Research on Charging Behaviour Strategies for Electric Vehicles Based on Machine Learning

Simeng Wang *

Beijing Bayi School International Department, Beijing, China

* Corresponding Author Email: simeng.wang@bayims.cn

Abstract. With the popularity of electric vehicles, accurately predicting charging demand in the workplace has become the key to optimising the grid load and improving the operational efficiency of charging stations. To cope with the uncertainty of users' charging behaviours, this study is based on real charging records from the ACN-Data platform and uses the Random Forest Model to systematically identify the core features that affect charging energy. The analysis results show that charging duration is a decisive factor in predicting charging energy, with a contribution of more than sixty percent. The amount of charging requested by the user is also crucial, with a contribution of twenty-one per cent. This suggests that the user's own charging behavioural characteristics are central to driving demand. This study further provides key insights and data to support the development of high utility and low computational cost charging management strategies. In the future, lightweight random forest models can be explored to be embedded in the edge computing units of charging stations for intelligent scheduling decisions. Building an adaptive charging network with user behaviour prediction as the core will be a key direction to achieve grid interaction optimization and enhance user experience.

Keywords: Charging behaviour strategies, Electric vehicles, Machine learning.

1. Introduction

With the rapid development and widespread adoption of the global electric vehicle industry, workplace charging facilities have become an important part of urban energy infrastructure. However, the inherent randomness and uncertainty of users' charging behavior pose severe challenges to the operational efficiency of charging stations and the load balance of the regional power grid. In this context, how to accurately predict the short-term charging demand at workplaces and optimize power resource allocation, alleviate peak grid pressure, and enhance user experience has become one of the key issues urgently to be addressed in the current smart grid and smart transportation fields. Currently, research in this area has shifted from early statistical analysis to more complex machine learning modeling [1-3]. Many scholars have attempted to use models such as support vector machines and neural networks to predict charging loads, and have preliminarily verified the influence of factors such as time periods and temperature [4][5]. However, existing research mainly focuses on macro load prediction and has relatively shallow exploration of the micro user decision-making characteristics that drive charging behavior and their complex coupling mechanisms with the external environment (such as temperature). At the same time, research based on real, large-scale scenario data is still relatively scarce, which limits the practicality and interpretability of the constructed models. Therefore, this study, based on the ACN-Data platform covering real charging records for twelve consecutive months, introduces the random forest regression model, aiming to systematically identify the core characteristics affecting charging energy, deeply reveal the joint effect of user behavior patterns and physical constraints, and hope to provide key insights and data support for the development of highly practical and low-computational-cost charging management strategies. This study analyzes real charging data, predicts the charging energy at workplaces, identifies key influencing factors, and proposes optimization scheduling strategies.

2. Method

2.1. Dataset

This study utilizes the workplace electric vehicle charging dataset provided by the ACN-Data platform, which encompasses approximately 9,850 complete charging order records. The data collection period spans a continuous twelve months, fully covering different seasonal cycles and typical user charging behavior patterns. Each record includes multiple dimensions of features such as the start time, end time, actual charging duration, charging energy, charging equipment type, environmental temperature, meteorological parameters, and the date and time of the charging event. Charging duration and charging energy are the core behavioral indicators, while environmental temperature and meteorological conditions provide important references for analyzing the impact of external factors. This dataset has significant advantages in terms of sample size and feature completeness, and can effectively support the systematic analysis and modeling research on workplace electric vehicle charging behaviors [6][7].

2.2. Model selection and construction

In this study, a random forest regression model was employed for charging load prediction, mainly due to its unique advantages in handling complex nonlinear systems [8][9].

2.2.1 Requirements for modeling nonlinear relationships

The charging behavior of electric vehicles in the workplace is a typical multi-factor coupled system, where variables such as charging duration, environmental temperature, and time distribution have complex interactions and nonlinear characteristics. Traditional linear models are unable to fully capture this multi-dimensional complexity. However, random forests, through the mechanism of ensemble learning, can automatically identify the nonlinear correlations among variables, providing a more flexible framework for modeling charging behavior.

2.2.2 Fine characterization of temperature sensitivity

The influence of environmental temperature on charging efficiency exhibits a significant non-uniform characteristic, showing differentiated sensitivity levels in different temperature ranges. Random forest, through its inherent decision boundary learning ability, can adaptively identify the key thresholds of temperature influence, achieving segmented and refined modeling of temperature sensitivity, thereby more accurately reflecting the complex patterns of energy consumption changes under low-temperature conditions.

2.2.3 Intelligent identification of time period effects

The daily distribution of charging demand exhibits distinct multi-peak characteristics and time-period dependencies. There are systematic differences in user behavior patterns across different time periods. The random forest model can automatically learn the typical charging characteristics of each time period from the data, effectively capturing the electricity usage patterns during weekdays and weekends, as well as peak and off-peak periods, and achieving intelligent identification and modeling of time period effects.

3. Results

3.1. Data Visualization Analysis

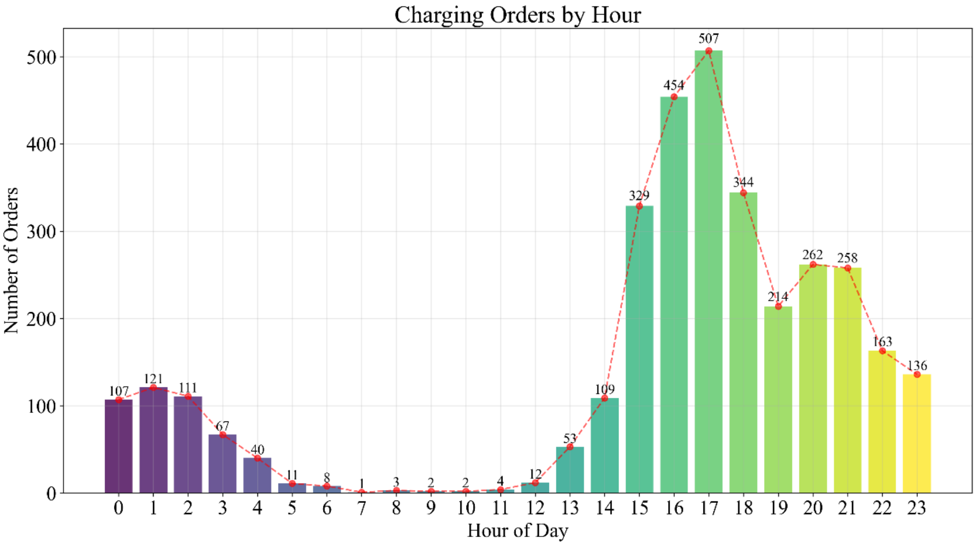

Figure 1. Daily Distribution Characteristics of Charging Orders

Figure 1 presents the distribution characteristics of charging orders within 24 hours, showing a distinct bimodal structure, which profoundly reflects the intrinsic relationship between users' travel behavior and charging demands. The secondary peak around 10 a.m. coincides with the immediate charging needs of users after arriving at their workplaces; while the peak throughout the day from 3 p.m. to 5 p.m., with an order volume exceeding 500 per hour, is highly corresponding to the concentrated charging behavior before leaving for work in the afternoon. The order volume is nearly zero during the early morning and dawn period, further indicating that this scenario is mainly daytime usage; although there is still some charging activity in the evening, the intensity is significantly lower than the afternoon peak. The overall distribution is highly consistent with the typical commuting rhythm, revealing that workplace charging behavior is deeply influenced by users' daily routines, and demonstrating the strong regularity and predictability of charging demands in the time dimension.

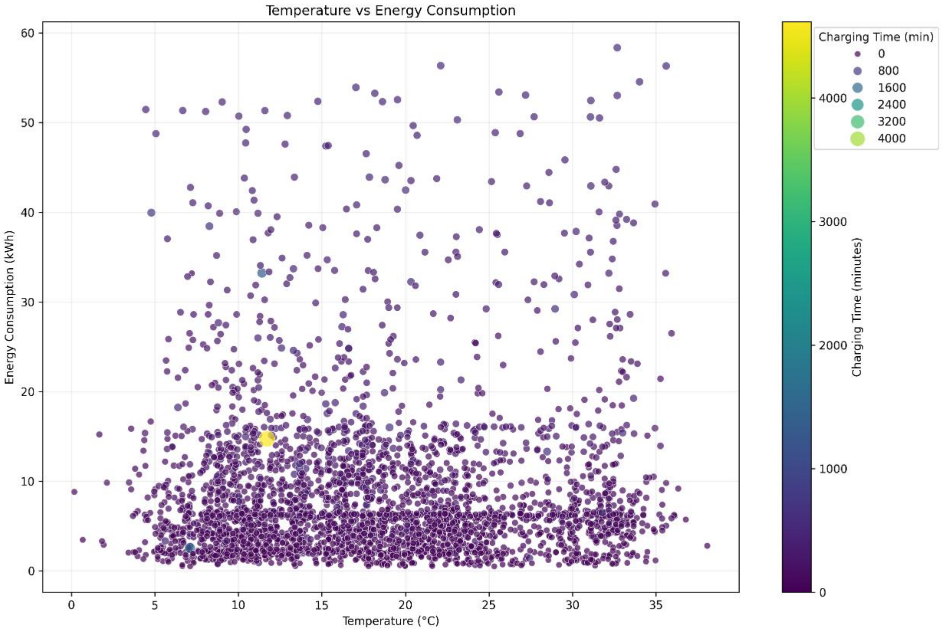

Figure 2. Visual Analysis of Charging Duration Based on the Relationship between Temperature and Energy Consumption

Figure 2 reveals the significant impact of environmental temperature on the energy consumption of a single charge, demonstrating the dual mechanism through which temperature conditions, in conjunction with battery chemical properties and user behavior, influence the charging process. In low-temperature environments, the increase in internal resistance of the battery leads to a decrease in energy conversion efficiency, while users tend to increase their charging demands to address battery range degradation, jointly pushing up the energy consumption level. As the temperature enters the suitable range of 20–25°C, the battery efficiency approaches its optimal state, users' concerns about range reduction are alleviated, and charging behavior becomes more planned, resulting in a more concentrated and stable distribution of energy consumption. A few high-energy consumption discrete points in the figure correspond to cases of long charging periods, possibly due to the deep replenishment needs of specific user groups or special vehicle usage scenarios. This phenomenon indicates that temperature not only affects the battery performance itself but also indirectly shapes the overall characteristics of charging energy consumption by altering users' psychological expectations and behavioral patterns.

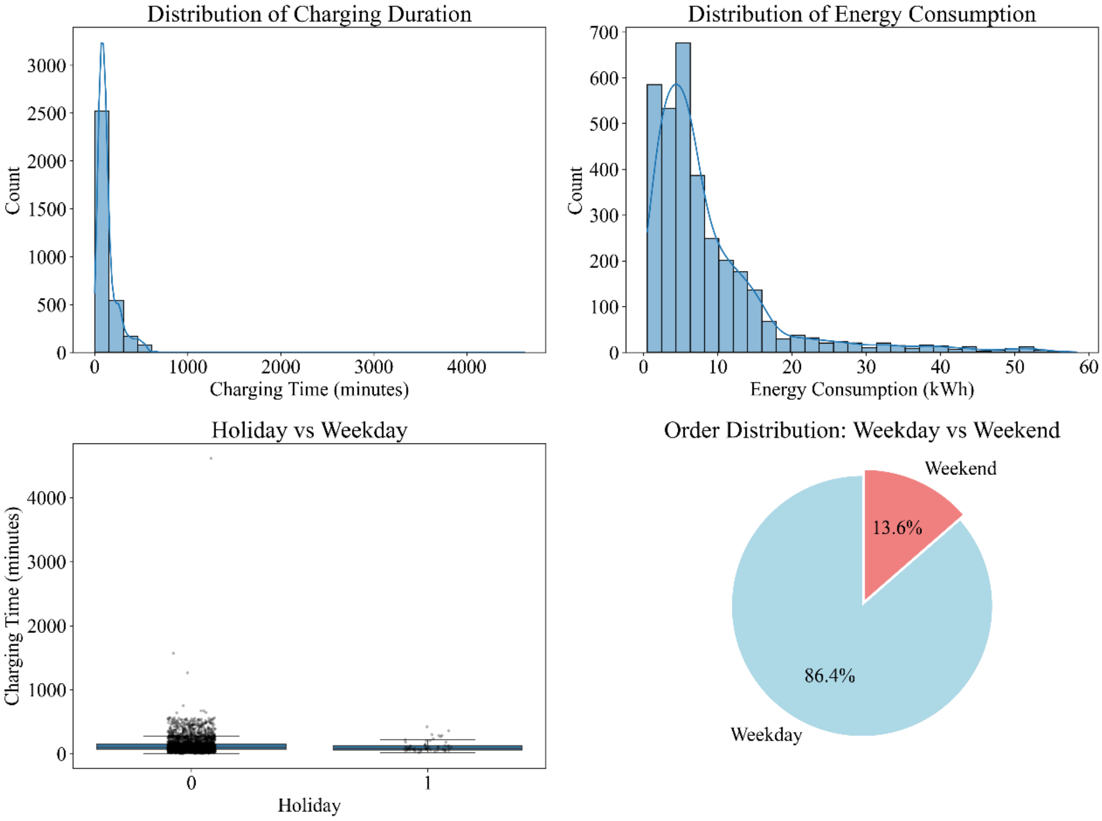

Figure 3. Multi-dimensional Feature Analysis of Charging Behavior

In Figure 3, the analysis of "charging time" reveals the typical charging behavior patterns of users in the working environment. The single charging duration shows a highly right-skewed distribution. The vast majority of charging behaviors are concentrated within 200 minutes, with the peak occurring in shorter periods, indicating that users generally use workplace charging for television as a supplementary energy supply during daily commutes, and tend to complete rapid charging during working hours. A small number of extremely long charging records may correspond to special usage scenarios, such as long-term business trips or abnormal parking situations. These abnormal patterns are worthy of separate analysis in subsequent studies.

The analysis of energy consumption in Figure 3 further confirms the "recharging" characteristic of workplace charging. The distribution of single-charging energy consumption shows that the energy consumption is concentrated in the range of 5-15 kWh, which precisely meets the range requirements for daily commutes. A few high-energy consumption orders may correspond to vehicles with larger battery capacities or special vehicle needs. This distribution pattern indicates that users are more inclined to complete partial charging at the workplace rather than fully charging, reflecting a rational charging strategy based on user needs.

Figure 3 compares the charging patterns during holidays and working days to reveal the impact of user behavior flexibility on charging duration. Although the median charging durations for both periods are similar, holidays show greater dispersion and more long-duration charging records, indicating that in an environment with weaker time constraints, users' charging behaviors are more flexible. This reflects that workplace charging not only meets basic travel needs but may also serve more diverse leisure activity scenarios. The date distribution characteristics presented in Figure 3 reinforce the scene dependence of workplace charging. Working days account for 86.4% of the total orders, highlighting the strong correlation between this scenario and professional activities. This highly concentrated usage pattern not only verifies the accuracy of the positioning of workplace charging facilities but also provides an important basis for the efficient allocation of charging resources, indicating that special attention should be paid to facility carrying capacity and scheduling optimization during the peak working day hours.

3.2. Evaluation of Prediction Model Performance

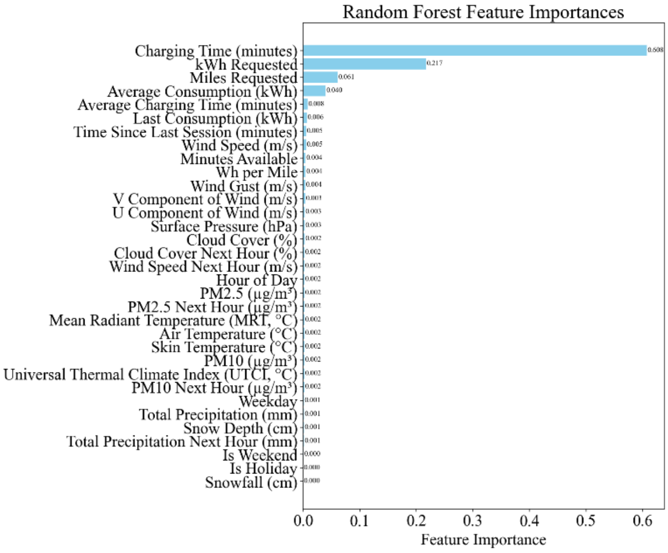

Figure 4. Analysis of Feature Importance Calculated by Random Forest Model

To verify the effectiveness of multi-factor coupling prediction, this study constructed a random forest regression model based on charging duration, environmental temperature, charger type, and time period characteristics to predict charging energy. The evaluation results of the prediction performance based on the random forest regression model showed that this model demonstrated excellent fitting ability and generalization performance in the charging demand prediction task. Specifically, the coefficient of determination (R^2) between the predicted values and the actual values reached 0.887, indicating that the model could effectively explain 88.7% of the variation in the target variable. At the same time, the average absolute percentage error (MAPE) of the model prediction was 22.06%, and the root mean square error (RMSE) was 3.054. This result reflects that the model maintains a high explanatory power while having acceptable prediction accuracy, which can provide reliable support for charging demand prediction in practical application scenarios. This result indicates that machine learning methods can effectively capture the nonlinear characteristics of charging behavior, providing a reliable technical path for high-precision prediction of charging load in work environments. The results of the feature importance analysis calculated by the random forest model in Figure 4 show that there are significant differences in the contribution of each feature variable to the charging energy prediction. Among them, charging duration (target.chargingTime) was identified as the most influential predictor, with an importance score of 0.6078, accounting for more than 60% of the total explanatory power. This result is in line with the physical nature of energy transmission, indicating that charging energy is mainly governed by the objective physical law of charging time.

The user's request for charging capacity (behavior.kWhRequested) was regarded as the second most important feature, contributing 21.67% of the explanatory power, highlighting the significant influence of user behavior preferences on charging energy. The planned driving distance (behavior.milesRequested) and the vehicle's average energy consumption (behavior.avgConsumption) each accounted for 6.12% and 3.96% of the importance, respectively, reflecting the auxiliary predictive role of travel demands and vehicle energy efficiency characteristics in determining charging needs.

In contrast, the importance of other behavioral characteristics (including average charging duration, last charging energy consumption, and charging time interval, etc.) is all below 1%, indicating that these variables provide only limited supplementary information. Overall, the distribution of feature importance clearly reveals the dual influence mechanism of charging energy prediction by both physical laws and user behavior, providing a theoretical basis for feature selection and model optimization.

4. Discussion

Based on the aforementioned visualization analysis results, this paper proposes a dynamic charging facility scheduling strategy for typical scenarios. In response to the charging peak phenomenon that occurs during the afternoon working hours, it is recommended to implement a resource allocation mechanism based on demand elasticity. By establishing a time-based reservation and dynamic power regulation system, while ensuring the charging needs of high-priority users, a moderate power restriction is imposed on non-urgent charging requests, and the peak-valley price leverage is used to guide users to shift to non-peak hours, thereby balancing the system load and user experience. The power company can conduct ultra-precise predictions of regional charging loads for the next few minutes to several hours based on the real-time usage duration of charging piles and user request data. This provides key data support for real-time power grid scheduling, frequency stability, and prevention of overloading.

In low-temperature environment scenarios, the significant increase in charging energy consumption requires the adoption of a dedicated scheduling plan that adapts to temperature changes. It is recommended to prioritize the allocation of charging resources to vehicles with battery pre-heating functions, and automatically increase the basic power configuration during cold periods to compensate for battery efficiency losses. At the same time, a winter charging guarantee zone should be established to provide exclusive service channels for users with concerns about battery range. This temperature compensation mechanism can effectively alleviate the system pressure caused by the decline in charging efficiency in low-temperature environments.

In response to the behavioral differences between working days and holidays, an elastic scheduling system should be established. During working days, a standardized service model should be adopted, with a focus on ensuring the charging supply during commuting hours. During holidays, a flexible management model should be employed, allowing for a relaxation of the single charging duration limit and increasing temporary charging points to meet the diverse travel needs. This differentiated resource allocation strategy not only conforms to the user behavior characteristics in each scenario but also effectively improves the utilization efficiency of facilities.

5. Conclusion

This study is based on the real charging data from the ACN-Data platform and uses the random forest model to conduct an in-depth analysis of the charging behavior of electric vehicles in workplaces. The research found that the charging behavior has a high degree of regularity, presenting a bimodal distribution within the day that coincides with the commuting rhythm, and the charging energy is mainly driven by two core factors: charging duration and the amount of power requested by the user. The model successfully captured these nonlinear relationships and achieved high-precision predictions. Based on this, the study proposed dynamic scheduling strategies for peak hours, low-temperature environments, and differences between working days and holidays, such as time-based reservations, power regulation, and temperature compensation mechanisms. The significance of this study lies in providing data-driven decision-making basis for charging station operations and grid load optimization, and revealing the dual mechanism where user behavior and physical laws jointly shape charging demands. Looking to the future, lightweight machine learning models can be embedded in the edge computing units of charging stations to build an adaptive charging network centered on user behavior prediction, ultimately achieving the dual goals of grid interaction optimization and user experience improvement.

References

[1] Zhang Y, et al. A comparative study of statistical vs. machine-learning methods for EV charging-load forecasting. IEEE Transactions on Smart Grid, 2022, 13(4): 3123-3135.

[2] Wang Z, Wang J. Short-term charging load forecasting using hybrid SVM-XGBoost with adaptive weighting. Applied Energy, 2021, 304: 117611.

[3] Liu Siyuan, et al. Electric vehicle charging load forecasting based on deep belief network. Proceedings of the CSEE, 2023, 43(5): 1823-1832.

[4] Luo X, et al. Non-linear effects of ambient temperature on electric vehicle charging energy: A quantile regression approach. Energy, 2021, 237: 121589.

[5] Wang Lei, et al. A modified electric vehicle charging load model considering ambient temperature sensitivity. Power System Technology, 2020, 44(9): 3382-3389.

[6] Lee J, Lee H. ACN-Data: One-year publicly available dataset of 2.1 million charging sessions in California. Scientific Data, 2022, 9: 412.

[7] Zhang Hongcai, et al. Charging behavior profiling based on full log data from public charging operators in Shanghai. Electric Power Information and Communication Technology, 2023, 21(3): 45-52.

[8] Zhao P, et al. Identifying key determinants of electric vehicle charging energy using random forest regression: Evidence from 50,000 charging events. Energy Reports, 2021, 7: 4400-4409.

[9] Chen Qiujie, et al. Importance analysis of influencing features for electric vehicle charging energy based on random forest. High Voltage Engineering, 2022, 48(8): 3101-3109.