction after V2G activities. The above metrics are designed to comprehensively assess V2G outcomes in city power stabilization, economic input, and widely concerned V2G impacts on batteries and range anxiety.

Comment 7

Data Validation: Is there a reference for the dataset used for validation? Please provide details on its source and reliability.

Response to Comment 7

Thank you very much for the comment. The datasets used for validation include 2022 Shenzhen P-PEV Trajectories Dataset (2022 SPTD), 2022 Beijing P-PEV Trajectories Dataset (2022 BPTD), 2019 Shenzhen Parking Census Dataset (2019 SPCD), and 2019 National PEV Big Data Statistics (2019 NPBDS). The details of them are introduced as follows:

(a) 2022 SPTD and 2022 BPTD

The two datasets are divided from a larger P-PEV trajectories dataset, provided by the Shenzhen Municipal Transport Bureau. Due to the related confidentiality agreement, we are not allowed to share the original data directly. Full data is available on request.

(b) 2019 SPCD

The dataset is provided by the Shenzhen Municipal Development and Reform Commission. Due to the related confidentiality agreement, we are not allowed to share the original data directly. Full data is available on request.

(c) 2019 NPBDS

The dataset is from the book *Annual Report on the Big Data of New Energy Vehicle in China*. The book is written by Beijing Yiwei big data research center of new-energy vehicles. The ISBN of the book is 9787111661924, recorded by National Library of China¹⁹. The data used for validation is provided in the Source Data folder as one of our manuscript items.

Comment 8

Manuscript Organization & Writing: Improve the manuscript structure to ensure that figures appear in the order they are mentioned in the text. Proofread the paper to correct typos. For example, on page 5, line 137, you reference Figures 1a and 1b, but these figures do not exist. Additionally, Table 1 does not present EV parameters, despite the statement in lines 389–392 on page 15.

Response to Comment 8

Thank you very much for your suggestions about the manuscript structure and proofreading. Our modifications and explanations are as follows:

(a) Proper locations of figures

We add several new figures in the revised manuscript, and we adjust the locations of all figures to make them appear properly. We use a smaller font size for figure captions in the revised manuscript.

(b) Paper proofreading

We apologize for causing your misunderstanding. In our first manuscript, page 5, line 137, our references are Supplementary Fig. 1a, 1b rather than Fig. 1a, 1b. The figures are provided in the supplementary information file named “Unlocking the potential of private EVs for shaving the electricity peak load of the city: Supplementary Information”, which we believe we have submitted as one of the manuscript items. To our knowledge, in multiple existing works published in Nature Communications^{20,21,22}, supplementary materials are referenced in plain texts (*e.g.*, ‘Supplementary Fig. 3’) rather than as hyperlinks. Similarly, in our first manuscript, page 15, lines 389-392, we refer to Supplementary Table. 1 rather than Table. 1.

References

1. Jiang, S., Yang, Y., Gupta, S., Veneziano, D., Athavale, S., & González, M. C. (2016). The TimeGeo modeling framework for urban mobility without travel surveys. *Proceedings of the National Academy of Sciences*, 113(37), E5370-E5378.
2. Barthélemy, M. (2011). Spatial networks. *Physics reports*, 499(1-3), 1-101.
3. Fu, Z., Liu, X., Zhang, J., Zhang, T., Liu, X., & Jiang, Y. (2025). Orderly solar charging of electric vehicles and its impact on charging behavior: A year-round field experiment. *Applied Energy*, 381, 125211.
4. Liu, Y. S., Tayarani, M., & Gao, H. O. (2022). An activity-based travel and charging behavior model for simulating battery electric vehicle charging demand. *Energy*, 258, 124938.
5. Gamma distribution. (2025, March 7). In Wikipedia.
https://en.wikipedia.org/wiki/Gamma_distribution
6. Yang, P., & Nehorai, A. (2014). Joint optimization of hybrid energy storage and generation capacity with renewable energy. *IEEE Transactions on Smart Grid*, 5(4), 1566-1574.
7. Li, Q., & Vittal, V. (2016). Non-iterative enhanced SDP relaxations for optimal scheduling of distributed energy storage in distribution systems. *IEEE Transactions on Power Systems*, 32(3), 1721-1732.
8. Yang, S., Gao, H. O., & You, F. (2023). 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.
9. Wang, X., Nie, Y., & Cheng, K. W. E. (2019). Distribution system planning considering stochastic EV penetration and V2G behavior. *IEEE transactions on intelligent transportation systems*, 21(1), 149-158.
10. Li, W. (1994). Reliability assessment of electric power systems using Monte Carlo methods. Springer Science & Business Media.
11. Orfanoudakis, S., Diaz-Londono, C., Yılmaz, Y. E., Palensky, P., & Vergara, P. P. (2024). Ev2gym: A flexible v2g simulator for ev smart charging research and benchmarking. *IEEE Transactions on Intelligent Transportation Systems*.
12. Sagaria, S., van der Kam, M., & Boström, T. (2025). Vehicle-to-grid impact on battery degradation and estimation of V2G economic compensation. *Applied Energy*, 377, 124546.
13. Wang, D., Coignard, J., Zeng, T., Zhang, C., & Saxena, S. (2016). Quantifying electric vehicle battery degradation from driving vs. vehicle-to-grid services. *Journal of Power Sources*, 332, 193-203.
14. Gong, J., Wasylowski, D., Figgner, J., Bihn, S., Rücker, F., Ringbeck, F., & Sauer, D. U. (2024). Quantifying the impact of V2X operation on electric vehicle battery degradation: An experimental evaluation. *ETransportation*, 20, 100316.
15. Ebrahimi, M., Rastegar, M., Mohammadi, M., Palomino, A., & Parvania, M. (2020). Stochastic charging optimization of V2G-capable PEVs: A comprehensive model for battery aging and customer service quality. *IEEE Transactions on Transportation Electrification*, 6(3), 1026-1034.
16. Le, L. B. (2023). Building energy management and Electric Vehicle charging considering battery degradation and random vehicles' arrivals and departures. *Journal of Energy Storage*, 64, 107141.

- 1798 17. Antoniadou-Plytaria, K., Steen, D., Carlson, O., & Ghazvini, M. A. F. (2020). Market-based
1799 energy management model of a building microgrid considering battery degradation. IEEE
1800 transactions on smart grid, 12(2), 1794-1804.
- 1801 18. Nunna, H. K., Battula, S., Doolla, S., & Srinivasan, D. (2016). Energy management in smart
1802 distribution systems with vehicle-to-grid integrated microgrids. IEEE transactions on smart
1803 grid, 9(5), 4004-4016.
- 1804 19. National Library of China. (n.d.). OPAC. <http://opac.nlc.cn/>
- 1805 20. Tao, S., Liu, H., Sun, C., Ji, H., Ji, G., Han, Z., ... & Sun, H. (2023). Collaborative and
1806 privacy-preserving retired battery sorting for profitable direct recycling via federated
1807 machine learning. Nature Communications, 14(1), 8032.
- 1808 21. Ma, R., Tao, S., Sun, X., Ren, Y., Sun, C., Ji, G., ... & Zhou, G. (2024). Pathway decisions
1809 for reuse and recycling of retired lithium-ion batteries considering economic and
1810 environmental functions. Nature communications, 15(1), 7641.
- 1811 22. Tao, S., Ma, R., Zhao, Z., Ma, G., Su, L., Chang, H., ... & Zhou, G. (2024). Generative
1812 learning assisted state-of-health estimation for sustainable battery recycling with random
1813 retirement conditions. Nature Communications, 15(1), 10154.
- 1814 23. Longshine Technology Group Co., Ltd. (2024). *Xindiantu - Aggregated charging service*
1815 *platform*. <https://www.longshine.com/solutionInfo/81>

Responses to Reviewer Comments

Title	Unlocking vehicle-to-grid potential of load shifting in China’s megacities considering comprehensive real-world behaviors
Revised Title	
Authors	Kaisan Li, Xinxin Li, Zuxun Xiong, Shengyu Tao, Gucheng Zhao, Yi Jiang, He Qi, and Yi Zhang
Revised Authors	
Journal	Nature Communications
Manuscript ID	NCOMMS-24-79794B

Table of Contents

1	Response to Reviewers.....	3
2	Response to Reviewer #1	3
3	Response to Reviewer #2	5

4 **Response to Reviewers**

5 **Response to Reviewer #1**

6 I appreciate the authors' substantial effort in addressing my previous concerns through extensive
7 revisions, and I recognize the significant improvements made to the manuscript. Specifically, the
8 expanded scope to include multiple Chinese megacities (Beijing, Shanghai, Guangzhou, Shenzhen),
9 the enhanced realism in the charging behavior model, and the clarified methodological novelty in
10 terms of detailed, large-scale, user-centric modeling are commendable and reflect a meticulous and
11 rigorous research process.

12

13 Nevertheless, after careful consideration, my primary concern remains regarding the level of novelty
14 and the transformative insights provided by this manuscript. While the authors clearly demonstrate
15 detailed modeling rigor, systematic analysis, and careful validation, the fundamental conceptual
16 framework—leveraging electric vehicles (EVs) and vehicle-to-grid (V2G) technologies for urban
17 grid load balancing—is well-established and extensively explored in existing literature.
18 Consequently, the incremental insights provided, despite their depth, may not sufficiently meet the
19 high-impact and broad audience criteria typically expected for publication in *Nature*
20 Communications.

21

22 Therefore, I regretfully conclude that the manuscript, though methodologically sound and
23 significantly improved, does not offer the necessary level of general or impactful insights for this
24 particular journal.

25

26 However, the quality and detailed analysis presented would certainly be well-received in specialized
27 or topical journals on energy or transportation. Hence, I strongly encourage the authors to consider
28 submitting this refined manuscript to journals.

29 Dear Respected Reviewer,

30 We sincerely thank you for your previous constructive comments and for your positive
31 assessment of our revised manuscript. We are especially grateful for the recognition of our
32 substantial efforts in addressing previous concerns.

33 We understand your perspective regarding the novelty and transformative insights. Following
34 your comments in the last round, we deeply rewrite the manuscript to distinguish our work from
35 prior theoretical or small-scale V2G studies. Although we agree that the general topic of V2G has
36 been extensively studied, we believe our contributions in combining the geographical scope and the
37 fine-grained perspective are more than incremental in the field of V2G. Since we do not receive
38 specific comments from you and other reviewers on further improvements in this round, it is
39 regretfully difficult for us to move forward.

40 However, we greatly appreciate your time and guidance once again. Your feedback throughout

41 the review process has been invaluable in helping us strengthen the manuscript.

42

43 **Response to Reviewer #2**

44 Thanks for the efforts of the authors to revise this manuscript. All my comments have been well
45 addressed.

46 Dear Respected Reviewer,

47 Thank you for your time and guidance. Your comments help us significantly improve the quality
48 of our manuscript to meet the high standards for publication in Nature Communications.