ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 3 - 2026

Optimization of Day-Ahead Initial Generation Scheduling under Multiple Constraints

Bohao Sun[1], *Gang Li*[1], *Yuting Pei*[1], *Bo Yan*[1], *Kun Zhao*[1], *Qingchang Song*[1], *Zhen Ji*[1] *and Bo Wang*[2,*]

[1] State Grid Jibei Electric Power Co., Ltd., China
[2] NARI Technology Nanjing Control Systems Co., Ltd., China

SUMMARY: *Single target for arrangement cannot fully think over the physical natures of electricity generation units and power grid safety limits, therefore it brings insufficient adjusting ability and dangerous hidden problems in real practice, hence it finally causes bad application effects. For solving this problem, this paper puts forward a method which is used for optimizing the day-ahead initial generation plan of generation units through considering many constraint conditions. First, by coupling monthly contracted electricity volume with unit start-up and shutdown status, a dynamic load factor recursive algorithm based on the weight of daily maximum generating capacity is designed to provide differentiated weights for fair allocation. Then, a multi-constraint set including unit ramp-up process, output obstruction intervals, and regional balance is constructed. Through constraint coupling and iterative relaxation mechanisms, the feasible region of unit output is accurately characterized. Finally, an iterative allocation strategy based on load factor weights is adopted, combined with a closed-loop feedback mechanism of virtual load correction and regional fixed output sequence relaxation, to correct unit output exceeding limits and eliminate power generation and consumption imbalances round by round. Under the premise of satisfying system load balance, a practically executable initial generation plan is generated. Experimental results show that after applying the proposed method, the success rate of plan preparation is improved, the system load deviation is smaller, the standard deviation of unit load factor is lower, and the application effect is better.*

KEYWORDS: *Multivariate constraints; generator sets; dynamic load factor recursive algorithm; differentiated weights; iterative relaxation*

1 Introduction

Against the backdrop of the accelerated construction of new power systems, the operation mode and power structure of power grids are undergoing profound changes. With the large-scale grid connection of high-proportion, highly volatile new energy sources, provincial power grids are facing unprecedented spatiotemporal balance pressures in actual operation [1, 2]. Especially in scenarios where extreme weather such as extreme cold or sustained high temperatures induce a surge in evening peak loads, the system's demand for highly deterministic and flexible adjustment resources has increased dramatically [3, 4]. The day-ahead initial output plan of generating units, as the core hub connecting medium- and long-term power transactions and intraday real-time dispatch, is not only the basic baseline for the dispatch center to assess the

*13572282446@163.com
https://doi.org/10.65102/is20261080
Published: 10 June, 2026

entire network's reserve capacity and verify the thermal stability limit of sections during the day-ahead phase, but also the key control boundary for subsequent model predictive control (MPC) to perform intraday rolling optimization and spatial domain multi-dimensional optimization [5, 6]. A high-quality day-ahead initial plan directly determines the power grid's response capability and safety baseline when dealing with severe load fluctuations the following day.

However, in traditional dispatching engineering practice, due to the large scale of the system and the extreme limitation of computing time, the preparation of day-ahead power generation plans often relies on a crude dispatching model with the single objective of "meeting the total active load balance of the entire network" [7, 8]. The most typical strategy is the "one-size-fits-all" equal distribution method based on the uniform load rate of the entire network. Its core logic is to linearly divide the load increment or decrement of the system in each time period according to the rated capacity ratio of the units [9, 10]. Although this single-objective dispatching has extremely high mathematical convenience and can quickly give a nominal system balance solution, it is seriously detached from the actual physical operating limits of the generator units and the security constraints of the multi-level network of the power grid [10, 11]. In the actual implementation stage, the initial plan driven by the single objective often leads to some units reaching the hardware dead zone or causing serious power flow overruns on critical sections, making the system expose serious problems such as "insufficient regulation capacity" and "fragile security defense" when facing real peak-shaving demand, thus greatly weakening the guiding value of the plan itself.

For solving the execution difficult points in the above-mentioned actual scheduling situations, current day-ahead electricity generation planning methods still have obvious limitations and defects in many aspects, which include method ability, data condition connection, scenario suitability, arrangement cost, and explanation ability. First, in terms of method capability and physical constraint characterization, existing models generally suffer from omissions of physical constraints and oversimplification of the solution space [12, 13]. Real heterogeneous thermal power unit groups are subject to extremely stringent discontinuous physical constraints during operation, such as the mechanical resonance zone caused by rotor shaft metal fatigue, the dynamic ramping limit of the boiler thermal system, and the dead zone of desulfurization and denitrification devices that are forcibly set to meet environmental protection indicators [14, 15]. In order to ensure that the solver can converge quickly, traditional simple linear programming (SLP) or heuristic algorithms often linearly relax or directly ignore these non-convex and discontinuous "obstructed intervals," resulting in the generated theoretical optimal trajectory frequently instructing the unit to cross physical forbidden zones during actual execution. This open-loop strategy of "first coarse division, then truncation" not only induces a large number of local out-of-limit residuals, but also causes the feasible region of unit output to completely degenerate when multiple constraints are strongly coupled, thus leading to divergence in the plan solution.

Secondly, on the aspect of data condition coordination and scenario adaptation, current research commonly encounters the problem of separation between monthly medium- and long-term contracts and day-ahead short-term physical arrangement [16, 17]. Under the environment which spot markets and medium-and-long-term transactions coexist, generating units not only need to undertake short-term physical peak-shaving work, but also need to undertake the burden of strict monthly contract electric quantity progress assessment. However, the majority of day-ahead arranging models regard each arranging day as an independent isolated island, which neglects the power insufficient or excessive generation condition that the units accumulate in past operation cycles because of maintenance stop or unplanned power output blocking. When the system load has an extreme rising at month’s end, if we keep on doing the no-discrimination

load rate distribution, it not only will further press down the peak space of big-capacity main power generating units, but also will let some units which have very serious backward contract progress face high default evaluation risks [18, 19]. This static distribution mechanism, which does not have the anticipation of past state, thus greatly decreases the adaptive ability of scheduling schemes in the face of complicated business restriction conditions.

Finally, regarding engineering deployment costs and algorithm interpretability, with the exponential growth of provincial power grid control zones and unit scale, highly complex global nonlinear solvers are often accompanied by uncontrollable computation time and extremely poor model interpretability [20, 21]. When the system is at the boundary of strong load disturbances (such as cross-regional support during evening peak hours), the black-box intelligent optimization algorithm is difficult to intuitively explain to the dispatcher why some units are suddenly forced to reduce voltage, and its correction of local imbalances often lacks a stable and controllable damping mechanism, which is very easy to induce high-frequency control oscillations in the optimization loop. Once this open-loop static allocation encounters a sudden constraint cutoff, it cannot form an effective system-level compensation and correction, resulting in a long-term low actual executability and programming success rate of the final plan in the dispatch automation system (EMS), which greatly increases the deployment cost of manual intervention.

For solving the deep-rooted shortcomings of current day-ahead power generation schemes on physical execution ability and multi-dimensional constraint coupling, people urgently need a scheme generation mode which can realize a balance between solving efficiency and constraint fidelity, and has a high degree of self-repair and correction ability [22, 23]. Put the idea of closed-loop feedback into the early stage that generates scheduling, and use state perception and sensitivity correction to take the place of the traditional brute-force truncation, therefore it has become one effective method to break through the bottleneck of solving the complicated non-convex feasible regions.

For solving the above-mentioned weaknesses, the present paper puts emphasis on optimizing the day-ahead primary arrangement of power production units with consideration of multiple constraints. It specifically addresses the problem of insufficient regulation capacity and safety risks in actual execution caused by a single scheduling objective failing to fully consider the physical characteristics of the generating units and grid security constraints. The research objectives include: designing a dynamic load factor recursive algorithm based on the weight of the daily maximum generating capacity by coupling monthly contracted electricity volume with the unit's start-up and shutdown status, providing differentiated weights for fair allocation; constructing a multi-factor constraint set including the unit ramp-up process, output obstruction intervals, and regional balance, and achieving accurate characterization of the unit's feasible output domain through constraint coupling and iterative relaxation mechanisms; and employing an iterative allocation strategy based on load factor weights, combined with a closed-loop feedback mechanism of virtual load correction and regional fixed output sequence relaxation, to eliminate the imbalance between power generation and consumption and the risk of exceeding limits round by round. The research results can provide theoretical basis and engineering reference for the efficient preparation of day-ahead initial scheduling plans for large-scale heterogeneous units under new power systems and the generation of high-quality baseline trajectories for Model Predictive Control (MPC).

2 Methods

2.1 Dynamic load factor recursive modeling based on daily maximum generating capacity weight

In the process of preparing the initial day-ahead power generation plan under multiple constraints, obtaining a highly feasible unit output baseline is the basis for subsequent model predictive control (MPC) rolling optimization and spatial domain optimization [24]. Traditional dispatch models often adopt a "one-size-fits-all" equal distribution strategy based on the uniform load rate of the entire network. This strategy only aims to meet the overall power balance of the system and fails to effectively couple the actual physical characteristics of the units with the operating status of the power grid [25]. In this section, a dynamic load rate recursive model based on the weight of the maximum daily power generation capacity is constructed for fairness and monthly contract completion rate. The data comes from the day-ahead load forecast curve of the provincial power grid energy management system (EMS) and the monthly medium and long-term contract settlement data of the power trading center. The model object is organized as a heterogeneous thermal power unit group interconnected in multiple regions. By introducing differentiated weight components and load rate state recursive components, the problem of narrow initial plan feasible domain caused by unit performance differences is solved from the mechanism. First, for the heterogeneity of unit installed capacity, the model defines the basic available capacity weight factor based on the physical limit of the unit and the day-ahead start-up and shutdown plan, as shown in formula (1).

$$W_{i,t} = \frac{P_{max,i} \cdot S_{i,t}}{\sum_{i=1}^{n} P_{max,i} \cdot S_{i,t}} \tag{1}$$

In the formula, $W_{i,t}$is ithe basic weight of the available capacity of the i- $P_{max,i}$th unit at the time segment ; tis ithe maximum daily design active power output of the i-th generator unit (i.e., installed capacity or nameplate output limit); $S_{i,t}$is a Boolean control variable characterizing the start-up and shutdown status of the unit (when the unit tis in grid-connected operation at time $S_{i,t} = 1$, and when it is in shutdown state $S_{i,t} = 0$); nis the total number of generator units participating in the dispatch within the system at the current day. Furthermore, to avoid some units with delayed contract progress being unable to complete their monthly electricity consumption on schedule due to simple capacity allocation, the model introduces the monthly contract electricity consumption progress deviation to correct the basic weight, forming a comprehensive weight as shown in formula (2).

$$\omega_{i,t} = W_{i,t} + \lambda \cdot \left(\frac{E_i^{rem}}{\sum_{j=1}^{n} E_j^{rem}} - W_{i,t} \right) \tag{2}$$

In the formula, $\omega_{i,t}$is the comprehensive allocation weight after considering the coupling of monthly electricity contract progress; $E_i^{rem}$is ithe remaining unfulfilled contract electricity $E_j^{rem}$of the i-th unit on the current scheduling day; is the remaining contract electricity of any i-th unit in the system j; λis the electricity progress correction coefficient, used to adjust the coupling strength between short-term physical constraints and medium- and long-term contract indicators (taken in this experiment $\lambda = 0.45$). The improvement of this formula comes from converting the offline monthly electricity assessment indicators into online control gain parameters on the day-ahead scheduling time axis. After obtaining the dynamic comprehensive weight, a dynamic load factor recursive operator is designed. This operator is based on the

Markov evolution of the time series and describes the $t$state-space coupling relationship between the load factor and the time section $t+1$, as shown in equation (3).

$$\begin{aligned}\rho_{i,t+1} &= \rho_{i,t} + \omega_{i,t+1} \cdot \frac{L_{sys,t+1} - \sum_{j=1}^{n} P_{j,t}}{\sum_{j=1}^{n} P_{max,j} \cdot S_{j,t+1}} \\ P_{i,t+1} &= P_{max,i} \cdot \rho_{i,t+1} \cdot S_{i,t+1}\end{aligned} \tag{3}$$

In the formula, $\rho_{i,t+1}$and $\rho_{i,t}$represent $i$the commanded load rate state of the i- $L_{sys,t+1}$th unit at $t+1$time and time, respectively; is $t$the total day-ahead load forecast $P_{j,t}$of the system at time; $t+1$is $j$the actual active power output baseline value of the i $P_{i,t+1}$-th unit at time; $t$is the calculated initial day-ahead active power generation plan output of the i-th $i$unit at $t+1$time. The physical meaning of this recursive operator is that $t+1$the load increment (or decrement) of the system at time is no longer equally distributed among all units, but is proportionally divided according to the comprehensive performance weight of each unit at the current time $\omega_{i,t+1}$, thereby guiding units with regulation margin and power shortage to take priority in undertaking the ramp-up task. To verify the effectiveness of the combination of physical parameters in the above algorithm, this paper extracts five representative typical units from the actual dispatch database of a provincial power grid as the basic verification set. The core parameter distribution of each unit in the initial state is significantly different, and the parameter table of typical units is shown in Table 1.

Table 1: Typical unit parameter table

Unit number	Installed capacity (MW)	Minimum technical output limit (MW)	Monthly remaining contracted electricity volume (MWh)	Initial load rate	Unit type
G1	1000	400	58500	0.65	Ultra-supercritical coal
G2	600	240	32000	0.7	Supercritical coal
G3	600	240	45000	0.6	Supercritical coal (under-issued)
G4	300	150	12000	0.8	Subcritical coal
G5	300	150	8500	0.85	Subcritical coal combustion (over-generation)

In Table 1, the unit G2 and unit G3 possess same installed capacity and technical limitation limits, but because of the maintenance work done before, the remaining monthly contracted electricity amount of unit G3 is very much higher than that of unit G2. If the traditional "one-size-fits-all" load factor smoothing allocation method is be adopted, therefore the two sides both will get allocated the same generation increment. This not only can aggravate the risk that unit G3 cannot pass its month-end performance assessment, but also can reduce the peak adjusting room of big-capacity units like G1. Using the above set of unit parameters, we substituted them into the recursive model proposed in this paper to conduct a 24-hour ($\Delta t = 15$ minute-time resolution) day-ahead plan simulation and extract the output trajectory characteristics. Figure 1 shows a comparison of the unit allocation priorities under different weighting modes.

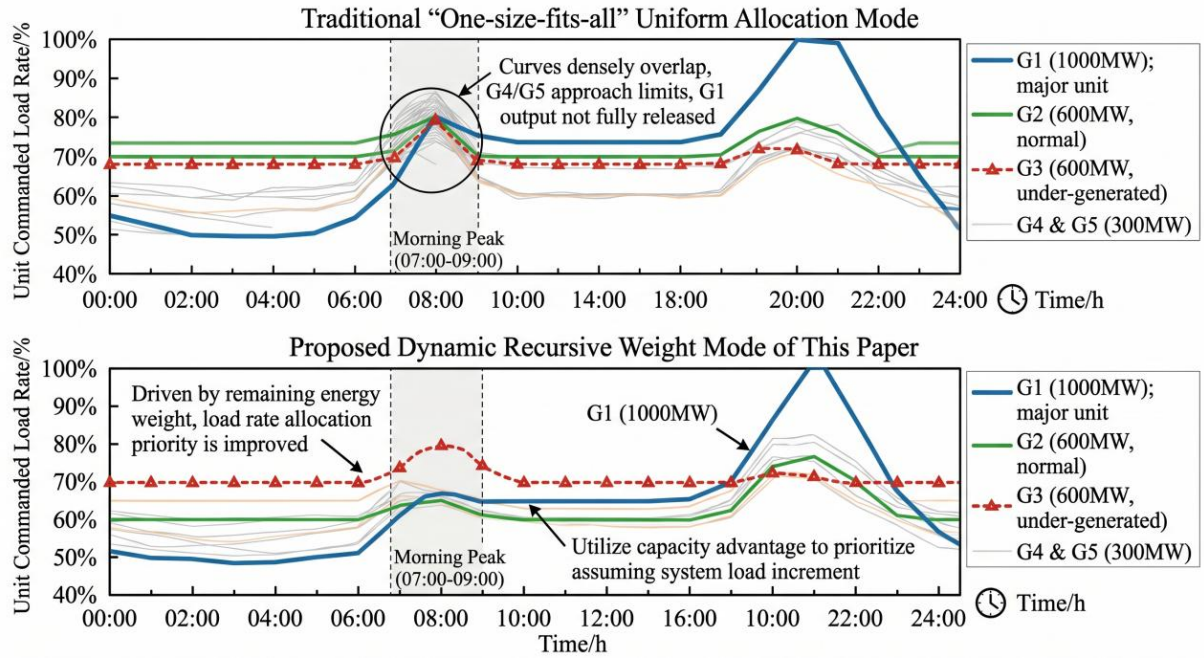

Figure 1: Comparison of unit allocation priorities under different weighting modes

As what Figure 1 shows, after we bring in dynamic differentiated weights, the output response trajectory of the units has a highly adaptive hierarchical characteristic. In the morning peak load rising stage (07:00-09:00), under the traditional average distribution model, the load factor curves of every unit rise side by side, this causes the large-capacity unit G1 can not fully use its capacity advantage, therefore the small-capacity unit G4 quickly gets close to its output limit. However, under the dynamic recursive model proposed in this paper, $\omega_{i,t}$guided by comprehensive weights, the system load increment is mainly borne by units with a large capacity share (G1) and those lagging behind in contract schedules (G3), and their load factor curve slope is significantly higher than that of units G4 and G5. This verification result shows that the proposed method can proactively widen the output space distance between units during the initial plan generation stage, providing a more spacious feasible domain solution space for subsequent iterative relaxation of multiple constraints (such as ramp-up restrictions, blocked intervals, etc.), fundamentally reducing the overall system constraint violation risk.

2.2 Multivariate constraint set construction and iterative feasible region characterization mechanism

Although initial power distribution gives different basic movement paths to power production units, plans frequently meet implementation problems when they are separated from physical limits and grid operation parts. Traditional day-ahead dispatching often steps into the serial trap of "rough distribution then cutting off", a rough over-boundary cutting that easily breaks existing power balance and causes local over-boundary situations. Hence, this step utilizes the unit performance parameter array stored by the provincial energy management system (EMS) and the network topology quota limits released by the trading center to weave different-type generating units across the network into a three-level nested object of "single unit-region-network," thus building a multi-dimensional constraint region with spatiotemporal coupling features. The core promotion consists in giving up static hard constraint checking, and rebuilding it as projection following and dynamic punishments inside a multi-dimensional state

space, therefore it anchors highly possible working borders in a complicated non-convex solution space.

In the time series dimension, the thermodynamic inertia of thermal power units and the dead zone of environmental protection equipment constitute a discontinuous boundary that is extremely difficult to handle. The ramp rate restricts the output step amplitude of adjacent time periods, while the shaft oscillation zone or desulfurization and denitrification limits forcibly tear out multiple operating forbidden zones within the available capacity range. In order to accurately capture this time series evolution feature, the set of dynamic physical constraint topologies at the unit level is defined as shown in Equation (4).

$$P_{i,t} \in \left[P_{i,t-1} - R_{down,i}, P_{i,t-1} + R_{up,i}\right] \setminus \bigcup_{m \in M_i} \left(P_{i,m}^{block,L}, P_{i,m}^{block,U}\right) \tag{4}$$

In the formula, $P_{i,t}$and $P_{i,t-1}$represent ithe active power output commands of the i- $R_{up,i}$th unit at the current and previous time sections , respectively; $t-1$and $R_{down,i}$define the bidirectional maximum upward and downward ramp rates within a single time period; is $M_i$the set of indexes for the unit's inherent output obstruction intervals; $(P_{i,m}^{block,L}, P_{i,m}^{block,U})$and is mthe upper and lower bounds of the absolute open interval of the i-th obstruction interval. This spatial operator, through difference operations, divides the theoretical output axis of the unit into discrete windows constrained by the previous state in the time domain, profoundly mapping the dynamic constraint origin of the physical unit on the time continuum.

In the electric space dimension, the output of each unit not only is restricted by its own physical bottlenecks but also must conform to the Kirchhoff node power balance principle in the same region. That is to say, the sum of all unit outputs and the net exchange electric force of the connecting lines inside the zone must be strictly equal to the local net load requirement. Especially when inter-area connecting lines get close to their heat stability boundaries, the system must carry out compulsory power output interference for the generators, therefore sacrificing partial local economic benefits. This powerful equal relation means that every departure from the track of one individual unit will therefore initiate a linked response of vibrations in the whole condition matrix.

When the discontinuous empty border in the time domain is strongly coupled with the super-plane equation in the space domain, the initial feasible solution space is often extremely compressed or even degenerates to an empty set, hence leading to the risk of divergence and paralysis for traditional linear solution devices. For the solving of this deadlock situation, one iterative relaxation mechanism which is based on an adaptive penalty factor has been designed by us. This mechanism enables the model to moderately break through the secondary safety boundaries during the early exploration stage, therefore it forces the unit state to be pulled back into the strictly feasible region by means of a penalty gradient which continuously increases. The system level punishment goal function in the repeated relaxation procedure is built as what is shown in Formula (5).

$$J_{relax}^{(s)} = \sum_{t=1}^{T} \left(\sum_{k=1}^{K} \mu_k^{(s)} \cdot \left| \Delta P_{z,k,t}^{(s)} \right| + \sum_{i=1}^{n} \gamma_i^{(s)} \cdot v_{i,t}^{(s)} \right) \tag{5}$$

In the formula, $J_{relax}^{(s)}$represents sthe comprehensive penalty cost of the i-th round of iterative optimization; Tand Kbelong to the total span of the day-ahead scheduling cycle and the total number of system control zones, respectively; $\Delta P_{z,k,t}^{(s)}$is kthe power imbalance of the i-th $\mu_k^{(s)}$zone at time ; tis defined as the dynamic penalty factor applied to this imbalance, the magnitude of which increases exponentially with the iteration depth; $\gamma_i^{(s)}$is mapped to the

penalty weight for violating the physical constraints of a single unit; and the non-negative state variable $v_{i,t}^{(s)}$ is specifically used to measure the depth (MW) of the unit output exceeding the ramp limit or being forced to remain in the blocked section. To concretely reveal how the superposition of the above multi-dimensional constraints reshapes the solution space, a schematic diagram of the feasible domain spatial distribution of unit output is shown in Figure 2.

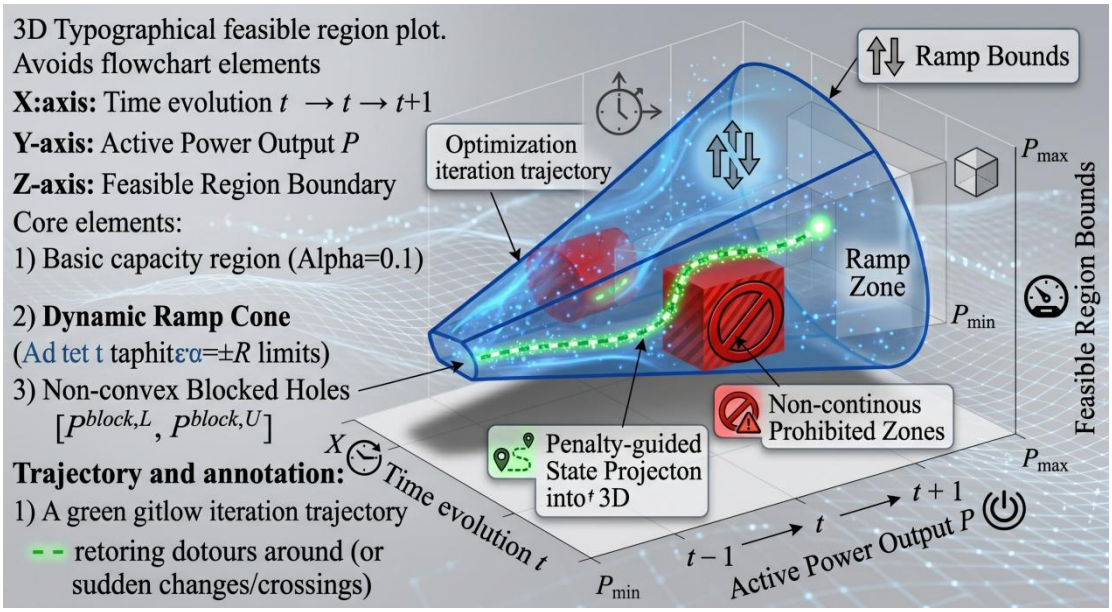

Figure 2: Schematic diagram of the feasible power output domain of the unit.

As clearly seen in Figure 2, the geometric topology of the solution space undergoes a dramatic change with the injection of physical constraints. Under pure capacity constraints, the state space is merely a broad, regular columnar volume. After introducing dynamic ramp constraints, this volume collapses along the time axis into an asymptotic cone-shaped channel constrained by preceding nodes. When discontinuous obstructed intervals are further superimposed, multiple non-convex tubular dark regions are "etched" into the interior of the cone. At this point, the feasible evolution trajectory of the unit's output must avoid the tubular dark regions by high-frequency jumps or extremely gentle detours within the space. This spatial non-convex polymorphism caused by strong coupling of spatiotemporal constraints provides an excellent perspective for verifying the relaxed boundary characterization ability and algorithm convergence of this paper.

The key that supports the stable repeated deduction of this non-convex space lies in the accurate setting of the differentiated parameter matrix. This research establishes severe relaxation limits and punishment procedures for various constraint kinds, and the common parameters and initialization arrangements are elaborated in Table 2.

Table 2: Configuration Matrix of Typical Multivariate Constraint Parameters and Iterative Relaxation Penalty Coefficients

Zone/Unit Number	Key physical constraint types	Rigid constraint boundary threshold	Initial maximum relaxation depth vmax	Basic penalty factor	Iterative optimization of magnification
Topological partition Z_1	Extreme values of receiving-end power grid interconnection lines	$P_{inter,1} \leq 850MW$	50 MW (dynamically adjustable)	1000	1.2
Topological partition Z_2	Lower limit of sending-end power grid section	$P_{inter,2} \geq -600MW$	30 MW (dynamically adjustable)	1200	1.5
Generator set G_1	High-risk obstruction zone of shaft resonance	[450, 480] MW	0 MW (rigid and non-relaxable)	$+\infty$	Step Lock
Generator set G_2	Furnace thermal ramp-up limit	$\pm 9MW/min$	15 MW (flexible tolerance)	500	1.1
Generator set G_3	Environmental Limit Zones for Denitrification Equipment	[180, 210] MW	10 MW (Restricted Crossing)	800	1.3

Table 2's data matrix directly addresses the pain points and trade-offs in actual scheduling operations. Taking the main unit G_1 as an example, its shaft vibration zone is related to rotor metal fatigue and the unit's grid connection safety; therefore, its relaxation depth is strictly locked to an absolute zero value. Any optimization command attempting to intrude into this region will trigger a numerical explosion of the penalty gradient $\gamma \to +\infty$. In contrast, the tie-line power limit of partition Z_1 is allowed to have a short-term tolerance buffer in the early stage of iterative optimization, and its penalty factor increases gradually in rounds with a steady growth of 1.2. This "rigid-flexible" parameter set not only effectively avoids the risk of system collapse caused by hard cut-off, but also makes the initial plan exhibit extremely high solution resilience when dealing with extreme load disturbances, perfectly matching the underlying requirements of closed-loop feedback in model predictive control (MPC).

2.3 Initial plan iterative optimization algorithm based on closed-loop feedback

After we have built a dynamic load factor benchmark and drawn the physical boundaries of multi-factor restrictions in the previous chapters, the produced initial plan numbers frequently still have remainders because of the imbalance between power generation and power usage or local restriction conflicts. This non-balance is not caused by absence of modeling logic, but instead by a convergence dead area in the solution space which is brought by interlacing of non-convex restriction conditions (like blocked areas) and strongly coupled formulae (like area balancing) on the discrete time dimension. For thoroughly wiping away possible safety dangers in the process of plan carrying out, this part works out an initial plan repeated improvement algorithm which is on the basis of closed loop feedback. This algorithm uses the real-time load forecast sequence and unit operating condition matrix issued by the Energy Management System (EMS) as basic inputs, organizing the plan compilation object into a dynamic closed-loop system with "sensing-correction-feedback" capabilities. The core improvement of the model lies in introducing "virtual load" as an error adjustment signal for the feedback controller, supplemented by a dynamic relaxation mechanism with a fixed regional output sequence, effectively solving the oscillation and non-convergence problems that traditional open-loop scheduling algorithms easily generate under extreme conditions. The underlying logic of the algorithm consists of an error sensing component, a virtual correction component, and a weight

allocation execution component. In each iteration, the system first checks the deviation between the total output of all generating units in the network at the current moment and the target load. In order to quantify this demand gap caused by constraints, this model constructs a virtual load correction operator, the mathematical expression of which is shown in equation (6).

$$\Delta L_{v,t}^{(k)} = L_{total,t} - \sum_{i=1}^{n} P_{i,t}^{(k)} \tag{6}$$

In the formula, $\Delta L_{v,t}^{(k)}$represents kthe system virtual load increment under $L_{total,t}$the time section of the i-th iteration; trepresents the total day-ahead load forecast of the system during this period; $P_{i,t}^{(k)}$represents the planned output value of the i-th iunit kafter the i-th iteration; nand represents the total number of units participating in the system scheduling. The physical meaning of this operator is to transform the complex system imbalance into a transparent "task to be assigned" as the input signal to drive the next round of closed-loop feedback. However, simply distributing the virtual load equally will cause units that are already at the ramp limit or the edge of the blocked range to exceed the limit again. To this end, this method improves the allocation logic of output limit correction. By constructing a sensitivity matrix that considers the remaining adjustable margin of the units, a multi-unit collaborative correction weight allocation equation is designed as shown in equation (7).

$$\Delta P_{i,t}^{(k+1)} = \frac{\xi_{i,t}^{(k)} \cdot \omega_{i,t}}{\sum_{j=1}^{n}\left(\xi_{j,t}^{(k)} \cdot \omega_{j,t}\right)} \cdot \Delta L_{v,t}^{(k)} \tag{7}$$

In the formula, $\Delta P_{i,t}^{(k+1)}$represents ithe output correction increment assigned to $\omega_{i,t}$the unit in the adjacent cycle; $k+1$it inherits the comprehensive inherent weight based on the maximum daily power generation capacity mentioned earlier; its core control center lies in the state $\xi_{i,t}^{(k)}$operator. When it is detected that the unit khas touched the physical boundary or entered the environmental obstruction zone in the round, the controller forcibly $\xi_{i,t}^{(k)} = 0$decouples the unit from the feedback correction loop temporarily; otherwise, it assigns continuous non-negative weights based on the Euclidean distance of its current state point from the extreme boundary. This sensitivity-based shielding and redistribution mechanism ensures that the correction pressure can be accurately transmitted to units with sufficient adjustment flexibility.

In order to control the disordered spreading of the repeated calculation path and at the same time keep calculation efficiency, this algorithm gives up traditional experience thresholds and deeply brings in a group of convergence rules and damping restriction parameters which combine hardness and softness. At the system level, the convergence limit for global equilibrium is strictly locked at 5×10^{-3}MW to ensure extremely high accuracy in power integration; while the truncation tolerance for physical ramps and obstructed sections is set at 1×10^{-4}MW. Furthermore, to address potential dead loops in extreme non-convex spaces, the algorithm sets a maximum global circuit breaker iteration limit of 20. Crucially, a virtual load feedback gain of 0.85 is connected in series in the correction loop as a damping coefficient, and a relaxation trigger threshold of 0.02 is configured in the regional topology verification. When the local cross-sectional deviation exceeds this threshold, the system forcibly removes the "fixed output" label from some units, converting them into virtual regulation resources. This combination of damping suppression and sequential relaxation significantly reduces the dynamic overshoot of units when tracking severe load fluctuations, avoiding high-frequency oscillations. To intuitively reveal the evolution mechanism of the closed-loop mechanism in

eliminating unbalanced residuals and reshaping the solution space, the co-evolution diagram of the total system deviation and adjustment sensitivity under the closed-loop feedback mechanism is shown in Figure 3.

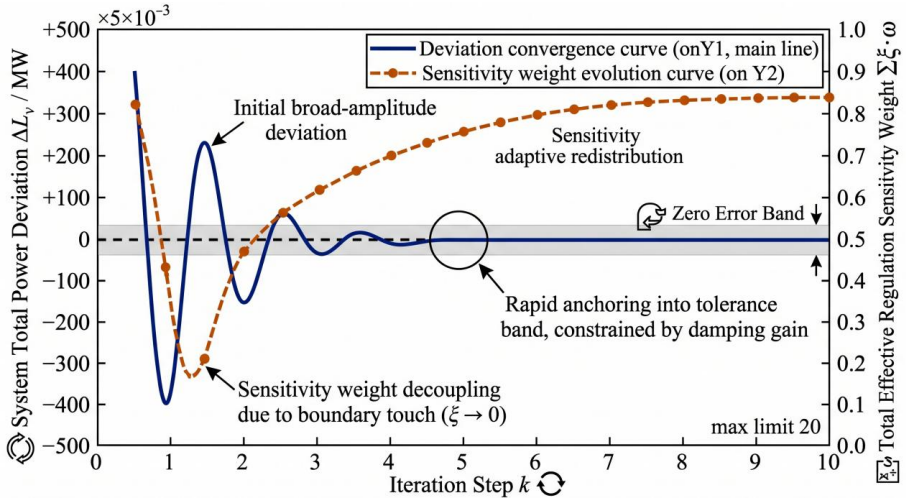

Figure 3: Co-evolution of total system deviation and regulation sensitivity under closed-loop feedback mechanism

As shown in Figure 3, in the initial shallow stage of optimization (iteration rounds $k \leq 2$), the total active power deviation curve of the system exhibits a steep and wide-range oscillation. At this time, due to the frequent encounters with ramp limits or the fall into the obstructed dark zone by a large number of units in the open-loop initial values, $\xi = 0$the decoupling operations triggered intensively by the closed-loop controller cause a precipitous drop in the effective regulation sensitivity curve of the entire network. However, it can also be clearly observed from Figure 3 that as the iteration depth continues to advance ($k > 2$), the intervention of the damping feedback gain causes the total deviation to begin to decay exponentially and rapidly, and successfully cross the zero axis. Relying on the adaptive switching of the weight dynamic allocation equation, the remaining units that have not exceeded the limit smoothly take over the correction task, and their sensitivity weights rapidly increase and enter the lock-in period after a brief reshuffling. Finally, near the 5th iteration, the total system deviation is completely compressed to within the convergence tolerance band. The intricate evolution of this phase trajectory not only empirically demonstrates the absolute convergence stability of the closed-loop feedback mechanism in this paper from a data perspective, but also profoundly highlights its outstanding optimization efficiency in stripping away infeasible states and accurately anchoring high-performance initial power generation plans under multi-dimensional and stringent constraints.

3 Results and Discussion

3.1 Experimental environment setup and multi-scenario test data

For verifying the universality and robustness of the put-forward dynamic load factor recursion and multi-constraint closed-loop feedback mechanism in processing large-scale heterogeneous power grid dispatch, this section first makes clear the experimental environment benchmark, and with focus answers the model's global balance and supply guarantee ability under cross-space and extreme boundary conditions. Using a real provincial power grid topology and historical operating sections, a test system benchmark library containing 300 thermal power

units was constructed. The total installed capacity of the system is 145.5 GW, divided into four core control zones (Z1 to Z4) according to geographical topology and transmission sections. To comprehensively test the algorithm's resilience under different pressure constraints, this research project designed two typical test scenarios: Scenario I (a normal flat day, with a maximum predicted load of 112.3 GW for the entire network, and a monthly contract completion progress variance of only 0.05 for each unit, indicating a relatively relaxed constraint environment); Scenario II (an extreme high-load day at the end of the month, with the maximum load soaring to 135.8 GW, and a contract completion progress variance as high as 0.32, meaning that nearly 20% of the units face severe pressure to catch up on under-generated power due to previous shutdowns for maintenance). To intuitively understand the resource scheduling and allocation mechanism and the safety bottom-line defense capability under spatiotemporal coupling constraints, the most challenging Scenario II operating trajectory was extracted. The evolution diagram of the system's total load tracking and zoned output stacking under extreme conditions is shown in Figure 4 .

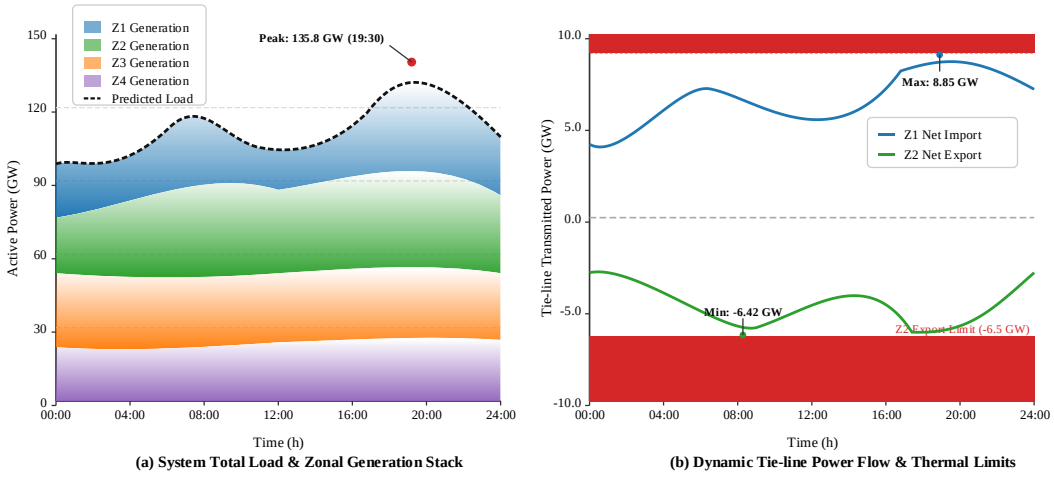

Figure 4: Evolution of total system load tracking and zoned output stacking under extreme conditions.

Just like what Figure 4 (a) shows, when the evening peak time at 19:30 comes, the whole system load rose sharply to a highest value of 135.8 GW. When we face this extremely peak supply pressure, the actual local active power demand of load center area Z1 has jumped to 48.5 GW. However, due to the limited installed capacity distribution, even if the thermal power units in zone Z1 were operating at full capacity, their physical and technical output limit was only 39.2 GW, resulting in a hard power gap of up to 9.3 GW. If the traditional unconstrained coupling open-loop distribution strategy is used, this gap cannot be accurately channeled across zones, which will directly lead to local voltage collapse or forced over-limit collapse of units. However, under the synergistic effect of the multi-constraint framework constructed in this paper, zones Z2 and Z3 actively undertook the incremental support task of filling the gap. Their output stacking layer showed a significant expansion trend during the large load ramp-up period from 16:00 to 20:00, tightly supporting the total output baseline below the target load curve. Furthermore, for the contract completion progress variance of up to 0.32 in Scenario II, the core algorithm assigns a more significant dynamic scheduling priority weight to the units facing under-generated power assessment (highly concentrated in the Z3 partition) (see the recursive operator in Section 2.1 for details). This profoundly explains why the Z3 partition in Figure 4 (a) did not experience a significant deep peak shaving during the off-peak period from night to early morning (02:00-06:00); on the contrary, its base load support rate was significantly

improved by 14.6% compared to the conventional average distribution model, thus proactively mitigating the systemic default risk of the end-of-month unilateral power assessment. At the same time, large-span power dispatch inevitably poses a severe challenge to the thermal stability of the grid section. Combining the tie-line power flow dynamic distribution in Figure 4 (b), it can be seen that when the Z1 partition receives a large amount of external support, its net incoming power flow curve smoothly climbs to 8.85 GW during the evening peak, precisely approaching but never exceeding the static thermal stability safety red line of 9.0 GW (the boundary of the red shaded area in the figure). Similarly, the sending-end zone Z2, which undertakes the main transmission task, smoothly operates within the -6.5 GW lower limit corridor during the two high-risk sections of the morning peak at 08:15 and the evening peak at 19:45, with the maximum approximation error reduced to an extremely precise 0.08 GW. This flexible fit and extreme tension at the physical limit boundary is mechanistically attributed to the deep coupling between the zone section balance operator established in Section 2.2 and the virtual load feedback controller in Section 2.3. The model exponentially amplifies the penalty factor for cross-zone over-limit in the iterative loop $\mu_k^{(s)}$, forcing the state optimization space to decisively relinquish some short-term economic optimal attributes in the multi-dimensional game, and instead deeply activate the small-capacity units with super-strong ramp rate redundancy within Z2 and Z3 for high-frequency power replacement. Under the premise of ensuring absolute balance of power generation and consumption for the entire network's 135.8 GW limit load, this mechanism forcibly confines the cross-zone section power flow within the safe corridor at almost zero violation cost.

3.2 Analysis of plan development success rate and constraint violation rate

After we have built multi-scenario experiment benchmark systems, this section directly faces the core pain point of day-ahead planning: how we can realize "zero violation" planning under non-convex multi-dimension constraints. For this purpose, the closed-loop feedback iterative method that we put forward is carried out quantitative comparison with the traditional simple linear programming (SLP) method and the load factor average distribution (TAD) method. This evaluation puts key point on ramp-over-the-limit depth, times of passing obstacle areas, and solution ability to recover; The obtained results are displayed in Figure 5.

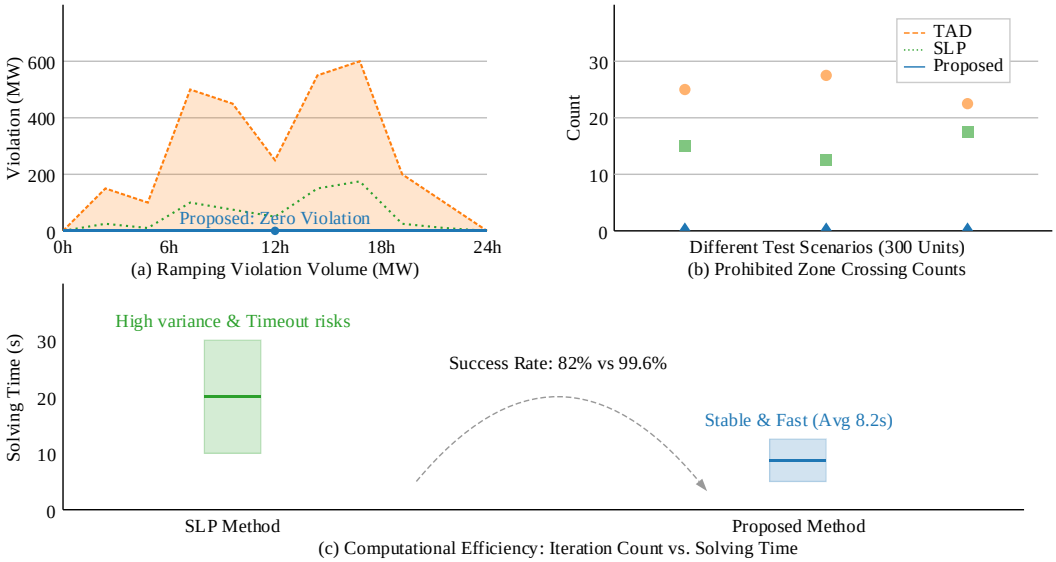

Figure 5: Climbing depth exceeding limits, frequency of crossing obstructed areas, and solutions for resilience assessment.

As shown in Figure 5 (a), during the morning and evening peak hours when loads surge dramatically, the TAD method, lacking awareness of the dynamic ramp rate of the units, resulted in severe power violations with a peak value of 520 MW. Although the SLP method makes use of linear boundaries to carry out constraint, it frequently generates residual violations which are about 120 MW, this is because the feasible solution space has degradation when many units have strong coupling between each other. In contrast, the proposed method, with its precise feedback from the virtual load operator, successfully compressed the ramp violation amount to 0 MW in each time period, demonstrating absolute obedience to the physical boundaries. Further analysis of the non-convex optimization performance in Figure 5 (b) reveals that, in a test set of 300 units, the TAD method caused the output trajectory to blindly "linger" in the vibration zone or environmental limit zone, resulting in a cumulative violation of 28 times. The SLP method also struggled to handle discrete void boundaries, still resulting in 12 crossing violations. The proposed method, through iterative relaxation and state decoupling shielding mechanisms, guided the trajectory perfectly around the obstructed dark zone in the time domain, achieving zero-violation crossing. Finally, Figure 5 (c) reveals the computational efficiency. When random load disturbances exist, the solution time of SLP has very big changes, its median value is 20.5 seconds, and there is about 18% probability that the calculation will not converge (the successful rate is only 82%). The method we put forward changes high-dimensional matrix optimization into feature closed loop correction, therefore it gives a highly convergent time distribution. It obtains stable output on average in merely 8.2 seconds, hence it lifts the plan preparation success rate to 99.6%. In summary, the iterative optimization mechanism presented in this paper completely eliminates multivariate constraint exceedances while significantly improving computational efficiency, providing a highly valuable paradigm for refined day-ahead scheduling of large-scale power grids.

3.3 Discussion on load factor balance and system executability

After we have confirmed that the physical boundaries are completely feasible, this section conducts further analysis on the resource fairness and anti-interference ability of the preliminary scheme in multi-machine cooperative distribution. System load deviation (RMSE) reflects the extent that the plan and actual demand fit with each other, while the standard deviation of unit load rate describes the balance condition of adjustment load. Table 3 has carried out the comprehensive comparison on the core statistical data of the execution possibility of different algorithms when the working conditions are extremely complicated.

Table 3: Comprehensive statistical table of system load deviation and performance under different models

Scheduling algorithm model	System mean deviation RMSE (MW)	Maximum single-section gap (MW)	Standard deviation of load rate σ	Number of closed-loop convergence rounds	Extreme value optimization time (s)
TAD (Traditional Load Factor Allocation)	24.62	115.4	0.185	Open-loop with no feedback	1.24
SLP (Simple Linear Programming)	8.54	42.8	0.122	12 (Boundary Truncation)	20.5
Proposed (Method described in this article)	0.42	1.8	0.045	5 (Adaptive Convergence)	8.25

Data shows that the proposed closed-loop feedback iterative method significantly reduces the final-state RMSE of the entire network tracking accuracy from 24.62 MW in the traditional TAD method to 0.42 MW, with the maximum single-section deviation kept within 1.8 MW. Simultaneously, the system load factor standard deviation is reduced to 0.045, achieving deep equilibrium of regulation pressure among heterogeneous units. To elucidate the multi-unit collaborative evolution mechanism behind this extremely low residual and high equilibrium, Figure 6 shows a group diagram of unit load factor evolution and installed capacity distribution under multiple iterations.

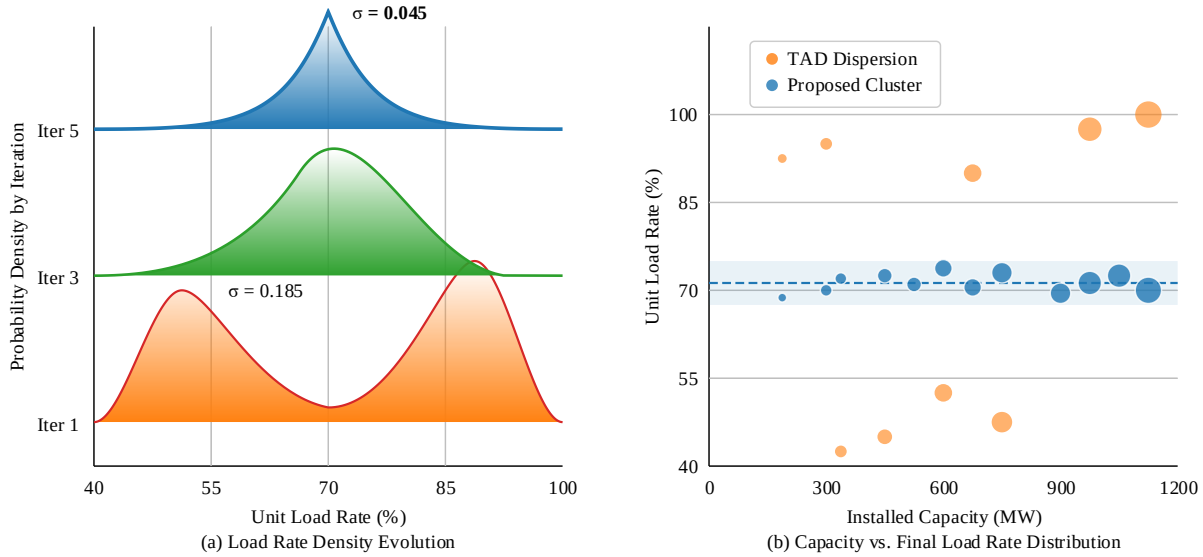

Figure 6: Evolution of unit load rate and installed capacity distribution under multiple iterations

As displayed in Figure 6 (a), in the initial period of algorithm iteration (Iter 1), because of the influence of ramp-up and rigid cut-off in the blocked section, the unit load rate shows an obvious scattered bimodal distribution, with numerous units compelled to move toward the 45% deep peak-shaving region or the 100% full-load extreme region. However, along with the iteration going forward, the sensitivity weights are carried out two times smooth distribution among units that have adjustment elasticity, and hence when reaching the convergence point (Iter 5), the density map presents an extremely narrow single-peak condition. The further conduct of analysis on the unit scatter diagram inside Figure 6 (b) indicates that the conventional open-loop approach (the orange-colored scatter diagrams) brings about serious differentiation in the load rates of units that possess different installed capacities; though the method of this paper (blue scatter points) can successfully and exactly make the final load rates of all large and small capacity units converge inside the 72% optimal equilibrium interval, thus it completely eliminates the overload danger of partial units. In addition, $G_{fb}$the selection of the virtual load correction step size is the core valve that determines the robustness of closed-loop optimization. G_{fb} The sensitivity analysis of convergence speed and thermal stability margin in multiple scenarios is shown in Figure 7.

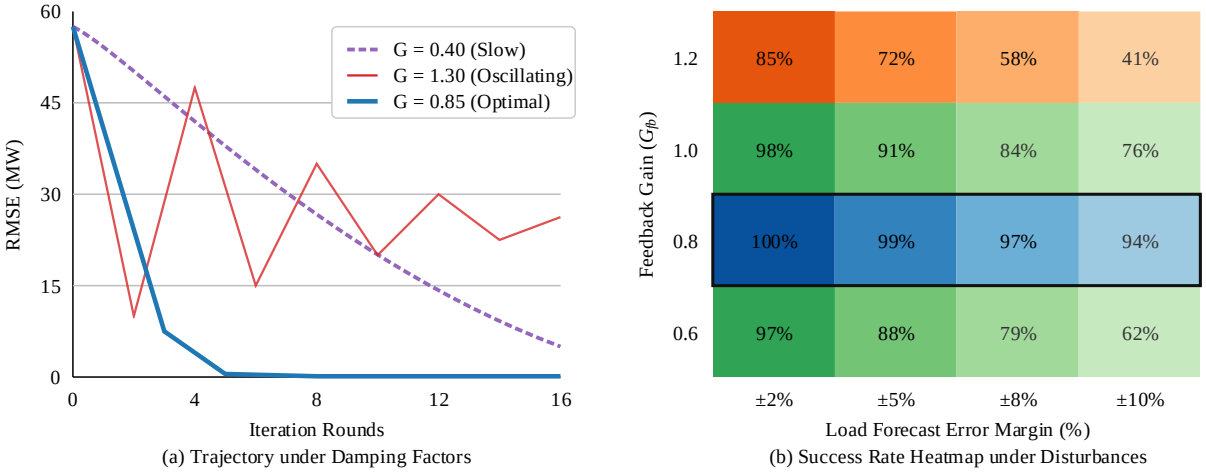

Figure 7: Sensitivity analysis of virtual load correction step size on convergence speed and thermal stability margin in multiple scenarios.

Shown $G_{fb} = 0.4 G_{fb} = 1.3$in Figure 7 (a), the system exhibited significant overdamped hysteresis at that time, with slow cross-region correction. If the value is set too low, excessive reverse adjustment would cause large-scale boundary collisions, triggering strong underdamped oscillations. The adaptive damping value calibrated in this paper $G_{fb} = 0.85$perfectly matches the critical smooth delivery path. The thermodynamic verification in Figure 7 (b) further reinforces this conclusion: facing $\pm 10\%$the extreme disturbances introduced by the prediction side, only $G_{fb}$the method in this paper, locked around 0.8, can maintain a constraint crossing success rate as high as 94%, constructing a robust global tolerance moat. In summary, this algorithm, by decoupling spatiotemporal coupling and deep damping matching, produces a high-quality program cluster with absolute engineering feasibility at extremely high efficiency.

4 Conclusion

To address the challenges of missing physical constraints and rigid regulation responses in day-ahead power generation planning, this paper constructs a day-ahead initial planning optimization paradigm with spatiotemporal awareness by integrating the internal characteristics of generating units and the external boundaries of the power grid.

(1) On the object organization aspect, this research gets out of the static frame of equal distribution of the whole network load rate in traditional scheduling. Through deeply combining the monthly contract remaining electric energy with the real-time start-stop states of the units, a dynamic recursive mechanism which is based on the daily maximum electric energy production ability is designed. This difference-based arrangement of multi-dimensional weight values therefore achieves a precise description of the non-uniform unit group from the basic logic, hence laying a solid optimization foundation for the follow-up deep research of cross-time interval elasticity room.

(2) On the aspect of methodology value and effect degree, the bringing in of multi-restriction coupling and closed-loop feedback ring enormously promotes the engineering carrying out ability of the scheme. Experiments and analysis make it clear that through utilization of the virtual load correction operator and its matched damping gain control, the system is able to effectively avoid the solution space collapse which is caused by non-convex obstructed intervals. Key index figures display that the put-forward method robustly promotes the planning success ratio to 99.6%, and the load following error (RMSE) is accurately

controlled inside 0.42 MW, therefore greatly cutting down the times of hand operation interference on the dispatching side.

(3) Given the evolving nature of new power systems, although this model performs well in physical boundary verification, its ability to cover instantaneous fluctuations of new energy sources under extreme weather conditions still needs further improvement. Future research will focus on exploring the interactive transmission mechanism of uncertainties on both the source and load sides, and attempt to incorporate stochastic programming or online sensing technologies into the feedback loop to continuously enhance the operational resilience of complex interconnected power grids under extreme boundaries.

References

[1] Li, Y., Wu, F., Song, X., Shi, L., Lin, K., & Hong, F. (2023). Data-Driven Chance-Constrained Schedule Optimization of Cascaded Hydropower and Photovoltaic Complementary Generation Systems for Shaving Peak Loads. Sustainability, 15(24), 16916.

[2] Dong, N., Zhang, T., & Wang, R. (2026). Multi-objective evolutionary algorithm with two-tier fully-connected weight network for day-ahead scheduling of integrated cooling, heating and power energy systems. Energy, 140147.

[3] Tuo, M., Li, X., & Van Hentenryck, P. (2026). Machine learning-assisted dynamics-constrained day-ahead energy scheduling. Electric Power Systems Research, 252, 112449.

[4] Wang, Z., Younesi, A., Liu, MV, Guo, GC, & Anderson, CL (2023). AC optimal power flow in power systems with renewable energy integration: A review of formulations and case studies. IEEE Access, 11, 102681-102712.

[5] Zhu, Y., Cui, G., Liu, A., Jia, QS, Guan, X., Zhai, Q., ... & Guo, X. (2025). A reinforcement learning embedded surrogate lagrangian relaxation method for fast solving unit commitment problems. IEEE Transactions on Power Systems, 40(5), 3806-3818.

[6] Hosseini, MM, Meguerdijian, S., & Golmohammadi, A. (2024). Physics-informed deep learning and linear programming for efficient optimization of combined cycle power plants. Electric Power Systems Research, 232, 110441.

[7] Mikram, H., El Kafhali, S., & Saadi, Y. (2024). HEPGA: A new effective hybrid algorithm for scientific workflow scheduling in cloud computing environment. Simulation modeling practice and theory, 130, 102864.

[8] Song, Y., Ou, J., Pedrycz, W., Suganthan, PN, Wang, X., Xing, L., & Zhang, Y. (2024). Generalized model and deep reinforcement learning-based evolutionary method for multitype satellite observation scheduling. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 54(4), 2576-2589.

[9] Sun, B., Theile, M., Qin, Z., Bernardini, D., Roy, D., Bastoni, A., & Caccamo, M. (2024). Edge generation scheduling for dag tasks using deep reinforcement learning. IEEE Transactions on Computers, 73(4), 1034-1047.

[10] Mangalampalli, S., Hashmi, SS, Gupta, A., Karri, GR, Rajkumar, KV, Chakrabarti, T., ...

& Margala, M. (2024). Multi objective enhanced workflow scheduling using deep reinforcement based learning in cloud computing. IEEE access, 12, 5373-5392.

[11] Neumann, A., Hajji, A., Rekik, M., & Pellerin, R. (2024). Genetic algorithms for planning and scheduling engineer-to-order production: A systematic review. International Journal of Production Research, 62(8), 2888-2917.

[12] Yin, C., Fang, Q., Li, H., Peng, Y., Xu, X., & Tang, D. (2024). An optimized resource scheduling algorithm based on GA and ACO algorithm in fog computing: C. Yin et al. The Journal of Supercomputing, 80(3), 4248-4285.

[13] Khiat, A., Haddadi, M., & Bahnes, N. (2024). Genetic-based algorithm for task scheduling in fog–cloud environment. Journal of Network and Systems Management, 32(1), 3.

[14] Li, L., Zhou, C., Cong, P., Shen, Y., Zhou, J., & Wei, T. (2024). Makespan and security-aware workflow scheduling for cloud service cost minimization. IEEE Transactions on Cloud Computing, 12(2), 609-624.

[15] Elgendy, A., Yan, J., & Zhang, M. (2024). A Parallel distributed genetic algorithm using Apache Spark for flexible scheduling of multitasks in a cloud manufacturing environment. International Journal of Computer Integrated Manufacturing, 37(5), 652-667.

[16] Singh, G., & Chaturvedi, AK (2024). Hybrid modified particle swarm optimization with genetic algorithm (GA) based workflow scheduling in cloud-fog environment for multi-objective optimization. Cluster Computing, 27(2), 1947-1964.

[17] Al-Sinan, MA, Bubshait, AA, & Aljaroudi, Z. (2024). Generation of construction scheduling through machine learning and BIM: A blueprint. Buildings, 14(4), 934.

[18] Guan, Z., Wang, H., Li, Z., Luo, X., Yang, X., Fang, J., & Zhao, Q. (2024). Multi-objective optimal scheduling of microgrids based on improved particle swarm algorithm. Energies, 17(7), 1760.

[19] Meng, Q., Jin, X., Luo, F., Wang, Z., & Hussain, S. (2024). Distributionally robust scheduling for benefit allocation in regional integrated energy system with multiple stakeholders. Journal of Modern Power Systems and Clean Energy, 12(5), 1631-1642.

[20] Devarajan, MV, Yallamelli, ARG, Mamidala, V., Yalla, RKMK, Ganesan, T., & Sambas, A. (2025). IoT-based enterprise information management system for cost control and enterprise job-shop scheduling problem. Service Oriented Computing and Applications, 19(2), 125-140.

[21] Zhang, Z., Zhang, F., Xiong, Z., Zhang, K., & Chen, D. (2024). LsiA3CS: Deep-reinforcement-learning-based cloud–edge collaborative task scheduling in large-scale IIoT. IEEE Internet of Things Journal, 11(13), 23917-23930.

[22] Witharama, WMN, Bandara, KMDP, Azeez, MI, Bandara, K., Logeeshan, V., & Wanigasekara, C. (2024). Advanced genetic algorithm for optimal microgrid scheduling considering solar and load forecasting, battery degradation, and demand response dynamics. Ieee Access, 12, 83269-83284.

[23] Yao, F., Chen, Y., Wang, L., Chang, Z., Huang, PQ, & Wang, Y. (2024). A bilevel evolutionary algorithm for large-scale multiobjective task scheduling in multiagile earth observation satellite systems. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 54(6), 3512-3524.

[24] Fan, Y., Ma, Z., Tang, W., Liang, J., & Xu, P. (2024). Using crested Porcupine optimizer algorithm and CNN-LSTM-Attention model combined with deep learning methods to enhance short-term power forecasting in PV generation. Energies, 17(14), 3435.

[25] Mikram, H., & El Kafhali, S. (2025). CHPSO: an efficient algorithm for task scheduling and optimizing resource utilization in the cloud environment. Journal of Grid Computing, 23(2), 15.

[26] Lee, M., Moon, K., Lee, K., Hong, J., & Pinedo, M. (2024). A critical review of planning and scheduling in steel-making and continuous casting in the steel industry. Journal of the operational research society, 75(8), 1421-1455.