

Article

An Intelligent Profiling and Classification Method for Load Adjustment Potential of Multi-Type Demand-Side Resources Considering Adjustment Willingness

Can Wang, Xuesong Shao, Shihai Yang, Huiling Su * and Yingwen Zhu

State Grid Jiangsu Marketing Service Center, Nanjing 210019, China; stategrid_cw@163.com (C.W.); shaoxs@js.sgcc.com.cn (X.S.); yangshihai@js.sgcc.com.cn (S.Y.); zhuyw8@js.sgcc.com.cn (Y.Z.)

* Correspondence: shl@js.sgcc.com.cn; Tel.: +86-13913823628

Abstract

The rapid development of new energy has caused a sharp increase in the stochasticity on the source side of the new power system (NPS), and extreme weather along with climate variability have also led to increased stochasticity in power demand on the load side; thus, how to achieve source-load matching and enable the load to track the source under the new situation is the key to the efficient operation of the power system. Aiming at the problem that existing load regulation potential evaluation mainly focuses on physical capacity, making it difficult to reflect users' subjective willingness to participate as well as the dynamic changes in regulation capability under different operating scenarios, this paper proposes a two-stage dynamic profiling classification method for multi-type power user loads considering regulation willingness. First, an evaluation index system is constructed from three dimensions, physical reliability, execution reliability, and behavioral willingness, to achieve the unified characterization of the regulation capabilities of heterogeneous resources such as industrial loads and electric vehicle (EV) aggregators. Second, the DBSCAN algorithm is adopted to identify typical annual operating scenarios. Finally, the Dynamic Time Warping (DTW) distance is introduced to improve the K-Means++ algorithm, achieving the profiling classification of user regulation potential. This paper takes a certain NPS demonstration park as an example for verification, and the results show that the annual operating scenarios can be divided into 4 types of typical days; the proposed DTW-K-Means++ method has better classification performance compared with traditional Euclidean distance clustering, can effectively identify the differences and dynamic migration characteristics of user regulation potential under different operating scenarios, and stably classifies users into three types of profiles: deep regulation type, agile response type, and rigid constraint type. The research results aim to provide reliable data support for the refined dispatch of the power grid by effectively quantifying the dynamic migration patterns of heterogeneous resources under variable scenarios.

Keywords: load regulation potential; user profiling; regulation willingness; DBSCAN clustering; DTW-K-Means++ algorithm

Academic Editors: Peter Van den Bossche and Grzegorz Sierpiński

Received: 11 May 2026

Revised: 16 June 2026

Accepted: 17 June 2026

Published: 29 June 2026

Copyright: © 2026 by the authors.

Published by MDPI on behalf of the

World Electric Vehicle Association.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](#)

[Attribution \(CC BY\)](#) license.

1. Introduction

The large-scale integration of new energy has promoted the rapid development of the new power system (NPS), but it has also significantly enhanced the stochasticity and

volatility of source-side output. At the same time, the frequent occurrence of extreme weather and complex, variable climate conditions have further increased the uncertainty of power demand on the load side [1]. Under the new situation where the stochasticity on both the source and load sides is simultaneously enhanced, how to achieve precise source-load matching and “load tracking the source” has become a key issue in ensuring the safe and stable operation of the power system and improving the accommodation level of new energy. Demand Response (DR), as an important regulation measure connecting the supply and demand sides, can fully excavate the flexible regulation capability on the load side and is one of the important approaches to achieving load tracking of the source [2].

With the continuous enrichment of DR resource types, heterogeneous resources such as industrial loads, electric vehicle (EV) aggregators, integrated energy systems (IES), and data centers are gradually becoming important entities for the flexible regulation of the power grid. Meanwhile, to alleviate the operational pressure on the distribution network caused by the uncoordinated charging of large-scale EVs, utilizing distributed renewable energy such as solar energy to promote vehicle-grid coordinated operation has also become a research hotspot in recent years [3]. In addition to EVs, multi-type demand-side resources such as multi-park IESs [4], data center clusters [5], community IESs [6], and industrial enterprises are gradually participating in Integrated Demand Response (IDR), becoming important entities supporting the flexible regulation of the NPS. However, due to significant differences in capacity scale, operational characteristics, and geographical distribution among various resources, massive heterogeneous resources present highly complex spatio-temporal distribution characteristics [7], causing the aggregated load to exhibit strong stochasticity and uncertainty. Therefore, accurately evaluating the load regulation potential of multi-type power users and constructing user profiles that can reflect the differences in their regulation capabilities have become an important foundation for improving the management and dispatch level of DR resources.

Currently, research on the evaluation of user-side load regulation potential has made certain progress. Existing studies mainly rely on user load characteristics and historical operational data, utilizing Non-Intrusive Load Monitoring (NILM) technology [8] as well as artificial intelligence and data-driven methods [9] to identify user groups with regulation capabilities, such as extracting the characteristics of flexible resources like buildings to construct user profiles to evaluate DR potential [10] or conducting potential evaluation on adjustable loads in industrial parks based on profiling technology [11]. Such studies can well reflect the objective operational characteristics of users, but their evaluation basis mainly focuses on the physical capacity of equipment and historical electricity consumption behavior, rarely considering the subjective decision-making factors when users actually participate in DR. In fact, user response behavior is often jointly influenced by multiple factors such as compensation revenue expectations, travel plans, energy consumption preferences, and satisfaction requirements. For example, residential users have differentiated demands for power rationing compensation levels [12], EV users possess obvious heterogeneity in travel behavior [13,14], and some users are also affected by social welfare perception and comfort needs [15]. In recent years, some studies have attempted to introduce user willingness factors into DR models, for example, considering DR behavior under a low-carbon game mechanism [16], designing a two-dimensional compensation mechanism [17], and constructing an uncertainty model containing a travel chain with user willingness [18]. With the deepening of research, frontier explorations in the DR field have begun to profoundly recognize the impact of user behavior complexity, and attempt to introduce bounded rationality and behavioral economics theories for modeling. For instance, the latest literature characterizes the heterogeneous bounded rationality and asymmetric risk perception of EV users based on prospect theory [19], incorporates abnormal behaviors such as response delays and insufficient duration into the IDR

optimization model [20], and further discusses the fatigue and learning effects of demand-side behavior under multi-scale extreme weather [21]. However, from a critical perspective, these studies still have obvious limitations: existing behavioral modeling mostly revolves in isolation around a single type of load (such as pure EVs or pure industrial loads), or focuses on theoretical games at the macroscopic planning level; there is still a lack of data-driven methods that incorporate the subjective participation willingness of multi-type heterogeneous resources like EVs and industrial loads into a unified framework for comprehensive quantification, leading to a relatively large deviation between the theoretical regulation capacity and the actual adjustable potential, making it difficult to accurately characterize the true regulation capabilities of heterogeneous load resources from a system level.

Furthermore, in terms of user profile construction, most existing studies establish a unified static label system or typical load curves based on continuous annual operating data [22,23]. Although the latest research, such as literature [24], proposes an EV charging behavior profiling method considering users' subjective psychological effects, its limitation lies in that the evaluation results are still presented as a static probability set for a single scenario. However, during the actual operation process of the power grid, user load characteristics undergo dynamic migration influenced by factors such as seasonal changes, meteorological conditions, and holidays; especially under typical scenarios like extreme heat, extreme cold, and major holidays, users' energy consumption behavior and response capabilities often exhibit obvious differences. Traditional static profiles are difficult to reflect the dynamic evolution patterns of user regulation potential under different operational scenarios, which easily leads to misjudgments by the dispatch department regarding available flexible resources during critical periods, thereby affecting the precise identification of DR resources. Therefore, constructing dynamic user profiles capable of reflecting the differences in typical operating scenarios holds important practical significance.

Regarding profiling classification methods, traditional studies mostly adopt clustering algorithms based on Euclidean distance and require pre-setting the number of clusters. Such methods are not only relatively sensitive to abnormal noise caused by factors like extreme weather, but also lack an effective clustering validity verification mechanism. Aiming at massive and complex load data, existing studies have utilized the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm to dynamically classify EV users, verifying its adaptive identification advantages in processing multidimensional feature data with noise [25]. Meanwhile, for heterogeneous resources such as industrial loads and EVs, their DR behaviors usually exhibit characteristics like time delay or random translation; traditional distance metric methods are difficult to effectively handle the time-series phase shift problem, which easily causes classification errors. To this end, some studies introduce the Dynamic Time Warping (DTW) algorithm to measure the time-series similarity of users' electricity consumption behaviors [26], or apply it to the optimization of building thermal load control [27]; there are also studies adopting the K-Means++ algorithm to improve regional load forecasting performance [28]. However, how to combine DBSCAN, DTW, and cluster analysis while simultaneously taking into account the differences in time-series characteristics of heterogeneous loads and the characteristics of users' regulation willingness, to achieve the construction of high-precision and high-robustness load profiles, still requires further research.

Aiming at the above problems, this paper proposes a two-stage dynamic profiling classification method for multi-type power user loads considering regulation willingness. First, an evaluation index system for load regulation potential is constructed from three dimensions: physical reliability, execution reliability, and behavioral willingness, incorporating the subjective and objective characteristics of multi-type loads such as EVs and

industrial enterprises into a unified evaluation framework to achieve a comprehensive characterization of the regulation capabilities of heterogeneous resources. Second, the DBSCAN algorithm is utilized to divide annual operating data into typical days, constructing dynamic user profiles oriented towards different operating scenarios, and revealing the migration patterns of load regulation potential as operational scenarios change. Finally, the DTW distance is introduced to improve the K-Means++ clustering algorithm, reducing the impact of time-series phase shifts on the profiling classification results, and improving the accuracy and robustness of heterogeneous load classification. The research method in this paper differs from traditional static evaluation by accurately characterizing the true regulation flexibility of heterogeneous resources under variable spatio-temporal environments through data-driven means, thereby providing data support for the dispatch department to formulate differentiated strategies.

2. Framework of Intelligent Profiling Process for Load Regulation Potential of Multi-Type Power Users Considering Regulation Willingness

To achieve the effective quantification and hierarchical profiling of power users' regulation potential, it is necessary to construct a multidimensional evaluation system encompassing physical characteristics, execution efficacy, and behavioral psychology. Considering that there are significant differences in the regulation mechanisms of different types of loads, and that user characteristics exhibit dynamic changes along with the external environment and typical day environments, the load characteristics of multi-type power users are first defined; subsequently, taking the time scale as the core variable, quantitative indicators applicable to specific typical days are defined from the three dimensions of physical reliability, execution reliability, and behavioral willingness, respectively. This is shown in Figure 1.

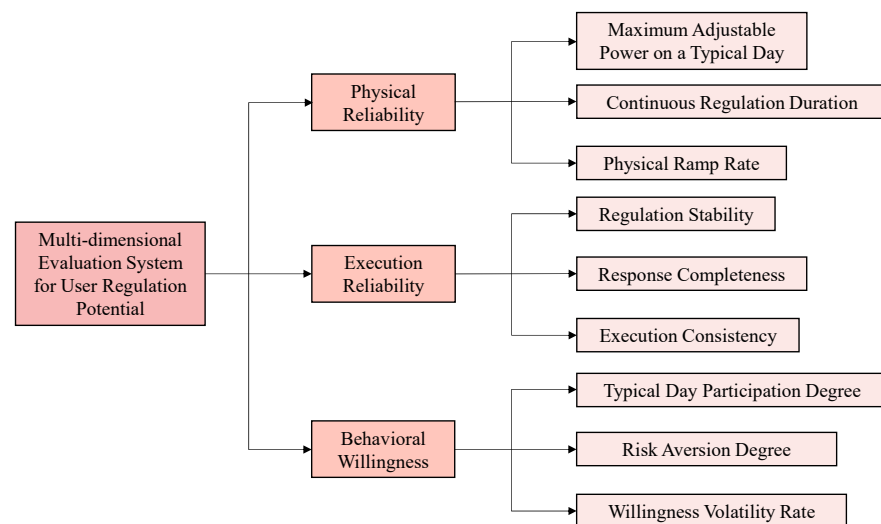


Figure 1. Framework of the multi-dimensional evaluation index system for user regulation potential.

2.1. Overall Framework of Dynamic Profiling for Load Regulation Potential

As shown in Figure 2, this framework mainly includes three core modules: data preprocessing, typical day identification in the time dimension, and profiling classification in the user dimension. The data cleaning and standardized preprocessing target the historical operating data of multi-type demand-side resources. It first removes anomalous values

and performs min-max normalization on heterogeneous indicators of different dimensions, providing standardized data input for subsequent algorithms. The first stage starts from the macroscopic time dimension, extracts the daily average load rate, daily peak-to-valley difference rate, and daily load fluctuation rate as operating features, and utilizes the DBSCAN density clustering algorithm to adaptively classify the operating scenarios of 365 days throughout the year. This stage aims to isolate four types of typical scenarios—normal production days, extreme peak load days, deep valley rest days, and meteorological anomaly days—providing a macroscopic environmental baseline for microscopic evaluation. Under the premise of clarifying the constraints of specific typical day scenarios, the second stage quantifies the regulation potential indicators of multi-type users from three dimensions: physical reliability, execution reliability, and behavioral willingness, respectively. Finally, the improved DTW-K-Means++ algorithm is introduced to eliminate the time-series phase shifts among heterogeneous loads, achieving the effective classification of user profiles (deep regulation type, agile response type, and rigid constraint type).

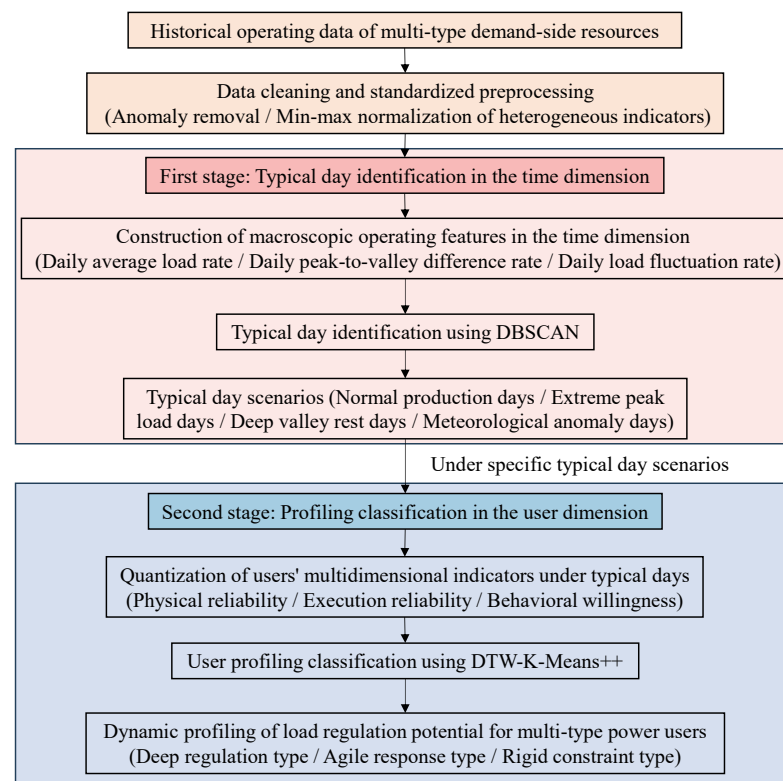


Figure 2. Overall process framework of dynamic load profiling considering regulation willingness.

2.2. Definition and Characteristic Analysis of Multi-Type Power User Loads

Multi-type power users primarily encompass industrial users and EV aggregators with demand response potential. These two types of loads have essential differences in electricity consumption behavior, regulation mechanisms, and constraint conditions. Among them, as the core mobile nodes in the heterogeneous resource system, EVs are not only a key link connecting the transportation network and the distribution network but also the core application target of the willingness profiling and classification method in this study.

Industrial users can generally be further subdivided into discrete manufacturing types and continuous production types. The production processes of discrete manufacturing users (such as machining enterprises) are relatively independent and possess strong time-shifting characteristics. They can participate in peak shaving and valley filling

by adjusting shifts or starting/stopping non-critical equipment, thus offering large regulation potential and high stability. The production processes of continuous production users (such as fine chemical enterprises) have a high degree of coupling. Their regulation flexibility is limited by process safety and rigid delivery requirements, typically only providing limited reduction capability under security load constraints.

EV aggregators, as mobile energy storage resources, possess dual attributes. In the spatial dimension, EV connection locations change with users' travel trajectories; in the temporal dimension, their regulation capability is strictly constrained by the State of Charge (SOC) of the battery, the expected grid-departure time, and the travel willingness of the owners. Compared to industrial loads, EV loads exhibit significant randomness and instantaneous response capability.

To ensure the physical interpretability and engineering computability of the multidimensional evaluation indicators for load regulation potential, this paper constructs a data mapping link composed of 'bottom-layer measurement data–basic physical characteristics–advanced abstract indicators'. In the bottom-layer data acquisition stage, relying on the real distribution network marketing system and the DR management platform, the user's basic load sequence $P_{i,t}^{base}$ and actual response power sequence $P_{i,t}^{res}$ with a time resolution of Δt , the real-time State of Charge (SOC) of vehicles under the jurisdiction of the EV aggregator, and the external incentive signal sequence S_t^{inc} are collected. Subsequently, through cleaning and normalization processing, the above-mentioned concrete physical measurement data are substituted into the multidimensional evaluation system constructed later, achieving the quantitative mapping of users' unobservable subjective willingness and risk preferences.

2.3. Indicators for the Physical Reliability Dimension

It characterizes the objective regulation boundary and resource baseline of the user under a specific typical day. This dimension integrates the physical operational constraints of the equipment and the user's tolerance for risk, which together constitute the effective capacity boundary for the user to participate in DR.

1. Maximum adjustable power on a typical day

This metric characterizes the average limit of power adjustment that a user can provide within the time periods of a specific typical day, constrained by equipment-rated parameters and operational status. It characterizes the extreme average value of power regulation that the user can provide within the time period contained in a specific typical day, limited by the equipment's rated parameters and operational state.

$$P_{i,d}^{max} = \frac{1}{N_T} \sum_{t \in T_d} (P_{i,t}^{base} - P_{i,t}^{min}) \quad (1)$$

where $P_{i,d}^{max}$ is the average maximum adjustable power of the i -th user under a typical day d (MW); $t \in T_d$ is the set of time steps contained in a specific typical day, and N_T is the total number of time steps; $P_{i,t}^{base}$ is the baseline load of user i at time t (MW); $P_{i,t}^{min}$ is the minimum rigid load allowed for equipment operation (MW).

For the EV aggregator, if V2G reverse discharging is considered, the calculation of $P_{i,d}^{max}$ needs to satisfy the battery physical constraints:

$$P_{v,t}^{V2G,max} = \min \left(P_v^{dis,max}, \frac{(SOC_{v,t} - SOC_{v,exp})C_v^{cap}}{\Delta T_{park} \eta_{V2G}} \right) \quad (2)$$

where $P_{v,t}^{V2G,max}$ is the maximum discharge power limit of the v -th vehicle (MW); $P_v^{dis,max}$ is the physical power upper limit of the charging and discharging equipment; $SOC_{v,t}$ and $SOC_{v,exp}$ are the battery SOC of the v -th vehicle at the current time and the expected grid-departure time, respectively; C_v^{cap} is the rated battery capacity of the v -th vehicle (MWh); ΔT_{park} is the remaining parking duration (h); η_{V2G} is the comprehensive V2G discharge efficiency.

2. Effective continuous regulation duration

It quantifies the length of time that the user can continuously respond when maintaining the maximum regulation power output, limited by the energy storage state or the rigid constraints of the production process within a typical day.

$$T_{i,d}^{dur} = \frac{E_{i,d}^{flex}}{P_{i,d}^{max}} \quad (3)$$

where $T_{i,d}^{dur}$ is the equivalent continuous regulation duration (h); $E_{i,d}^{flex}$ is the total flexible energy that can be dispatched by user i during typical day d (MWh).

For industrial users, $E_{i,d}^{flex}$ depends on the total power consumption of shiftable processes or the available capacity of thermal and cold storage equipment within the factory area. For the EV aggregator, $E_{i,d}^{flex}$ depends on the sum of the current remaining available discharge capacity of the battery packs of all online vehicles within its jurisdiction, and its specific calculation formula is as follows:

$$E_{i,d}^{flex} = \sum_{v=1}^{N_{ev}} (SOC_{v,t} - SOC_{v,min}) C_v^{cap} \eta_{V2G} \quad (4)$$

where N_{ev} is the total number of currently online vehicles under the jurisdiction of the aggregator; $SOC_{v,t}$ is the current SOC of the v -th vehicle at time t ; $SOC_{v,min}$ is the minimum power threshold to ensure that the owner of the v -th vehicle meets the return travel demand; C_v^{cap} is the rated physical battery capacity of the v -th vehicle (MWh); η_{V2G} is the comprehensive V2G discharge efficiency.

3. Physical regulation rate limit

It reflects the maximum physical power ramping up or ramping down capability that the user equipment can achieve per unit time. For continuous production industrial users, this indicator is limited by the safe and stable operation constraints of the production line; for EV charging and discharging facilities, this indicator is limited by the inverter topology and the short-circuit capacity limit at the grid connection point.

$$R_i^{lim} = \max_t \frac{|P_{i,t}^{base} - P_{i,t-1}^{base}|}{\Delta t} \quad (5)$$

where R_i^{lim} is the physical regulation rate limit (MW/h); $P_{i,t}^{base}$ and $P_{i,t-1}^{base}$ are the baseline load values at adjacent times; Δt is the time step. This indicator directly determines whether the user can quickly reach the target position in short-term, high-frequency emergency peak-shaving commands.

2.4. Execution Reliability Dimension Indicators

In evaluating the degree of coincidence between the actual power curve and the target curve after the user receives the dispatch command, it reflects the user's regulation precision and control quality under a specific typical day.

1. Regulation stability coefficient

It describes the degree of fluctuation deviation of the actual response power relative to the target command during single or multiple regulation events within a typical day for the user.

$$C_{i,d}^{stab} = 1 - \frac{1}{M} \sum_{m=1}^M \frac{D_{DTW}(P_{i,m}^{res}, P_{i,m}^{obj})}{D_{max}} \quad (6)$$

where $C_{i,d}^{stab}$ is the regulation stability coefficient (value range [0, 1]); M is the total number of regulation events occurring within a typical day; $D_{DTW}(\cdot)$ is the DTW distance function, used to eliminate time-series phase shifts; $P_{i,m}^{res}$ and $P_{i,m}^{obj}$ are the actual response power sequence and the target response power sequence of user i in the m -th event, respectively (MW); D_{max} depends on the maximum advance notice time window allowed by the power market rules (such as 24 h for day-ahead invitation or 2 h for intra-day agile regulation). A response dislocation exceeding this time window has physically lost the practical significance of peak shaving and valley filling, so it is used as the upper penalty limit of the warping distance for normalization processing.

2. Response completion rate

It represents the supply–demand matching degree between the regulation energy actually delivered by the user within a typical day period and the target energy issued by the dispatch center.

$$\gamma_{i,d}^{comp} = \frac{\sum_{m=1}^M E_{i,m}^{res}}{\sum_{m=1}^M E_{i,m}^{obj}} \quad (7)$$

where $\gamma_{i,d}^{comp}$ is the cumulative response completion rate of user i on typical day d ; $E_{i,m}^{res}$ and $E_{i,m}^{obj}$ are the integral electricity quantities of the actual delivery and the target demand in the m -th event, respectively (MWh).

3. Execution consistency index

It is used to examine whether there is significant dispersion in the performance quality of the user during multiple response behaviors under similar typical days.

$$I_{i,d}^{cons} = 1 - \frac{\sigma(\gamma_{i,m}^{comp})}{\mu(\gamma_{i,m}^{comp})} \quad (8)$$

where $I_{i,d}^{cons}$ is the execution consistency index; $\sigma(\gamma_{i,m}^{comp})$ and $\mu(\gamma_{i,m}^{comp})$ are the standard deviation and mean functions of the user's M single response completion rates ($\gamma_{i,m}^{comp} = E_{i,m}^{res} / E_{i,m}^{obj}$), respectively.

2.5. Behavioral Willingness Dimension Indicators

The regulation willingness of power users in actual DR is driven by their utility trade-offs and risk perceptions under specific scenarios. Based on Prospect Theory and the Bounded Rationality hypothesis in behavioral economics, the true willingness utility of users accepting dispatch is determined by the asymmetric game between expected

revenue (positive utility) and perceived risk loss (negative utility), and its micro-theoretical fundamental equation can be expressed as:

$$U = v(\Delta R) - \lambda v(\Delta C) \quad (9)$$

where $v(\Delta R)$ is the value function brought by economic compensation; $v(\Delta C)$ is the risk cost brought by changing energy consumption habits or battery degradation; λ is the loss aversion coefficient. Since micro-psychological parameters such as λ are hidden variables that cannot be directly observed in actual power grid engineering, this paper abandons the traditional subjective questionnaire survey method and instead adopts a data-driven ‘behavioral trajectory backward-deduction’ strategy. Namely, assuming that the user’s historical DR compliance behavior is the ultimate external manifestation of their internal utility game, by extracting statistical characteristics from real interaction records, the following three engineering-computable macroscopic mapping indicators are constructed, thereby achieving the quantitative characterization of the theoretical utility model without invading user privacy:

1. Typical day participation rate

According to random utility theory, whether a user accepts the invitation from the dispatch center is essentially a game between the positive utility brought by economic compensation and the negative utility brought by changing energy consumption habits. When the incentive level breaks through the user’s psychological threshold, the user will exhibit a positive “stimulus-response”. The proposed typical day participation rate is exactly the statistical mapping of this utility-maximizing rational decision-making result, reflecting the user’s initial willingness baseline under the current incentive system and the specific typical day environment.

$$\alpha_{i,d} = \frac{N_{i,d}^{act}}{N_{i,d}^{inv}} \quad (10)$$

where $\alpha_{i,d}$ is the participation rate of user i on typical day d ; $N_{i,d}^{act}$ is the number of times actually confirmed and executed; $N_{i,d}^{inv}$ is the total number of invitations issued by the system.

2. Risk aversion factor

It reflects the user’s sensitivity to uncertainty factors (such as default assessment and electricity price fluctuations) under a specific typical day. This is a dimensionless scalar coefficient that characterizes the user’s tendency to proactively shrink their actual available physical boundaries. To eliminate the dimensional differences between power and incentive signals, extreme value normalization processing is first performed on the historical actual response power sequence $P_{i,t}^{res}$ and the external incentive signal sequence S_t^{inc} to obtain the dimensionless sequences $\tilde{P}_{i,t}^{res}$ and $\tilde{S}_t^{inc}$.

$$\rho_{i,d} = 1 - \frac{Var(\tilde{P}_{i,t}^{res})}{Var(\tilde{S}_t^{inc})} \quad (11)$$

where $\rho_{i,d}$ is the risk aversion factor, with a value range between $[0, 1]$; $Var(\cdot)$ is the sequence variance calculation function. If the user is completely driven by economic interests, their response power will closely follow the fluctuations of the incentive signal, and the normalized variance of the two will tend to be consistent, at which point $\rho_{i,d} \approx 0$; conversely, if the EV owner adopts a conservative strategy out of concern for battery degradation, regardless of how high the incentive is, they will maintain a relatively small

response fluctuation, then $Var(\tilde{P}_{i,t}^{res}) \ll Var(\tilde{S}_t^{inc})$, and a larger numerical difference indicates a deeper degree of risk aversion.

Under the completely rational ‘economic man’ hypothesis, in pursuit of maximizing DR revenue, the fluctuation of the user’s actual regulation power should remain absolutely synchronous with the fluctuation amplitude of the external incentive signal. However, when limited by stringent industrial production safety constraints or the EV owner’s range anxiety regarding battery attenuation, the user will exhibit a significant loss aversion tendency. In order to avoid potential default risks or energy consumption inconvenience, the user will actively suppress the response amplitude, causing the actual power curve to still maintain a smooth fluctuation when facing high incentives. Therefore, the more significant the aforementioned normalized variance difference is, the more it proves from the physical external manifestation that the user has a deeper degree of risk aversion in their internal decision-making logic.

3. Willingness volatility rate

Prospect theory in behavioral economics points out that humans exhibit significant bounded rationality and loss aversion when facing uncertainty and risks. Under complex power grid operating scenarios (such as extreme peak load days or meteorological anomaly days), the industrial user’s concern about power outage default losses, or the EV owner’s perception of range anxiety, will be asymmetrically amplified. This dynamic change in risk perception causes the user’s decision-making logic to frequently swing between pursuing economic revenues and avoiding energy consumption risks. By constructing the willingness volatility rate indicator, this degree of behavioral swing caused by the intertwining of bounded rationality and environmental uncertainty is effectively quantified.

$$V_{i,d}^{will} = \sqrt{\frac{1}{K} \sum_{k=1}^K (\alpha_{i,k} - \bar{\alpha}_{i,d})^2} \quad (12)$$

where $V_{i,d}^{will}$ is the willingness volatility rate; K is the total number of sub-intervals into which the typical day period is divided; $\alpha_{i,k}$ is the participation rate of user i in the k -th sub-interval; $\bar{\alpha}_{i,d}$ is the average participation rate within that typical day.

3. Two-Stage Dynamic Profiling Clustering Model Based on Spatio-Temporal Decoupling

After completing the construction of the multidimensional evaluation indicators for the regulation potential of multi-type power users, to alleviate the dimensionality and time-series phase shift issues caused by the massive data of heterogeneous resources, this section proposes a two-stage clustering model based on DBSCAN and the improved DTW-K-Means++ algorithm. The first stage focuses on the macroscopic time dimension, performing dimensionality reduction on the operating days throughout the year through adaptive density clustering; the second stage delves into the microscopic user dimension, introducing the dynamic time warping distance to effectively classify user profiles.

3.1. First Stage: Typical Day Identification Model in the Time Dimension

3.1.1. Construction of Macroscopic Operating Features in the Time Dimension

Based on the macroscopic fluctuation conditions of the power grid load, this paper extracts three key indicators: the daily average load rate, the daily peak-to-valley difference rate, and the daily load fluctuation rate, to construct a three-dimensional vector that characterizes the daily load environment features. These three features can effectively

reflect the overall supply and demand pressure and regulation requirements of the power grid under different meteorological conditions and holiday arrangements. Due to the inherent differences in the distribution intervals of the above feature indicators, to avoid a single feature dominating the subsequent distance calculations, it is necessary to utilize the min-max normalization method in advance to uniformly map the three-dimensional feature vectors of the 365 days throughout the year to the $[0, 1]$ interval, thereby eliminating dimensional deviations.

3.1.2. DBSCAN Density Clustering Algorithm

The load behavior of power users is significantly influenced by day types and meteorological conditions, exhibiting diverse operating patterns. Since the number of typical operating days across the 365 days of the year is unknown, and power grid operating data often contains outliers caused by line maintenance, sudden faults, or collection errors, the traditional K-Means algorithm—which requires a predefined number of clusters and is sensitive to noisy data—is inadequate to meet the requirements of typical day extraction at this stage.

In contrast, DBSCAN does not rely on a predefined number of clusters; instead, it partitions cluster structures by deriving the density distribution of the sample space. This algorithm can effectively identify clusters of arbitrary shapes and automatically filter out low-density noise points, making it highly suitable for extracting representative typical days from massive daily load data.

The DBSCAN algorithm constructs clusters by defining core objects and their neighborhood densities. Assume a daily load feature dataset containing N samples $D = \{x_1, x_2, \dots, x_N\}$, where x_d is the normalized load feature vector of the d -th day. The algorithm is based on two core parameters: neighborhood radius ϵ and minimum number of samples $MinPts$.

Define the ϵ -neighborhood of sample x_i as the set of samples in space with a distance to x_i less than or equal to ϵ :

$$N_\epsilon(x_i) = \{x_j \in X \mid \text{dist}(x_i, x_j) \leq \epsilon\} \quad (13)$$

If the number of samples contained in the ϵ -neighborhood of sample x_i is greater than or equal to $MinPts$, i.e., $|N_\epsilon(x_i)| \geq MinPts$, then x_i is called a core object.

Based on core objects, the following concepts of density connectivity are defined:

(1) Directly density-reachable: If sample x_j is located in the ϵ -neighborhood of core object x_i , then x_j is said to be directly density-reachable from x_i .

(2) Density-reachable: If there is a sample sequence $p_1, p_2, \dots, p_n$, where $p_1 = x_i$, $p_n = x_j$, and p_{k+1} is directly density-reachable from p_k , then x_j is said to be density-reachable from x_i .

The algorithm forms a cluster by finding all density-reachable core objects and their boundary points and connecting them. Isolated points that are not density-reachable from any core object are marked as noise.

3.1.3. Typical Day Identification Process Based on DBSCAN

The typical day classification and identification process based on DBSCAN is shown in Figure 3, with specific steps as follows:

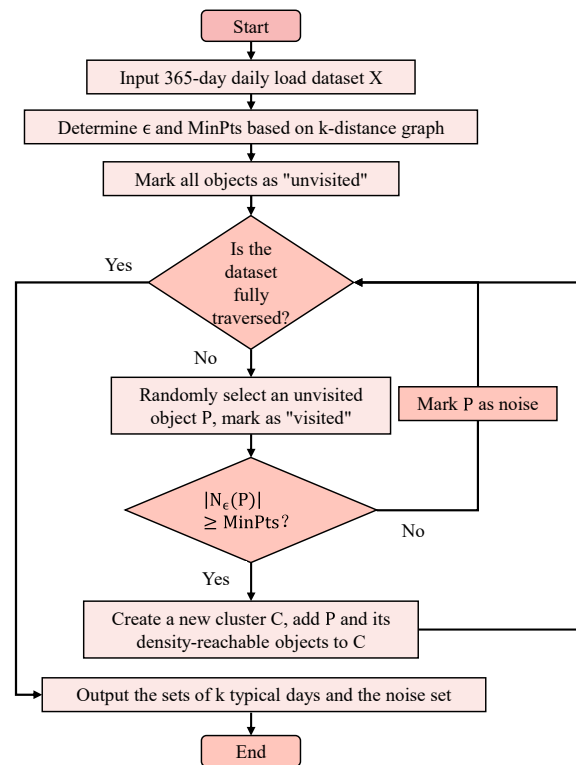


Figure 3. Flowchart of typical day classification and identification based on DBSCAN.

(1) Input the preprocessed full-time daily load feature matrix. Determine the optimal neighborhood radius ϵ and the minimum number of samples $MinPts$ using the inflection point of the sorted distance curve observed from the k -distance graph.

(2) Traverse the dataset, calculate the ϵ -neighborhood density of each sample point, and filter out the set of core objects that meet the density threshold requirement.

(3) Starting from any unvisited core object, find all connected core objects and boundary points through the density-reachable relationship to form an independent cluster; isolated samples that fail to be classified into any cluster are marked as noise data and removed.

(4) Iteratively execute the above process until all samples have been visited, and output the typical day set $\Omega = \{C_1, C_2, \dots, C_m\}$, where m is the adaptively identified number of typical days by the algorithm.

3.2. Second Stage: Profiling Classification Model in the User Dimension

3.2.1. DTW Distance Metric

After completing the first-stage time-dimension typical day dimensionality reduction based on the DBSCAN algorithm, the evaluation object transitions from macroscopic daily environmental characteristics to microscopic user multidimensional response characteristics. Aiming at the different data characteristics and regulation demands of these two stages, this paper adopts two different clustering algorithms, aiming to achieve the effective decoupling of time scenarios and user profiles:

In order to effectively overcome the phase shifts in heterogeneous resources in the microscopic user profiling stage and satisfy the engineering demand of the power grid to stably divide regulation echelons, this paper transforms the clustering strategy in the second stage and introduces an improved algorithm based on center partitioning. Specifically, the first stage focuses on macroscopic scenario identification. Since the number of

annual operating scenario categories is unknown, and they are highly susceptible to the interference of incidental extreme weather such as typhoons and plum rains, the DBSCAN algorithm, which does not require pre-setting the number of clusters and has strong anti-noise capability, is selected, enabling the adaptive extraction of typical scenarios and the effective stripping of meteorological abnormal noise. The second stage focuses on microscopic profile construction. Under the determined typical days, continuing to adopt density clustering can no longer meet the requirements, mainly for two reasons: At the physical level, the response actions of multi-type users often exhibit a time-series phase shift phenomenon of 'similar in shape but asynchronous in time'. At this time, traditional density clustering or Euclidean distance is highly prone to misjudging this natural dislocation of time as morphological differences. At the engineering level, actual power grid dispatch needs to stably and clearly divide massive users into specific categories with clear regulation echelon significance (such as deep regulation type, agile response type, etc.), rather than irregularly shaped density clusters. In summary, this stage is more suitable for adopting a center-based partitioning strategy.

Therefore, in the user profiling classification stage, this paper shifts the clustering strategy by introducing the Dynamic Time Warping (DTW) distance to mitigate the impact of time-series phase deviations, and combines it with the K-Means++ algorithm, which features a more stable initialization strategy, to achieve the refined classification of the regulation potential of multi-type heterogeneous resources.

On a specific typical day, the response behaviors of different users may exhibit characteristics that are similar in shape but asynchronous in time. The traditional Euclidean distance, based on point-to-point rigid alignment calculations, misinterprets such phase shifts in the time dimension as significant shape differences. To this end, this paper introduces the Dynamic Time Warping algorithm as the similarity metric standard.

Suppose there are two 24 h normalized response power sequences of users to be aligned, denoted as $P_A = (a_1, a_2, \dots, a_n)$ and $P_B = (b_1, b_2, \dots, b_m)$. DTW aims to find an optimal warping path $W = \{w_1, w_2, \dots, w_K\}$ that minimizes the cumulative distance along the path. This path must satisfy boundary, continuity, and monotonicity constraints.

The DTW distance is defined as the sum of distances along the optimal path:

$$DTW(P_A, P_B) = \min_W \left(\sum_{k=1}^K d(w_k) \right) \quad (14)$$

where $d(w_k)$ is the Euclidean distance at the path node. This problem can be solved using the dynamic programming method, and the recurrence formula for its cumulative distance matrix $D(i, j)$ is:

$$D(i, j) = d(a_i, b_j) + \min \begin{cases} D(i-1, j) \\ D(i, j-1) \\ D(i-1, j-1) \end{cases} \quad (15)$$

where $D(i, j)$ represents the shortest cumulative distance between the first i points of sequence P_A and the first j points of sequence P_B .

3.2.2. Improved K-Means++ Clustering Algorithm

The traditional K-Means algorithm randomly selects initial clustering centers, making it prone to falling into local optimal solutions and resulting in unstable clustering performance. The K-Means++ algorithm addresses this issue by optimizing the initialization strategy, based on the core principle that the distances between initial clustering centers should be as far apart as possible.

Suppose that r centers have been selected. For each sample x , its DTW distance to the nearest existing center is calculated as $D(x)$. The probability of selecting the next center x' , denoted as $Prob(x')$, is given by:

$$Prob(x) = \frac{D(x')^2}{\sum_{x \in X} D(x)^2} \quad (16)$$

This mechanism effectively avoids the crowding of initial centers, significantly improving the convergence speed and clustering accuracy of the algorithm.

3.2.3. Two-Stage Profiling Classification Process Based on DTW-K-Means++

The distance metric during K-Means++ clustering uses a mixed-weight distance. For time-series features containing 24 h operational characteristics (e.g., daily load response power sequences), DTW is used to calculate the distance to eliminate phase deviations; for static scalar indicators such as physical reliability and willingness participation degree, Euclidean distance or standardized absolute distance is still used for calculation; finally, the weighted sum of the two is taken as the overall distance metric for clustering.

The specific execution steps of the DTW-K-Means++ algorithm are shown in Figure 4:

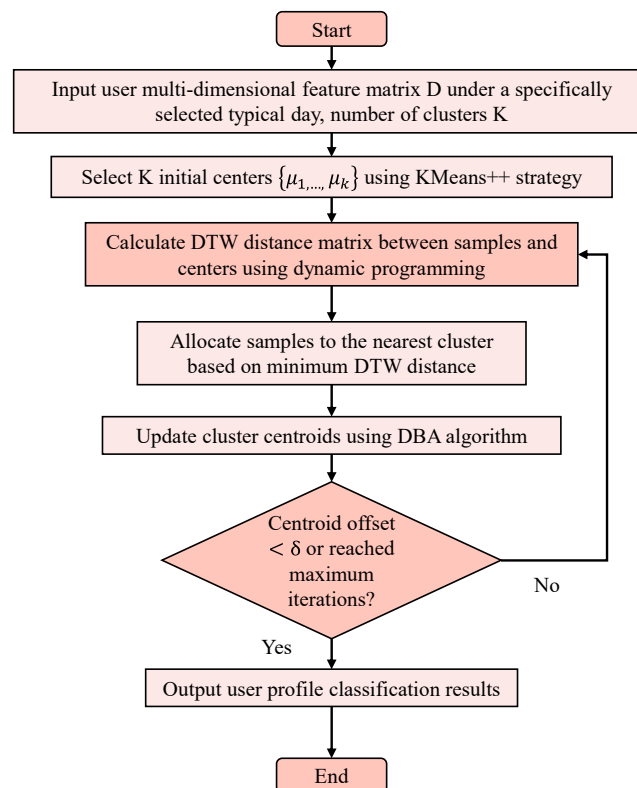


Figure 4. Flowchart of user profile classification based on DTW-K-Means++.

(1) Select the multi-dimensional user feature matrix under a specific typical day as input data and set the number of clusters K .

(2) Use the dynamic programming algorithm to calculate the DTW distance matrix between all pairs of user samples.

(3) Randomly select the first sample as the first cluster center, based on the probability formula $P(x')$, use the roulette wheel selection method to iteratively select the remaining $K-1$ initial centers.

(4) Traverse all user samples and, based on the principle of minimum DTW distance, allocate each sample to the cluster of its nearest cluster center.

(5) Due to the non-linear characteristics of the DTW space, the arithmetic mean is no longer applicable. Adopt the DTW Barycenter Averaging (DBA) algorithm to calculate the average sequence of samples within each cluster to update the cluster centers.

(6) Repeat steps 4 and 5, calculating the offset of the cluster centers. If the offset is less than the preset convergence threshold or the maximum number of iterations is reached, stop the iteration.

(7) Output the final user regulation potential profile classification results and the typical characteristic curves of each category.

4. Case Study

4.1. Data Sources and Multidimensional Feature Data Preprocessing

4.1.1. Case Study Data and Scenario Settings

To verify the applicability of the proposed method to heterogeneous resources, two typical types of demand response resources—industrial users and the EV aggregator—in a national-level NPS demonstration park in southern Jiangsu, China, are selected as the research objects. This region is affected by continuous high temperatures in summer and is vulnerable to typhoons and plum rains during the transition season between summer and autumn; moreover, it has dense industrial loads and a high EV penetration rate, making its source-load heterogeneous characteristics and meteorological sensitivity highly representative. The data acquisition time span is the entire year of 2023, the sampling interval is 1 h, and there are a total of 8760 time cross-sections throughout the year. To ensure the objectivity of the willingness modeling, this paper established strict user selection criteria during data collection: the selected users must have complete high-frequency metering data over the past year, and have retained at least 20 real day-ahead/intra-day DR invitation and execution interaction records in the power grid marketing system. To enhance the transparency and reproducibility of this data-driven research, the specific geographical and meteorological characteristics, load subdivision composition, current time-of-use electricity pricing rules, and DR project design parameters of the case study park are extracted in detail and summarized in Table 1.

Table 1. Basic load composition and market rule parameters of the case study park.

Data Dimension	Core Parameters
Geographical and Meteorological Location	Southern Jiangsu Province, China; characterized by extreme heat in summer, with typical plum rain and typhoon seasons
Overall Load Composition	300 typical users: 200 industrial users, 100 EV aggregators
Industrial Load Subdivision	Discrete manufacturing enterprises (100 users, IDs #001–#100): machining, electronic assembly, etc., with equipment start-stop flexibility.
	Continuous production enterprises (100 users, IDs #101–#180, #281–#300): chemical, metallurgical, etc., subject to rigid process safety constraints.
EV Infrastructure	Commercial district charging stations (40 users, IDs #181–#220): dominated by 120 kW DC fast chargers, heavily influenced by daytime passenger flow.
	Residential district charging stations (60 users, IDs #221–#280): dominated by 7 kW AC slow chargers/coordinated chargers, with concentrated load at night.
TOU	Peak periods (08:00–11:00, 17:00–22:00): approx. 1.12 CNY/kWh.
	Flat periods (11:00–17:00, 22:00–24:00): approx. 0.65 CNY/kWh.

DR Mechanism	Valley periods (00:00–08:00): approx. 0.31 CNY/kWh.
	Response types: Day-ahead invitation (24 h advance notice) and intra-day agile (2 h advance notice). Compensation benchmark: Effective peak-shaving response compensation is settled at 3.5–4.0 CNY/kWh, with penalty coefficients set for defaults.

To verify the generalization performance of the proposed method under the scenario of large-scale user integration, and simultaneously overcome the problem of limited samples caused by the privacy protection of high-frequency data in the actual distribution network, this paper adopts a ‘real seed + semi-synthetic enhancement’ strategy to construct an evaluation dataset containing 300 user samples. Although obtaining the full amount of measured real data presents objective difficulties, this paper extracts the historical load curves of 30 real large users from the case study park as seed data. Subsequently, data enhancement is performed on the seed data based on Monte Carlo simulation, introducing behavioral stochasticity simulation (superimposing a time translation factor obeying the local traffic congestion index) and energy consumption volatility simulation (superimposing Gaussian white noise). During this process, the physical capacity limits of the equipment are strictly imposed as rigid out-of-limit constraints, ensuring that the generated power fluctuations will not break through the objective boundaries of the real equipment.

Behavioral stochasticity simulation: Combining the local traffic congestion index and scheduling rules, a random time translation factor $\Delta t \in [-2h, +2h]$ is introduced for EVs and discrete manufacturing loads, to simulate the response delay caused by users returning late or using electricity during off-peak hours. Energy consumption volatility simulation: To avoid generating idealized artificial patterns, Gaussian white noise obeying a normal distribution $N(0, \sigma^2)$ is superimposed. The standard deviation σ is constrained to not exceed 5%–10% of the mean value of the baseline load, to ensure that the generated fluctuations will not break through the maximum adjustable boundary of the equipment’s physical capacity.

To verify that the generated 270 synthetic samples truly reflect the actual load behavior and do not produce an artificial “ideal clustering morphology”, this paper introduces the Kolmogorov–Smirnov (K-S) statistical test to compare the probability density distributions of the synthetic dataset and the real seed dataset. The test results show that the P -values of both at the confidence level of $\alpha = 0.05$ are greater than 0.05, accepting the null hypothesis, which proves that the synthetic data highly inherit the physical distribution morphology and thermodynamic evolution patterns of the real load. Figure 5 displays the original annual load time-series curves of typical users. Through macro-micro joint visualization technology, this paper extracts a normal spring week (h 2000–2120) and an extreme peak load summer week (h 5000–5120) in the Figure as local magnified windows, respectively. From the magnified views, the heterogeneous characteristics of the continuous flat base load of industrial users intertwined with the nighttime charging spikes of EVs can be clearly observed, further confirming the reliability of the evaluation dataset in this paper.

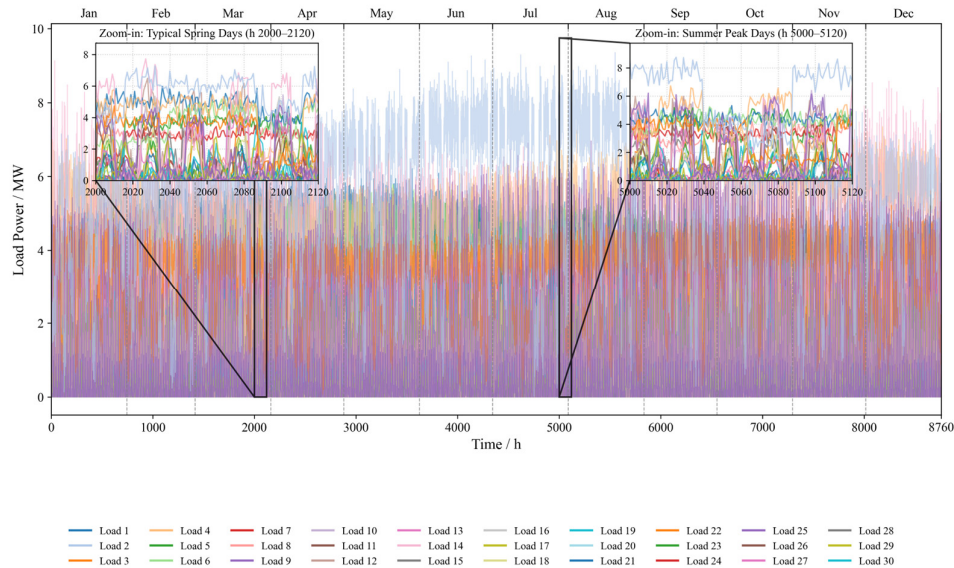


Figure 5. Original annual load curves of different types of typical users.

4.1.2. Normalization of Heterogeneous Indicators and Aggregation of Differentiated Weights

Since the indicators of the three dimensions—physical reliability, execution reliability, and behavioral willingness—have different dimensions and physical meanings, and there are differences between positive and negative indicators, standardized processing must be performed prior to clustering. Meanwhile, considering the essential differences between industrial users and electric vehicle aggregators in terms of physical constraint rigidity and behavioral decision-making flexibility, it is necessary to design a transparent standardized and differentiated weighting procedure.

1. Normalization of heterogeneous indicators based on the min-max method

To eliminate dimensional impacts and uniformly map all features to the [0, 1] interval, this paper adopts the min-max normalization method

For positive indicators where larger values are better, such as $P_{i,d}^{max}$, $\gamma_{i,d}^{comp}$, and $\alpha_{i,d}$, the normalization formula is:

$$\tilde{x}_{i,j} = \frac{x_{i,j} - \min(X_j)}{\max(X_j) - \min(X_j)} \quad (17)$$

For negative indicators where smaller values are better, such as $\rho_{i,d}$ and $V_{i,d}^{will}$, the normalization formula is:

$$\tilde{x}_{i,j} = \frac{\max(X_j) - x_{i,j}}{\max(X_j) - \min(X_j)} \quad (18)$$

where $x_{i,j}$ is the original calculated value of the j -th indicator for the i -th user; X_j is the set of this indicator for all users under a specific typical day; $\tilde{x}_{i,j}$ is the standardized scalar feature after normalization.

2. Design of differentiated weight matrix for resource heterogeneity

The physical operating constraints and behaviors of industrial users and electric vehicle aggregators are different. Industrial users (especially the continuous production type) are restricted by rigid production processes and safety bottom lines, making the

boundaries of their physical reliability dimension extremely strict; conversely, electric vehicle aggregators belong to behavior-driven resources, possessing relatively small absolute physical capacity, but their execution flexibility and subjective participation willingness play a decisive role in the final response results.

Therefore, targeting different resource types, differentiated dimensional weight vectors $\mathbf{W} = [w_{phy}, w_{exe}, w_{beh}]$ are constructed, corresponding to physical reliability, execution reliability, and behavioral willingness, respectively. The weights satisfy the normalization constraint: $\sum w = 1$. Industrial users (especially the continuous production type) are restricted by rigid production processes and safety bottom lines, making the boundaries of their physical reliability dimension extremely strict, and occupying an absolute dominant position in regulation.

To verify the scientific validity of the above differentiated weighting strategy based on business logic and to avoid the bias caused by subjective weighting, this paper introduces the CRITIC (Criteria Importance Through Intercriteria Correlation, an objective weighting based on correlation) method and Spearman rank correlation analysis to conduct an ex-post quantitative verification on the historical empirical DR event data of the case study park. Calculating the Spearman correlation coefficient matrix between indicators of different dimensions and the actual final compliance rate of users reveals that: for industrial users, the correlation coefficient between the “maximum regulable power on typical days” in their physical dimension and the final compliance rate is as high as 0.82, showing extremely strong physical capacity dominance; for electric vehicle aggregators, the correlation coefficient between behavioral willingness indicators such as “willingness volatility” and “risk aversion factor” and the final compliance rate reaches 0.76, indicating that the dominant role of subjective willingness significantly surpasses that of single physical capacity. Based on the above data-driven verification results, this paper configures benchmark weight vectors with different emphases for different resources: assigning industrial users a weight vector leaning towards the physical dimension, $\mathbf{W}_{IND}$, and assigning EV aggregators a weight vector leaning towards the willingness dimension, $\mathbf{W}_{EV}$. Through the above processing, a standardized feature matrix is provided for the subsequent clustering.

4.2. First Stage: Classification, Identification, and Analysis of Typical Days in the Time Dimension

The regulation potential of users is not static but is closely coupled with meteorological conditions and data types. Given the complexity of the annual operational data, this paper first utilizes DBSCAN to perform dimensionality reduction on the park’s operational data for the 365 days of the year, aiming to identify typical days with significant regulation needs from the time dimension, providing a time benchmark for subsequent refined user profiling.

4.2.1. DBSCAN Parameter Sensitivity and Classification Robustness Analysis

To further verify the stability of the classification results of the 365 typical days throughout the year against parameter changes, and to reduce the impact of parameter selection on the clustering results, this paper conducts a two-dimensional grid-based sensitivity analysis on the core parameters of the DBSCAN algorithm—the neighborhood radius (ϵ) and the minimum number of samples ($MinPts$). The parameter test range is set as $\epsilon \in [0.14, 0.22]$ and $MinPts \in [4, 8]$. In addition to adopting the overall Silhouette Coefficient (SC) to evaluate the clustering quality, the number of typical day clusters (Cluster Number) and the noise ratio (Noise Ratio) identified under different parameter combinations are further counted, comprehensively evaluating parameter stability from three

aspects: clustering quality, clustering structure, and abnormal sample identification. The results of each evaluation indicator varying with parameters are shown in Figure 6.

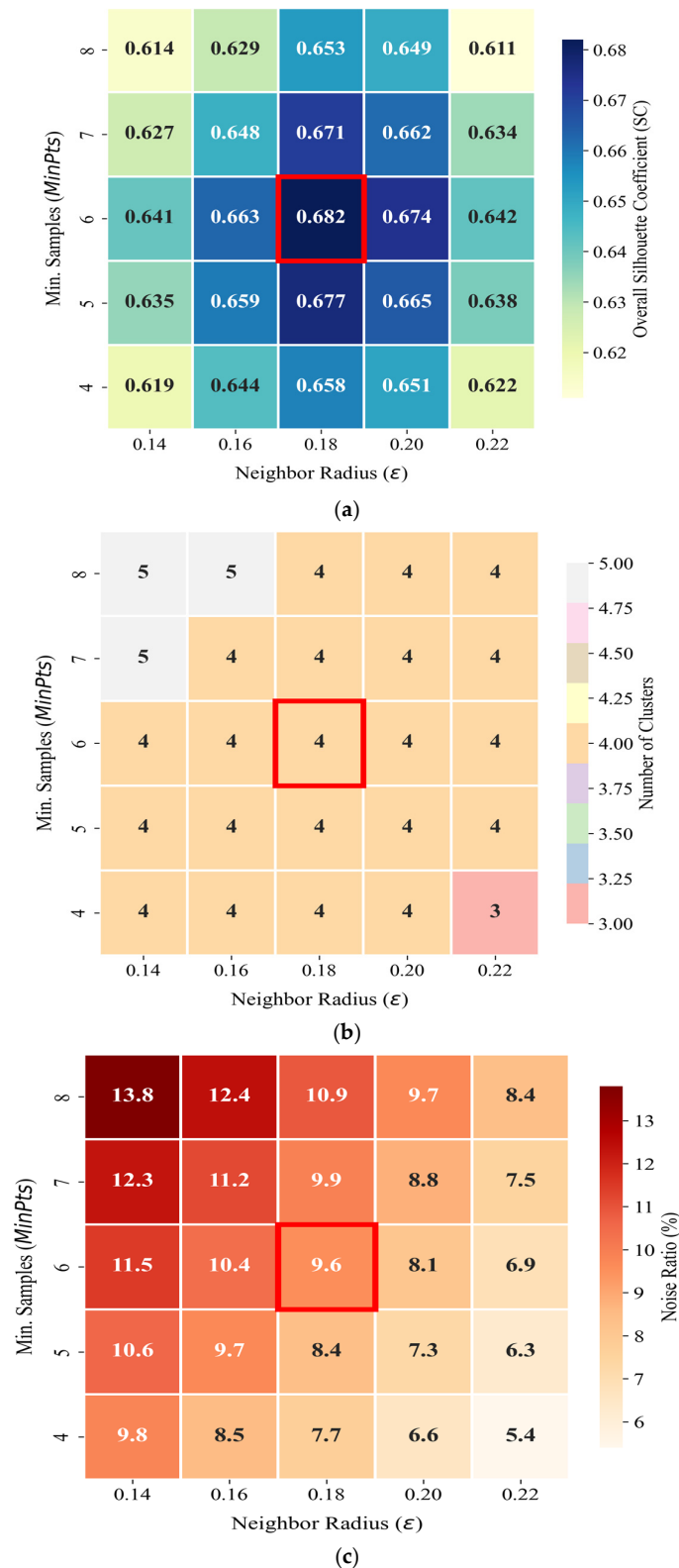


Figure 6. Sensitivity and robustness analysis of DBSCAN parameters. (a) Silhouette coefficient; (b) Cluster number; (c) Noise ratio.

Figure 6 shows that the optimal parameter combination determined based on the k-distance curve ($\epsilon = 0.18$, $MinPts = 6$) is located within the parameter stability region. The solid red boxes in the center of the three subplots intuitively indicate the specific location of this optimal parameter combination within the two-dimensional grids. As can be clearly observed from the numerical values within the red-boxed areas, the corresponding overall SC is 0.682, 4 types of typical days are identified, and the noise ratio is 9.6%. Under this parameter combination, the algorithm can effectively identify abnormal weather samples with obvious differential characteristics while guaranteeing clustering quality, providing a reasonable data foundation for the subsequent typical day division. It can be seen from Figure 6b,c that within the optimal parameter neighborhood where the red boxes are located, the number of typical day categories identified by the algorithm remains stable, with only minor changes occurring in the parameter boundary region; at the same time, the noise ratio presents a smooth changing trend as parameters vary, with no obvious abrupt change phenomenon occurring. Combining the variation pattern of the overall SC in Figure 6a, it can be found that within a reasonable parameter range, all evaluation indicators maintain good stability, indicating that the parameter combination highlighted by the red boxes not only has good clustering quality, but also can maintain a stable clustering structure and reasonable abnormal sample identification capability, possesses good robustness to parameter perturbations, and can meet the practical application needs of typical day identification.

4.2.2. Clustering Results and Feature Analysis of Typical Days

Upon the completion of data normalization, the feature vectors representing the 365 annual days serve as the input for clustering. The core of the DBSCAN algorithm lies in determining the neighborhood radius (ϵ) and the minimum number of points ($MinPts$). Initially, $MinPts$ is set to 6, considering the historical empirical frequency of meteorologically abnormal days in the power grid. Subsequently, a k-distance graph is plotted to identify the optimal neighborhood radius, as illustrated in Figure 7.

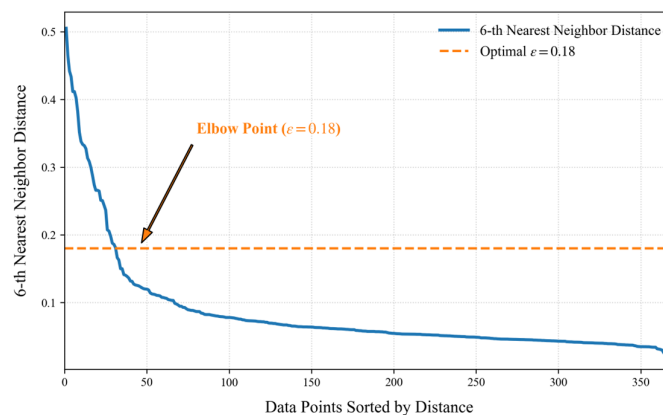


Figure 7. The k-distance graph of the DBSCAN algorithm and the selection of the ϵ inflection point.

The clustering results indicate that the 365 operating days of the year are adaptively partitioned into 4 categories of typical days. The results are presented in Figure 8. As can be seen from Figure 8, for the load fluctuations of a certain city over the 365 days of the year, the results identified through typical day classification can be summarized as follows: normal production days account for 61.6%, extreme peak load days account for 17.8%, deep valley rest days account for 11.0%, and meteorological anomaly days account for 9.6%. These four types of typical days exhibit clear boundary isolation in the three-dimensional feature space. Among them, the blue circular points (normal production

days) have relatively balanced indicators and are concentrated in the middle layer of the space; the orange triangular points (extreme peak load days) have significantly higher daily average load rates and peak-to-valley difference rates, reflecting a massive demand for peak shaving; the green square points (deep valley rest days) cluster in the low-value region, exhibiting deep valley characteristics of low load and low fluctuation.

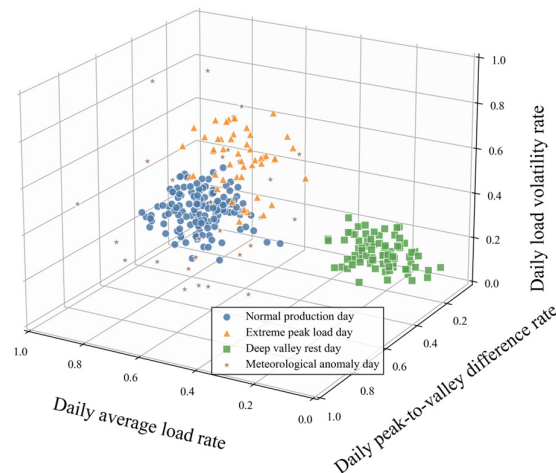


Figure 8. Three-dimensional scatter plot of typical day clustering.

It is worth noting that the meteorological anomaly days represented by purple star-shaped points do not form tight clusters but exhibit a highly discrete and generalized distribution across the entire domain. This exactly reflects the advantage of the DBSCAN algorithm: mechanically, the impact of extreme weather such as plum rains and typhoons on the power grid possesses extremely strong randomness, and some abnormal days show abnormally high daily load volatility rates (0.50 to 0.90). Meanwhile, at the algorithmic level, although a few abnormal days visually appear close to normal production days in the three-dimensional space, their local density fails to meet the connectivity threshold set by DBSCAN due to their accidental morphology. Thus, they are accurately identified by the algorithm as “outlier noise” and isolated. This density-based processing mechanism scientifically explains the boundary between normality and abnormality, effectively avoiding the profile distortion problem caused by forced classification in traditional clustering algorithms (such as K-Means).

The typical operational curves of the four types of typical day loads are shown in Figure 9. Combined with the climate and holiday arrangements of a certain city, the detailed physical characteristics of various typical days are shown in Table 2.

Table 2. Categories of typical days and corresponding date divisions in a certain city.

Typical Day Category	Date Classification	Physical Characteristics
Normal production day	1 March–28 April, 4–31 May, 7 October–30 November	Spring and autumn and climatically suitable periods
Extreme peak load day	15 July–15 August (continuous high temperature), 20–31 December (cold wave)	Summer high temperature or winter extreme cold periods
Deep valley rest day	21–27 January (Spring Festival long holiday), 29 September–6 October (Mid-Autumn and National Day long holiday)	Statutory long holiday shutdown periods
Meteorological anomaly day	15 June–10 July (plum rain season), 20–25 August and 5–10 September (typhoon impact period)	Abnormal weather days such as plum rains, rainstorms, or typhoons

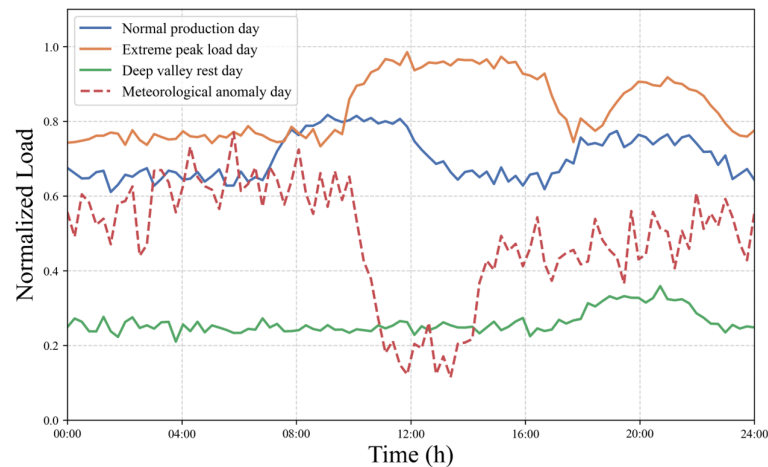


Figure 9. Normalized load curves of typical days in a certain city.

Normal production days are mainly concentrated in non-holiday periods from March to May and from October to November. During these spring and autumn seasons and climatically suitable periods, the temperature is mild, and the proportion of temperature-controlled loads such as air conditioning and heating in the power grid is extremely low. At this time, the production plans of industrial users proceed systematically, the commuting and charging behaviors of EV users are regular and predictable, the overall daily load curve shape is stable, and both the peak-to-valley difference rate and the load volatility rate remain at a low level. The power grid supply and demand pressure is small, and the system has a sufficient regulation margin.

Extreme peak load days mainly correspond to the continuous high-temperature days from mid-to-late July to early August, and the extreme cold wave days in late December. During the summer high temperature or winter extreme cold periods, strongly driven by the sharp climb in industrial and residential temperature-controlled loads (such as large-scale refrigeration or electric heating equipment), coupled with the impact of increased energy consumption per vehicle and higher charging frequency due to EVs turning on air conditioning/heating, the power grid often experiences extremely steep peak loads. At this time, the daily average load rate is at an annual high, the peak-to-valley difference significantly widens, and it is a critical node where the supply-demand contradiction is most prominent throughout the year, making it highly prone to triggering system capacity violation warnings and local overloads.

Deep valley rest days are typically represented by production shutdown periods during statutory long holidays such as the Spring Festival and National Day holidays. During this period, the vast majority of industrial enterprises and large and medium-sized commercial complexes in the whole society enter a state of shutdown and rest, retaining only basic security loads and maintenance loads. This causes the overall electricity consumption level to plunge into the annual minimum range, forming a distinct “deep valley” characteristic. Due to the extremely low daytime load base, if good lighting or wind conditions occur simultaneously, the large-scale centralized output of distributed renewable energy will face the dilemma of having no load for local consumption, easily triggering wind and solar curtailment phenomena. The core pain point of this type of typical day shifts from “peak shaving” to the severe challenges of renewable energy consumption and valley-filling peak regulation.

Meteorological anomaly days, on the other hand, are concentrated in the heavy rainfall days during the city’s plum rain season or severe weather days directly affected by typhoons in summer and autumn. Extreme meteorological conditions will severely

disrupt normal social operations and production order: on the one hand, causing a sudden drop in residents' willingness to travel, a sharp decrease in the load of EV public charging networks, and irregular spatio-temporal shifts; on the other hand, heavy rain and thick clouds will lead to violent fluctuations and even instant plunges in distributed photovoltaic output. The superimposition of dual strong uncertainties on both the source and load sides results in an extremely high daily load volatility rate on this type of typical day, and the system operational state exhibits extremely strong randomness and vulnerability.

In light of the different load characteristics of the above four types of typical days, differentiated yet mutually coordinated dispatching matching strategies need to be formulated in actual operation. Extreme peak load days are the critical nodes where the supply-demand contradiction is most acute and resources are most needed for deep regulation. To avoid facing the dilemma of no available resources on extreme peak load days, dispatching departments should guide some industrial users to transfer high-energy-consuming production tasks to normal production days in advance, utilizing the sufficient grid margin on normal production days to digest electricity demand, thereby reducing the rigid load base on extreme peak load days.

For normal production days, the regulation strategy should focus on economic optimization, guiding users to perform normalized peak shaving and valley filling through daily time-of-use electricity pricing mechanisms, and paying attention to retaining the response quotas of high-quality users to avoid causing regulation fatigue on extreme peak load days due to excessive dispatch.

For deep valley rest days, since the overall load is extremely low and renewable energy output is relatively vigorous, the core goal of dispatching shifts is to promote renewable energy consumption during concentrated power generation periods. Incentive mechanisms should be adopted to encourage resources with strong consumption capabilities, such as EV aggregators, to charge centrally during peak photovoltaic or wind power output periods.

For meteorological anomaly days, due to drastic load fluctuations and high forecasting difficulty, the dispatching strategy should pivot toward enhancing system robustness, reserving sufficient agile response resources with fast response speeds and high execution reliability as backups to cope with sudden power shortages or surges, ensuring the physical security of the grid under extreme disturbances. Through cross-typical-day production capacity transfer and differentiated dispatching matching, the comprehensive regulation potential of multi-type users can be maximized. The clear cluster-like distribution of the four types of typical days in the three-dimensional feature space, along with high physical boundary distinguishability, verifies the rationality of classifying typical days through multi-dimensional features. The effective separation of various typical days in the Figure also lays a data foundation for the subsequent screening of different types of potential users for specific typical days.

4.2.3. Baseline Comparative Validation of Classification Algorithms in the Time Dimension

First, this paper introduces the SC, DBI, and CHI indicators to verify the DBSCAN algorithm. The DBSCAN is baselined against K-Means, agglomerative hierarchical clustering, and GMM, and random interference is eliminated through 30 independent repeated experiments; the results are shown in Table 3.

Table 3. Performance benchmark comparison of typical day classification algorithms.

Clustering Algorithm	SC	DBI	CHI
DBSCAN	0.682 ± 0.013	0.431	854.3
Hierarchical Clustering	0.649 ± 0.035	0.962	346.5
GMM	0.654 ± 0.028	1.032	347.1
K-Means	0.694 ± 0.041	0.599	731.4

Statistical results show that DBSCAN exhibits extremely high robustness in typical day classification, with an SC standard deviation of 0.013, outperforming the random fluctuation performance of other algorithms, which proves the reliability of this method in scenario extraction.

4.3. Second Stage: Dynamic User Profiling and Migration Evolution Under Typical Days

4.3.1. Dual-Optimization Verification of the Optimal Number of Clusters K

(1) Data-driven quantitative optimization

The test interval for the number of clusters is set as $K \in [2, K_{\max}]$ (the upper limit is set to 8 considering the comprehensive computational cost). For each K value, the DTW-K-Means++ algorithm is run independently, and the overall Silhouette Coefficient (SC) and Sum of Squared Errors (SSE) are calculated. By observing the attenuation curve of SSE with the value of K , we look for the “elbow” inflection point with a significant decline (Elbow Method), while simultaneously seeking the position that maximizes the SC indicator. This step aims to screen out the candidate K values that are the most compact within clusters and the most separated between clusters from the mathematical dimension of pure data distribution.

(2) Engineering verification of business logic

The ultimate goal of clustering profiling is to serve the actual refined dispatch of the power grid. If the selected K value is too small, it cannot fully distinguish the essential differences between electric vehicles and industrial loads in terms of response speed and regulation capacity, thereby losing the significance of profiling classification; if the K value is too large, although the mathematical evaluation indicators may be higher, it will lead to extreme fragmentation of profiling tags, significantly increasing the communication and computational burden on the dispatch center for formulating strategies and issuing commands. Therefore, the candidate K values must be further aligned with the actual business echelon requirements of the power grid’s day-ahead/intra-day regulation (such as long-term peak shaving, short-term frequency regulation, base load guarantee, etc.).

According to the dual optimization mechanism, before conducting dynamic profile classification on the 300 typical loads, this section first performs a data-driven quantitative verification of the optimal number of clusters K for the DTW-K-Means++ algorithm. The testing interval for the number of clusters is set as $K \in [2, 8]$. The Sum of Squared Errors (SSE) and the overall Silhouette Coefficient (SC) are calculated under different K values. The results of the dual-metric optimization test are illustrated in Figure 10.

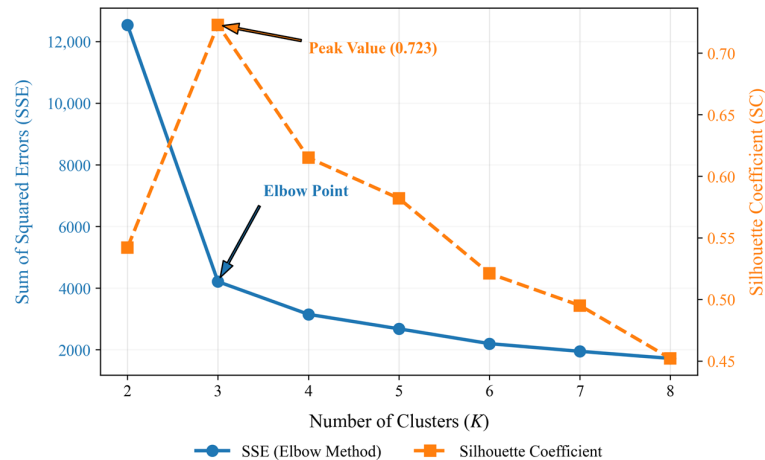


Figure 10. Dual-metric optimization curves of SSE and SC under different numbers of clusters.

According to Figure 10, it can be seen that as the number of clusters k increases, the sum of squared errors (SSE) continuously decreases; however, when $k = 3$, the curve exhibits an obvious elbow feature, and thereafter, the improvement in intra-cluster compactness brought by continuing to increase the number of clusters gradually weakens. At the same time, the SC reaches its maximum value when $k = 3$, indicating that this clustering scheme has the optimal intra-cluster compactness and inter-cluster separation. The SSE reflects intra-cluster compactness, and the SC comprehensively considers intra-cluster cohesion and inter-cluster separation; the joint evaluation of the two can effectively reduce the deviation caused by selecting the number of clusters based on a single indicator.

Therefore, from a statistical perspective, $k = 3$ can achieve a relatively good balance between clustering precision and model complexity, which not only avoids the mixing of different types of users caused by too few clusters, but also avoids the over-subdivision of similar users caused by too many clusters.

Further combining the multidimensional characteristics of users, it can be found that the three types of clustering results respectively exhibit different regulation capability characteristics: one type of users has high physical reliability and continuous regulation capability, and can undertake long-term deep regulation tasks; another type of users has a fast response speed and flexible execution, making them more suitable for undertaking short-term agile response tasks; the third type of users is affected by production process or behavioral constraints, and their adjustable potential is relatively limited. The three types of users have a good degree of separation in the feature space and simultaneously correspond to the typical regulation modes in the DR scenarios of the NPS.

Synthesizing the analysis of the SSE, SC, and the physical interpretability of the clustering results, this paper finally determines the number of clusters as $k = 3$, and the subsequent user dynamic profiling and regulation potential evaluation are all carried out based on this clustering scheme.

4.3.2. Effectiveness Validation of the DTW Time-Series Alignment Mechanism

To intuitively illustrate the methodological advantages of the DTW distance in addressing the time-series phase offset issue of heterogeneous loads in this study, the typical daily load sequences of two residential EV aggregator users (specifically, user IDs #235 and #248) are extracted from the case study for comparative analysis, as illustrated in Figure 5.

In actual physical scenarios, influenced by the superimposed effects of differences in EV owners' commuting habits and time-of-use (TOU) electricity price incentives, even resources of the same category with identical adjustment capacities often exhibit natural misalignments in their load peaks. As illustrated in Figure 11, the EV owner group of user #235 is accustomed to concentrated charging upon returning home at 17:00, whereas the group of user #248, influenced by a willingness to return late or shift peaks, only exhibits a charging peak at 19:00. The morphological profiles of both users' physical capacities and willingness are highly consistent, presenting merely a 2 h phase offset.

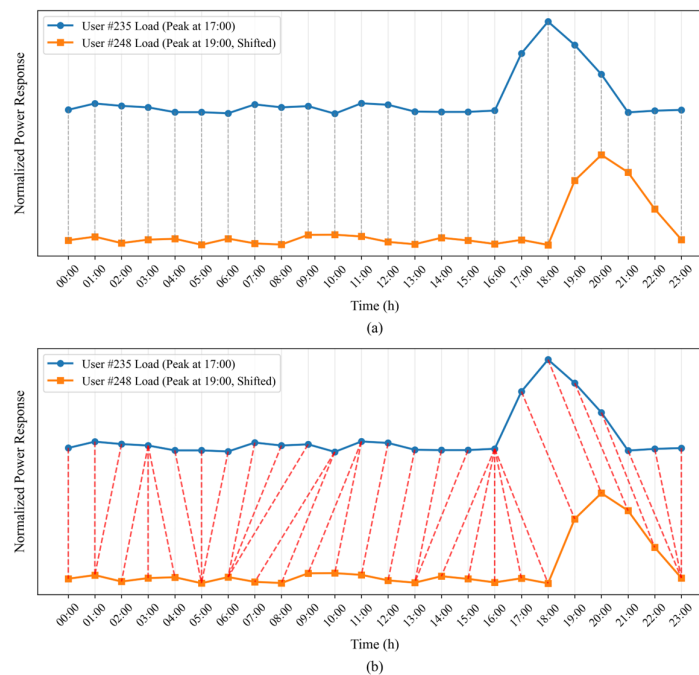


Figure 11. Comparison of alignment effects and numerical evaluation between Euclidean distance and DTW distance (a) Euclidean Distance (Rigid Alignment) | $D_{Euc} = 1.31$; (b) DTW (Optimal Path) | $D_{DTW} = 0.11$.

Under the traditional Euclidean distance metric (Figure 11a), the distance calculation strictly adheres to a rigid point-to-point mapping at identical time instances. This inflexible mapping amplifies the error penalty for the time misalignment of the peaks, resulting in a calculated distance value as high as $D_{Euc} = 1.31$. In traditional K-Means clustering, this easily leads the algorithm to misjudge users #235 and #248—who originally belong to the same category of agile-response resources—as two independent groups with drastically different morphologies. In contrast, after introducing the DTW distance metric (Figure 11b), the algorithm seeks the optimal alignment path through dynamic time warping, successfully identifying the nonlinear distortion on the time axis and morphologically aligning the misaligned charging peaks. The red dashed line intuitively illustrates the optimal time warping alignment path identified by the algorithm. It breaks the rigid constraint of one-to-one correspondence at the same moment in traditional metrics, achieving a non-linear connection and mapping alignment across the time axis between the charging peak feature point of user #235 at 17:00 and the similar feature point of user #248 at 19:00. The warped distance drops precipitously to $D_{DTW} = 0.11$.

This comparison of specific numerical values and alignment effects based on real case study data fully demonstrates that DTW can effectively penetrate the superficial

misalignments triggered by external behavioral randomness, precisely capturing the genuine physical morphologies of multi-type users in terms of load adjustment potential.

4.3.3. Multidimensional User Profiling Results Under Typical Day Constraints

First, the multidimensional feature matrix of 300 typical loads is extracted and calculated; this matrix specifically contains the evaluation indicators of the three dimensions—physical reliability, execution reliability, and behavioral willingness—as well as the user's 24 h normalized response power time-series sequence under the typical day. Selecting the user's multidimensional feature matrix under a specific typical day as input data, the DTW-K-Means++ algorithm is utilized to respectively perform multidimensional dynamic clustering on 300 user load samples under four typical day scenarios: normal production days, extreme peak load days, deep valley rest days, and meteorological anomaly days. Based on the differential characteristics of each cluster center in indicators such as physical reliability, execution reliability, and behavioral willingness, the three clusters are respectively named as deep regulation type, agile response type, and rigid constraint type. The 3D clustering dynamic profiles and tracking distributions of users under each typical day are shown in Figure 12.

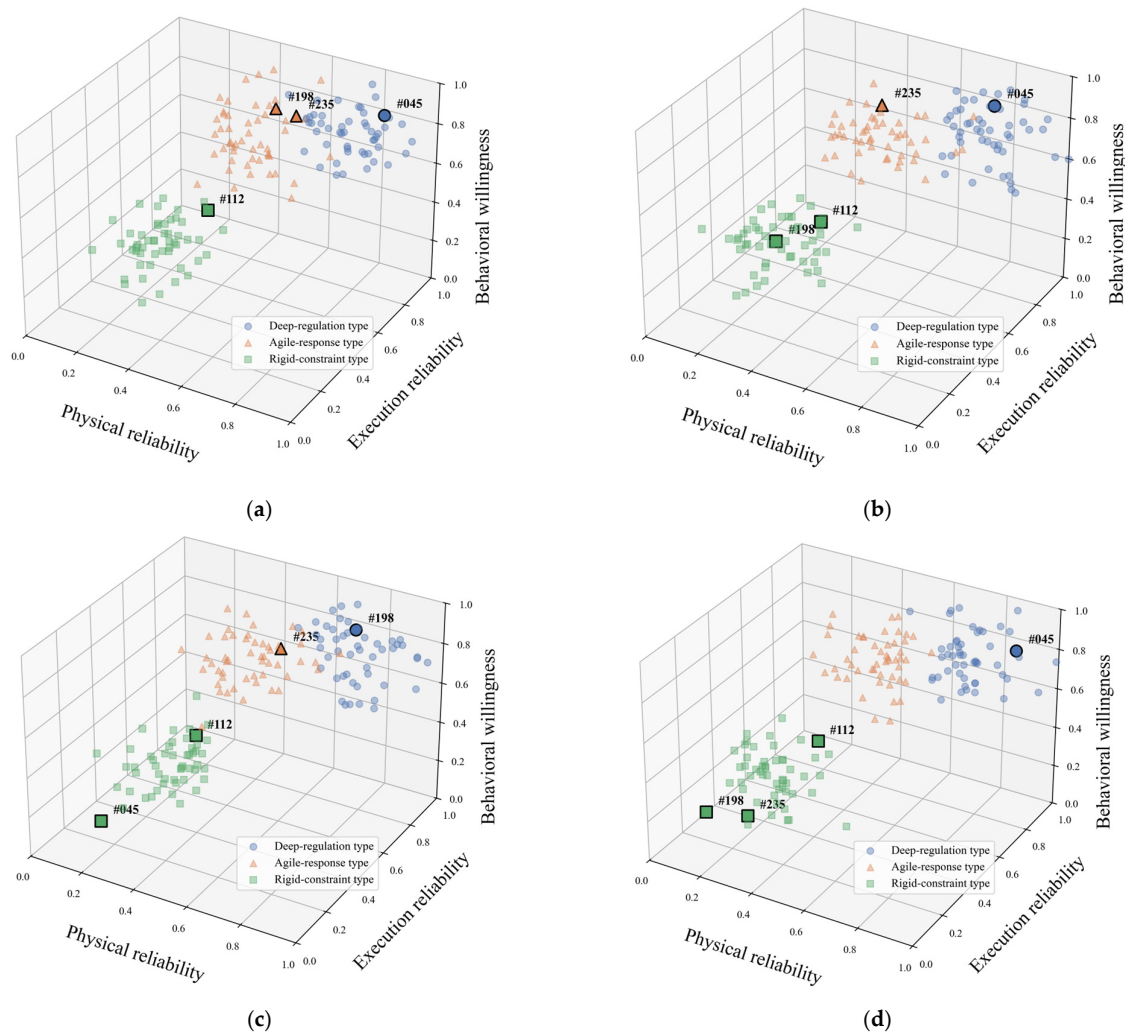


Figure 12. Profiles of user load regulation willingness under different typical day scenarios: (a) normal production days; (b) extreme peak load days; (c) deep valley rest days; (d) meteorological anomaly days.

Synthesizing the clustering results under the four types of typical days, it can be seen that for the 300 typical loads in the park, the results identified and classified by the DTW-K-Means++ algorithm can be summarized on average as follows: the agile-response type accounts for roughly 42.3% of the total, the deep-regulation type accounts for about 31.0%, and the rigid-constraint type accounts for about 26.7%. These three types of profile groups exhibit clear cluster distributions and boundary isolation in the three-dimensional feature space composed of physical reliability, execution reliability, and behavioral willingness. Specifically, the blue circular data points represent deep-regulation loads, mainly distributed in the high-value region of the three-dimensional space. Their physical reliability is concentrated in the 0.70 to 1.00 range, execution reliability in the 0.70 to 1.00 range, and behavioral willingness also in the 0.70 to 1.00 range, demonstrating extremely strong comprehensive regulation capacity and exceptionally high response proactiveness. The orange triangular data points represent agile-response loads, mainly distributed in the upper-middle region of the space. Their physical reliability is limited by equipment capacity and is concentrated in the 0.30 to 0.70 range, but their execution reliability and behavioral willingness are relatively high, distributed in the higher ranges of 0.50 to 0.90 and 0.50 to 0.90, respectively. This reflects that this type of load responds swiftly and proactively, but its absolute regulation margin has an upper limit. The green square data points represent rigid-constraint loads, clustered in the low-value region near the origin of the coordinate axes. Their physical reliability and behavioral willingness are in the extremely low ranges of 0.00 to 0.40 and 0.00 to 0.40, respectively, and their execution reliability generally fluctuates between 0.20 and 0.60. This indicates the low willingness of such loads to participate in the effective regulation of the power grid, as they are constrained by rigid production plans or rigid travel demands.

4.3.4. Dynamic Migration of Profile Categories and Behavioral Feature Analysis

To deeply explore the regulation mechanisms of different types of loads under each typical day, the loads are indexed with numbers. That is, the industrial users include 200 households in total, consisting of discrete manufacturing types (numbered #001–#100) and continuous production types (numbered #101–#180, #281–#300). The EV aggregators include commercial area charging stations (numbered #181–#220) and residential area charging stations (numbered #221–#280), totaling 100 households. A detailed analysis is conducted on the combinations of user categories under the four typical days to illustrate the differentiated regulation potential profiles of various load users across the four types of typical days.

Normal production days correspond to most of the climatically suitable periods throughout the year in a certain city, with smooth power grid operation. At this time, the production shifts in the discrete manufacturing enterprise #045 are regular; its maximum adjustable power on a typical day is large, and its effective continuous regulation duration is abundant. It performs excellently in the physical reliability dimension and maintains stable behavioral willingness, firmly ranking as a deep-regulation type. The continuous production enterprise #112 is bound by strict process safety constraints, and an extremely high risk aversion factor leads to a minimal available physical boundary, classifying it as a rigid-constraint type. The commercial area charging station #198 and the residential area charging station #235 have high typical day participation degrees and excellent regulation stability coefficients, scoring high in behavioral willingness and execution reliability, but they are limited by the upper limit of the individual capacity of their charging/discharging equipment.

Extreme peak load days occur during summer high temperatures or winter extreme cold periods. The contradiction between power grid supply and demand is prominent, and the deterioration of the external spatio-temporal environment directly leads to

significant differentiation in users' multi-dimensional indicators. Constrained by the capacity limits of internal transformers in shopping malls and squeezed by rigid loads such as high-temperature air conditioning, the maximum adjustable power of the commercial area charging station #198 on a typical day is significantly compressed. Its physical reliability plummets, downgrading it from an agile-response type to a rigid-constraint type. In contrast, EV owners at the residential area charging station #235 have their sensitivity to time-of-use electricity prices significantly amplified during this period. The willingness evolution trend shows a proactive response, demonstrating extremely high behavioral willingness and stable execution reliability, keeping it within the agile-response type.

Deep valley rest days occur during long holidays such as the Spring Festival and National Day, when industrial loads experience large-scale production shutdowns. Due to shutdown and maintenance, the total flexible energy of the discrete manufacturing enterprise #045 sharply decreases, and its physical reliability dimension indicators essentially fail, causing it to fall from a deep-regulation type to a rigid-constraint type. Conversely, driven by the holiday peak passenger flow, the charging load of the commercial area charging station #198 surges. Both its maximum adjustable power on a typical day and its effective continuous regulation duration increase significantly. Superimposed with an extremely strong valley-filling willingness, all its indicators leap comprehensively, transforming it into a high-quality deep-regulation resource.

On meteorological anomaly days, under extreme severe weather such as typhoons and rainstorms, the regulation potential of the entire network shrinks in a collapsing manner, which is intuitively reflected as an across-the-board decline in the user behavioral willingness dimension. The travel willingness of the vast majority of EV owners plummets, and the willingness volatility rate sharply increases. For example, the typical day participation degree of the residential area charging station #235 approaches zero, directly downgrading it to a rigid-constraint type. However, because the discrete manufacturing enterprise #045 has an indoor, fully enclosed production line and a self-owned power supply, its execution consistency index is less disturbed by weather, and it still remains a deep-regulation type at this time.

To more intuitively reveal the dynamic migration and drift rules of heterogeneous load groups resulting from spatio-temporal environmental changes on typical days, based on the tracking trajectories of 8 typical users, this paper constructs a three-dimensional time-series trajectory map of the dynamic migration of multi-type users' regulation willingness, as shown in Figure 13.

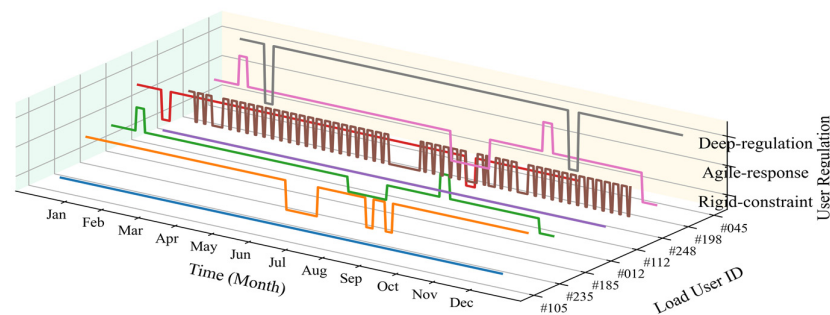


Figure 13. Dynamic migration diagram of load regulation willingness for multi-type users.

From the fluctuating trends of the three-dimensional trajectories in Figure 13, the dynamic evolution process of various load profiles can be clearly observed. The multiple step-like broken lines extending along the time axis in the Figure represent the category evolution paths of different individual loads under various typical days throughout the

year. The height of the lines on the Z-axis intuitively maps their corresponding profile categories: the bottom position represents the rigid-constraint type, the middle position represents the agile-response type, and the top position represents the deep-regulation type.

The trajectory patterns in the Figure exhibit significant differentiation characteristics. The operational trajectory of the continuous production enterprise (e.g., numbered #105) maintains a flat bottom-valley state across the entire time domain, confirming its absolute rigid characteristics due to strict process safety constraints, lacking effective regulation willingness and capability throughout the year. Meanwhile, #112 of the same category shows dense high-frequency periodic oscillations, reflecting the regular switching of willingness of some enterprises under flexible process scheduling. Relatively speaking, the curves of some loads exhibit dramatic potential energy conversions and right-angle jump characteristics. A typical user representing discrete manufacturing enterprises (e.g., numbered #045) maintains a high potential at the top during most periods, such as normal production days, thus exhibiting deep-regulation type characteristics. However, at certain specific nodes (such as deep valley rest days), due to long-holiday shutdowns or production plan adjustments, the trajectory experiences a vertical collapse-style downgrade, dropping to the rigid-constraint type at the bottom. Different from this, the operational trajectory of #012 steadily remains in the deep-regulation type throughout the year, reflecting its highly independent production elasticity.

Furthermore, the trajectory evolution of the EV aggregator exhibits extremely strong environmental sensitivity. The curves of commercial area EVs (such as No. #185 and #198) experience step fluctuations among different typical days, driven by passenger flow during specific periods like deep valley rest days, leaping against the trend to become the deep regulation type at the top, while restrictively falling during extreme peak load days in the middle of the year; whereas the trajectories representing residential area EVs (such as No. #235 and #248) mostly switch between the rigid constraint type at the bottom and the agile response type in the middle. Among them, No. #248 exhibits a significant short-period oscillation characteristic with a 7-day cycle. Out of the 52 weeks throughout the year, after excluding strong external interferences such as long holidays and meteorological anomaly days, approximately over 40 weeks maintain this regular state switching. Specifically, on weekdays, the commuting trajectories of vehicle owners possess high certainty. After vehicles return home at night, their continuous parking duration usually far exceeds the actual time required for a full charge. At this time, vehicle owners have clear expectations for their travel plans the next day, and thus their risk aversion tendency is relatively low. In order to obtain economic revenues from time-of-use electricity prices or DR compensation, the subjective willingness and physical flexibility of vehicle owners to participate in power grid dispatch are synchronously enhanced. Therefore, this load leaps up in the multidimensional feature space and is maintained at the agile response type; conversely, on rest days, residents' travel exhibits extremely strong randomness and fragmentation characteristics. To absolutely guarantee the freedom of travel during holidays and alleviate range anxiety, the decision-making logic of vehicle owners rapidly shifts from pursuing economic revenues to avoiding the risk of power shortage. This psychological reshaping causes their risk aversion factor to surge, and they actively shrink or close the vehicle's response permissions, and its trajectory subsequently drops rapidly to the rigid constraint type at the bottom. This normalized game, alternately dominated by subjective psychological expectations and objective spatio-temporal boundaries, shapes the dense periodic oscillation phenomenon of this load trajectory within over 40 standard weeks. At the same time, when situated in deep valley rest days or meteorological anomaly days, due to routine schedules being disrupted or hindered by extreme weather, its trajectory will undergo an obvious long-period collapse and downgrade. Furthermore,

No. #235 similarly experiences a cliff-like drop and remains at a trough during specific meteorological anomaly periods. This intuitively reflects the strong inhibitory and dynamic reshaping effects of environmental factors such as external meteorological anomalies and holidays on the subjective travel willingness and response capabilities of residential EV users.

The visualization results of the aforementioned three-dimensional time-series trajectory map demonstrate that users' adjustment potential is not a static and immutable inherent attribute; rather, it is a dynamic variable closely correlated with external factors such as meteorological conditions and date types. Under different typical operating scenarios, the states of user profiles exhibit obvious migration characteristics, indicating that employing a single static profile is inadequate to accurately characterize the genuine response boundaries of multi-type demand-side resources.

It is particularly worth noting that Figure 13 extracts only 8 typical users as representatives to achieve a clear visualization of the three-dimensional spatio-temporal trajectories. To compensate for the limitations of localized samples in supporting the dynamic migration conclusion and to verify the universality of this rule across the entire population, quantitative statistics are further compiled on the global dynamic migration characteristics of the full sample across the four categories of typical days.

The global migration statistical indicators reveal that the average number of state transitions per user across the entire population is 1.25. To intuitively demonstrate the replacement frequency of user profile categories, a pie chart depicting the distribution of category replacement frequencies among the population is plotted, as shown in Figure 14.

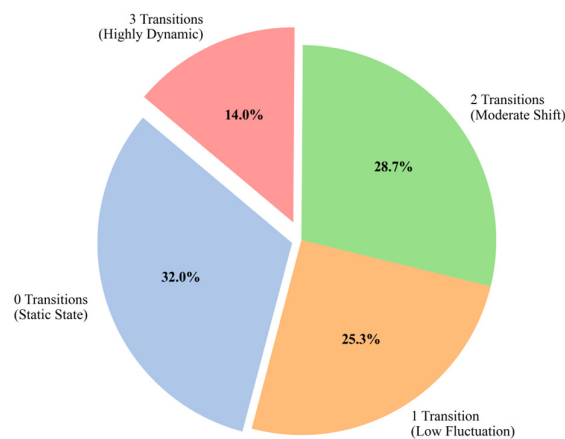


Figure 14. Percentage distribution of profile category replacement frequencies for the full user sample.

As can be observed from Figure 14, only 32.0% of users consistently maintain the same profile category across different typical days, which means that up to 68.0% of users experience at least one category transition. Furthermore, to reveal the essential differences in the driving forces of migration among different heterogeneous resources, an independent statistical separation of migration is conducted for industrial users and EV aggregators: the average number of transitions for industrial users is merely 0.82. Among them, continuous production industrial users, restricted by rigid processes, predominantly belong to the rigid-constraint profile; whereas discrete manufacturing users mostly experience 1 transition, primarily due to state degradation during deep-valley rest days.

EV aggregators exhibit extremely strong environmental sensitivity, with their average number of transitions reaching as high as 2.11. Especially for residential EV users, their migration characteristics highly conform to the inherent cycle and holiday logic. Taking users #235 and #248 from the trajectory map as examples, their state transitions are not random walks, but are strictly controlled by the alternating cycle of regular commuting on workdays and random travel on rest days/holidays.

The proposed method can effectively identify the dynamic evolution laws of user profiles under different typical operating scenarios, thereby validating the effectiveness of this method in evaluating the adjustment potential of multi-type demand-side resources.

4.3.5. Ablation Experiment of the Profiling Algorithm in the User Dimension

To further explore the contribution degree of each core module in the proposed DTW-K-Means++ algorithm, this section designs an ablation experiment. The comparison benchmarks include: a control group retaining K-Means++ but using Euclidean distance, traditional K-Means, K-Medoids, and GMM; the results are shown in Table 4.

Table 4. Ablation study and verification of user profile classification.

Clustering Model Setup	Distance Metric Standard	SC	DBI	CHI
DTW-K-Means++	DTW + Euclidean distance	0.725 ± 0.011	0.413	854.3
Euclidean-K-Means++	Euclidean distance	0.542 ± 0.023	0.854	832.4
Traditional K-Means	Euclidean distance	0.485 ± 0.038	0.981	745.2
K-Medoids	Euclidean distance	0.521 ± 0.038	0.912	789.3
GMM	Probability distribution	0.504 ± 0.026	0.933	760.8

Combining Table 4, the SC indicator of DTW-K-Means++ reaches 0.725 ± 0.011 , which is significantly superior to the benchmark group. The ablation experiment reveals: after removing DTW, the SC indicator plummets to 0.542, indicating that the DTW distance can effectively penetrate the time-series phase shifts in EV and industrial loads, and accurately capture the true regulation potential of multi-type loads.

To verify the scientificity of the performance improvement, this paper conducted a paired t-test on the 30 experimental sequences, and the results show that the improvement of the proposed method has extremely strong statistical significance ($p < 0.01$). Meanwhile, based on the noise injection experiment of Monte Carlo simulation, by injecting 5%, 10%, and 15% Gaussian random noise into the multidimensional scalar feature data of various users, 100 independent Monte Carlo clustering evolutions were performed. The simulation results show that when 5% and 10% data perturbations are injected, the mean values of the algorithm's overall clustering stability indicator (Adjusted Rand Index, ARI) are stabilized at high levels of 0.934 and 0.830, respectively; even when facing the extreme uncertainty impact of 15%, the basic clustering boundaries of the core profile groups in the entire network are still preserved. The above uncertainty quantitative analysis confirms that the clustering classification model established in this paper possesses excellent anti-noise performance and can effectively accommodate the random fluctuation deviations of heterogeneous loads under real physical scenarios.

4.4. Evaluation of Dispatch Benefits and Operational Performance Based on Dynamic Profiles

Although internal indicators such as the SC have been introduced to verify the mathematical compactness of the clusters, under the circumstance of unsupervised learning lacking absolute true labels, it is necessary to introduce actual engineering business scenarios to conduct downstream task validation on the DTW-K-Means++ framework. The

core purpose of constructing the dynamic profiling classification framework in this paper lies in providing reliable ex ante data support for the actual dispatch of the power grid under complex operating scenarios. Therefore, this section designs a day-ahead peak-shaving dispatch scenario targeting a certain extreme peak load day, integrates the dynamic profiling results output by this paper's framework into the dispatch business, and conducts external empirical verification on its potential dispatch benefits and operational performance.

Scenario setting: On a certain extreme peak load day (continuous high temperatures in summer), the dispatch center faces an emergency peak-shaving demand of 15 MW. This paper compares the proposed precise regulation strategy based on dynamic profiling (Strategy A) with the traditional conventional regulation strategy based on maximum physical capacity (Strategy B). Strategy B: The dispatch system solely relies on the equipment capacity on the ledger and blindly issues invitation commands in descending order of capacity until the cumulative target physical capacity reaches 20 MW (reserving a certain margin), without considering the users' subjective willingness and dynamic constraints on that day. Strategy A: Based on the classification results under the extreme peak load day in this paper, the dispatch system directly filters out users currently in the rigid constraint type, and preferentially issues targeted invitation commands to users with the highest comprehensive scores in the agile response type and deep regulation type clusters. The operational performance comparison results are shown in Table 5.

Table 5. Comparison of dispatch benefits and operational performance under different regulation strategies.

Operational Performance Indicator	Strategy B	Strategy A	Benefit Enhancement Evaluation
Target peak-shaving demand (MW)	15	15	\
Total number of actual invited users	76	44	Decreased by 42.1%
Number of rigid-constraint users	25	0	Reduced default risks
Final actual response energy (MWh)	12.5 (Failed to meet target)	16.2 (Robustly met target)	Increased by 29.6%
Effective dispatch response rate (%)	62.50%	94.30%	Increased by 31.8 percentage points
Invalid communication and compensation costs	Relatively high	Extremely low	Significantly reduced operational costs

As can be seen from the evaluation results in Table 5, although the traditional Strategy B covers sufficient capacity on the physical ledger, it blindly invites as many as 25 users who are actually under rigid constraints on that day (such as commercial centers requiring guaranteed air-conditioning for public livelihood due to high temperatures, or continuous production plants that must maintain cooling). This leads to large-scale defaults and insufficient response during the actual execution phase; the effective dispatch response rate is merely 62.5%, ultimately failing to accomplish the system's peak-shaving objective.

In contrast, Strategy A utilizes dynamic profiles to successfully achieve precise dispatching. By thoroughly circumventing high default-risk groups, it robustly overfulfilled the peak-shaving task by issuing only 44 commands. This approach not only substantially elevates the effective dispatch response rate to 94.3%, effectively guaranteeing the physical security of the power grid, but also vastly diminishes the invalid communication and computational overhead of the dispatch system, as well as the erroneous compensation costs in the demand response market. The aforementioned quantitative tests fully demonstrate the immense benefit potential of the profiling method proposed in this paper in practical dispatching operations.

5. Conclusions

This paper proposes a classification method for load regulation potential profiling that considers regulation willingness. It constructs a multidimensional evaluation indicator system covering physical reliability, execution reliability, and behavioral willingness, and designs a two-stage profiling framework with spatio-temporal decoupling: employing the DBSCAN algorithm to adaptively identify typical operating scenarios, and introducing Dynamic Time Warping (DTW) to improve the K-Means++ algorithm under specific scenarios, to eliminate time-series phase shifts among heterogeneous loads. Case study analysis shows that the proposed method can stably classify users into three groups: deep regulation type, agile response type, and rigid constraint type; further quantification reveals that 68.0% of users experience profile category migration across different typical scenarios, indicating that the regulation potential of demand-side resources possesses significant scenario dependency and dynamic change characteristics.

Validation results using the day-ahead peak shaving scenario as an example show that, compared with traditional invitation strategies based on maximum physical capacity, the precise invitation scheme proposed in this paper reduces the actual number of invited users by 42.1%, while the effective response rate increases from 62.5% to 94.3%, significantly improving the precision and effectiveness of demand-side resource dispatch. Although this paper achieves a comprehensive characterization of the load regulation potential of heterogeneous resources, there are still certain limitations: constrained by privacy protection and data acquisition conditions, the semi-synthetic dataset is still difficult to fully represent the complex irrational response behavior of users under extreme disturbances; meanwhile, current offline profiling evaluation methods based on historical data insufficiently consider real-time behavioral feedback under a multi-period dynamic electricity price environment. Future work can further integrate online learning and rolling update mechanisms to construct a closed-loop dynamic profiling model oriented toward bidirectional interaction scenarios.

The underlying load feature data utilized in this study originates from an actual power grid operation monitoring system. Constrained by a strict Non-Disclosure Agreement (NDA) signed with the cooperating power grid enterprise, and because the raw data contain spatio-temporal load information involving sensitive commercial production patterns and residential privacy, the relevant raw datasets and the algorithm source code developed based on this data system cannot be made directly available to the public.

Nevertheless, to support the reproducibility of the methodology in this study to the greatest extent possible, this paper has comprehensively disclosed the core technical details required for model reproduction within the main text, including the data preprocessing mechanism, and the min-max normalization of the indicator system alongside the heterogeneous weight calculation process. Furthermore, all core hyperparameters of the proposed model (such as the neighborhood radius $\epsilon = 0.18$ and the minimum number of points $\text{MinPts} = 6$ for the DBSCAN algorithm, as well as the optimal number of clusters $K = 3$ for the DTW-K-Means++ algorithm) avoid empirical presets; their complete data-driven optimization and verification processes have been thoroughly detailed.

Author Contributions: Author Contributions: Conceptualization, C.W. and H.S.; methodology, C.W.; software, C.W. and X.S.; validation, X.S., S.Y. and Y.Z.; formal analysis, S.Y. and Y.Z.; investigation, X.S. and S.Y.; resources, H.S.; data curation, C.W. and Y.Z.; writing—original draft preparation, C.W.; writing—review and editing, H.S.; visualization, S.Y.; supervision, H.S.; project administration, H.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Science and Technology Project of State Grid Corporation of China (No. 521001250026-079-ZN).

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DBA	DTW Barycenter Averaging
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DR	Demand Response
DTW	Dynamic Time Warping
EV	Electric Vehicle
IDR	Integrated Demand Response
NILM	Non-Intrusive Load Monitoring
SOC	State of Charge
TOU	Time-of-use
V2G	Vehicle-to-Grid

References

1. Chatuanramtharnghaka, B.; Deb, S.; Singh, K.R.; Ustun, T.S.; Kalam, A. Reviewing demand response for energy management with consideration of renewable energy sources and electric vehicles. *World Electr. Veh. J.* **2024**, *15*, 9.
2. Gao, H.; Li, Y.; He, S.; Tang, Z.; Liu, J. Distributionally robust planning for power distribution network considering multi-energy station enabled integrated demand response. *Energy* **2024**, *306*, 132460.
3. Jawad Alrubaie, A.; Salem, M.; Yahya, K.; Mohamed, M.; Kamarol, M. A comprehensive review of electric vehicle charging stations with solar photovoltaic system considering market, technical requirements, network implications, and future challenges. *Sustainability* **2023**, *15*, 8122.
4. Wang, Z.; Liu, Z.; Huo, Y. A distributionally robust optimization approach of multi-park integrated energy systems considering shared energy storage and uncertainty of demand response. *Electr. Power Syst. Res.* **2025**, *238*, 111116.
5. Bian, Y.; Xie, L.; Ma, L.; Cui, C. A novel two-stage energy sharing model for data center cluster considering integrated demand response of multiple loads. *Appl. Energy* **2025**, *384*, 125454.
6. Wu, S.; Pang, A. Two-stage optimal scheduling strategy for community integrated energy system based on uncertainty and integrated demand response model. *Renew. Energy* **2025**, *251*, 123373.
7. Shang, Y.; Shang, W.-L.; Cui, D.; Liu, P.; Chen, H.; Zhang, D.; Zhang, R.; Xu, C.; Liu, Y.; Wang, C.; et al. Spatio-temporal data fusion framework based on large language model for enhanced prediction of electric vehicle charging demand in smart grid management. *Inf. Fusion* **2026**, *126*, 103692.
8. Wei, Y.A.; Fan, J.; Meng, Q.; Debnath, K.B.; Yang, Y.; Liu, J.; Lei, Y. EOLD: A reinforcement learning-based energy-optimised load disaggregation framework for demand-side energy management. *Renew. Energy* **2025**, *252*, 123536.
9. Shahid, A.; Ahmadiahangar, R.; Rosin, A.; Blinov, A.; Korotko, T.; Vinnikov, D. Leveraging the machine learning techniques for demand-side flexibility—A comprehensive review. *Electr. Power Syst. Res.* **2025**, *238*, 111185.
10. Petrucci, A.; Vallianos, C.; Candanedo, J.A.; Buonomano, A.; Athienitis, A. Large-scale load profiling for energy flexibility in residential buildings: A data-driven approach for load aggregation through representative clusters. *Energy Build.* **2025**, *346*, 116195.
11. Fan, Y.; Jiang, T.; Huang, Q.; Ju, P. Profile-based demand response potential assessment for industrial parks. *Autom. Electr. Power Syst.* **2024**, *48*, 41–49.
12. Xu, S.; Yang, Z.; Deng, N.; Wang, B. Residents' willingness to be compensated for power rationing during peak hours based on choice experiment. *Appl. Energy* **2024**, *367*, 123335.
13. Comi, A.; Crisalli, U.; Hriekova, O.; Idone, I. Analysis of the willingness to shift to electric vehicles: Critical factors and perspectives. *Vehicles* **2025**, *7*, 4.
14. Gangadharaiah, R.; Brooks, J.; Boor, L.; Kolodge, K.; Su, H.; Jia, Y. Pooled rideshare in the U.S.: An exploratory study of user preferences. *Vehicles* **2025**, *7*, 2.
15. Dupont, A.; Breal, J.B.; Hayel, Y.; Andrianesis, P. Optimal price menu design of electric vehicle charging stations providing demand response. *Electr. Power Syst. Res.* **2024**, *235*, 110798.

16. Li, C.; Yan, Z.; Yao, Y.; Deng, Y.; Shao, C.; Zhang, Q. Coordinated low-carbon dispatching on source-demand side for integrated electricity-gas system based on integrated demand response exchange. *IEEE Trans. Power Syst.* **2024**, *39*, 1287–1303.
17. Hao, S.; Zu, G. Study on a dual-dimensional compensation mechanism and bi-level optimization approach for real-time electric vehicle demand response in unified build-and-operate communities. *World Electr. Veh. J.* **2025**, *17*, 4.
18. Cai, L.; Yang, C.; Li, J.; Liu, Y.; Yan, J.; Zou, X. Study on frequency-response optimization of electric vehicle participation in energy storage considering the strong uncertainty model. *World Electr. Veh. J.* **2025**, *16*, 35.
19. Zheng, S.; Sun, Y.; Wang, H.; Sun, Y.; Chen, Y.; Mo, X.; Liu, X.; Wang, S. Optimization modeling of integrated demand response considering users' abnormal response behaviors. *Proc. CSEE* **2026**, 1–18. <https://doi.org/10.13334/j.0258-8013.pcsee.260370>.
20. Xie, S.; Zhang, C.; Wang, J.; Shao, Z.; Xiao, S.; Liu, Z. Study on 'guidance-response' game equilibrium of power-traffic coupled system driven by heterogeneous bounded rationality. *Proc. CSEE* **2026**, 1–13. <https://doi.org/10.13334/j.0258-8013.pcsee.260045>.
21. Wang, Y.; Yu, S.; Xue, Y.; Wang, X.; Zheng, Z.; Lin, Z. Flexible resource planning considering multi-scale extreme weather and demand-side resource behaviors. *Autom. Electr. Power Syst.* **2026**, 1–19. Available online: <https://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CAPJ&dbname=CAPJLAST&filename=DLXT20260507001> (accessed on 16 June 2026).
22. Huang, W.; Wang, J.; Wang, J.; Zeng, H.; Zhou, M.; Cao, J. EV charging load profile identification and seasonal difference analysis via charging sessions data of charging stations. *Energy* **2024**, *288*, 129771.
23. García Hinde, Ó.; Gómez Verdejo, V.; Martínez-Ramón, M. Forecast-informed power load profiling: A novel approach. *Eng. Appl. Artif. Intell.* **2020**, *96*, 103948.
24. Dong, X.; Li, X.; Wang, M.; Yu, X.; Mu, Y.; Jia, H. Probabilistic evaluation method for dispatchable potential of electric vehicles based on user charging behavior profiling. *Autom. Electr. Power Syst.* **2026**, 1–14. Available online: <https://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CAPJ&dbname=CAPJLAST&filename=DLXT20260430004> (accessed on 16 June 2026).
25. Yu, Y.; Qian, X.; Chen, X.; Lv, C.; Wang, Y. Dynamic classification and multi-feature online aggregation method of electric vehicles for V2G. *Electr. Power Eng. Technol.* **2025**, *44*, 37–48.
26. Pan, H.; Yang, C.; Wei, X.; Chen, D.; Liang, C.; Yao, M. Research on anomaly detection of user electricity consumption behavior based on DTW algorithm. *Electr. Equip. Ind.* **2025**, *10*, 116–120.
27. Maturo, A.; Vallianos, C.; Buonomano, A.; Athienitis, A.; Delcroix, B. Clustering-driven design and predictive control of hybrid PV-battery storage systems for demand response in energy communities. *Renew. Energy* **2025**, *253*, 123390.
28. Shang, R.; Ma, Y. Electric vehicle charging load forecasting based on K-Means++-GRU-KSVR. *World Electr. Veh. J.* **2024**, *15*, 12.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.