

Power Grid Operational Risk Assessment Based on Source-Grid Multi- dimensional Feature Fusion and Multi-gate Mixture-of-Experts Model

Shupeng Zhang

State Grid Jiangsu Electric Power Co., Ltd

Chenguang Qiu

State Grid Jiangsu Electric Power Co., Ltd

Chaojie Ding

State Grid Jiangsu Electric Power Co., Ltd

Jie Li

State Grid Jiangsu Electric Power Co., Ltd

Xing Luo

luoxing3212026@163.com

State Grid Nanjing Power Supply Company

Article

Keywords:

Posted Date: October 27th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11506817/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Power Grid Operational Risk Assessment Based on Source-Grid Multi-dimensional Feature Fusion and Multi-gate Mixture-of-Experts Model

Shupeng Zhang¹, Chenguang Qiu¹, Chaojie Ding¹, Jie Li¹ and Xing Luo^{2,*}

¹State Grid Jiangsu Electric Power Co., Ltd., Nanjing 210000, Jiangsu Province, China

²State Grid Nanjing Power Supply Company, Nanjing 210008, Jiangsu Province, China

*Correspondence: luoxing3212026@163.com

Abstract

With the increasing penetration of renewable energy, the stochasticity of wind and photovoltaic (PV) outputs and complex power-flow constraints impose higher requirements on the spatiotemporal identification of power grid balance operation risks. Existing risk assessment methods often overlook the spatiotemporal coupling and dynamic characteristics of the system and rely heavily on wind/PV data-driven approaches, causing the assessed risk distribution to deviate from actual operation results; neural networks also suffer from poor interpretability. Therefore, this paper proposes a power grid operational risk assessment method based on a multi-gate mixture-of-experts (MMoE) model. First, a two-stage dispatch model driven by empirical forecast error distributions is established to simulate actual dispatch errors and generate risk labels. Second, an MMoE model integrating source-side and schedule-side tabular features, grid-side topology features, and historical temporal features is constructed, and task-specific gating networks are used for joint prediction. Meanwhile, the gating weights and graph attention coefficients of the model provide adaptive interpretability analysis for balancing risks. The determination accuracies for curtailment and security of supply events are 92.7% and 84.2%, respectively, with false-alarm rates of 4.8% and 21.0%. Results for a representative operating interval demonstrate risk-association interpretation based on gating weights and graph attention coefficients, achieving high-accuracy risk prediction and causal tracing.

Keywords: power system risk ; intelligent assessment ; digital-analog hybrid drive ; multi-gated hybrid expert model ; interpretability analysis.

1 Introduction

The pursuit of carbon peaking and carbon neutrality has positioned the construction of a new-type power system, with renewable energy as its main body, as the core direction of the energy transition. The share of renewable energy, particularly wind and photovoltaic power, in the generation mix has been rising steadily, and its output is characterized by significant stochasticity and volatility^[1-2]. As the share of demand-side responsive loads continues to grow, uncertainties on both the source and load sides are further amplified^[3]. Under such conditions, insufficient peak regulation capability may lead to coexisting curtailment and loss-of-load risks, making the imbalance between upward and downward

supply-demand regulation a primary challenge for grid operation [4-5]. In addition, curtailment risk and security of supply risk are negatively correlated in time sequence, and their joint assessment faces the challenge of gradient conflict. Therefore, accurate assessment of system risk levels is essential for the planning and operation of the new-type power system [6-7].

Traditional power system risk assessment methods can be divided into deterministic analysis and probabilistic analysis. Deterministic methods rely on contingency analysis to prevent accidents, but they are computationally intensive and difficult to apply in real time [8]. Probabilistic methods incorporate uncertain factors such as equipment failure probability, weather, personnel, and geological disasters into modeling, and directly quantify risks using real-time data, yielding results that better reflect reality [9-10]. However, with the increasing penetration of renewable energy, risk factors in the new-type power system have become more multi-source and risk manifestations more diverse, rendering traditional risk assessment methods inadequate [11-12]. Existing probabilistic models mostly rely on scenario generation techniques [13-14], which obtain labels by enumerating a large number of uncertain scenarios and solving the unit commitment model individually [15-16]. Their generation process is decoupled from the actual dispatch model, and the generated scenarios often deviate from real operation modes [17-18]. More importantly, existing methods often ignore the spatiotemporal coupling characteristics of physical boundaries and rely excessively on wind and solar data-driven approaches, leading to risk labels that deviate from actual operation results. Moreover, most models are black boxes and lack interpretability [19-20]. In addition, traditional black-box models cannot directly identify risk sources and can only rely on post-hoc explanation methods such as SHAP and GNNExplainer, which are both computationally expensive and unstable [21-22]. The above deficiencies can be summarized into three problems: first, the neglect of spatiotemporal coupling characteristics of physical boundaries; second, excessive dependence on wind and solar data-driven approaches, which causes risk labels to deviate from actual operation; and third, poor black-box interpretability, which cannot support dispatch decisions.

To address these three problems, this paper proposes a power grid operational risk assessment method based on a source-grid multi-dimensional feature fusion multi-gate mixture-of-experts model.

(1) To address the neglect of spatiotemporal coupling characteristics of physical boundaries, a two-stage dispatch model driven by empirical forecast error distribution is established, and a dedicated temporal expert (GRU) is designed to extract the temporal correlation features of renewable output and load, achieving source-grid multi-dimensional feature fusion.

(2) To address the excessive dependence on wind and solar data-driven approaches and the deviation of risk labels from actual operation, an improved MMoE model is proposed, introducing source-load uncertainty-driven independent gating networks and a gradient coordination mechanism to alleviate gradient conflicts and complex coupling constraints, enabling joint assessment of two negatively correlated risks.

(3) To address the poor black-box interpretability, the gating weights and graph attention coefficients are used to form an adaptive interpretability analysis for balancing risks, enabling risk tracing without post-hoc computation.

2 Risk Label Generation Method

2.1 Definition of Power Grid Operational Risk

This paper studies two types of operational risks: renewable energy absorption limitation and system upward-regulation pressure, which are characterized by curtailment power and reserve call amount, respectively. Considering that severe load-shedding events are rare in actual operation, the upward reserve call degree of intraday redispatch relative to the day-ahead plan is used as a proxy indicator of security of supply pressure. This indicator reflects the scale of upward resources required to cope with forecast deviations. The corresponding risk is referred to as security of supply risk.

2.1.1 Curtailment Risk

Curtailment risk characterizes the extent to which renewable output exceeds the system absorption capacity and is determined by the curtailed energy index. The total curtailed energy is defined as the sum of curtailed energy in each period:

$$R_{\text{curt}} = \sum_{t=1}^T \text{Curt}_t \quad (1)$$

where T is the number of periods. Each 15-minute interval is one dispatch interval, yielding 35,040 intervals per year. For single-day dispatch, $T = 96$. Curt_t is the system-wide curtailed energy in period t :

$$\text{Curt}_t = \sum_{b \in B_w} (W_{b,t}^{\text{av}} - Pw_{b,t}) + \sum_{b \in B_v} (V_{b,t}^{\text{av}} - Pv_{b,t}) \quad (2)$$

where B_w and B_v are the sets of buses connected to wind and PV power, respectively; $W_{b,t}^{\text{av}}$ and $V_{b,t}^{\text{av}}$ are the available wind and PV output of bus b in period t ; $Pw_{b,t}$ and $Pv_{b,t}$ are the actual grid-connected wind and PV output of bus b in period t . when $\text{Curt}_t \geq 1000$ MW, the dispatch interval is marked as a curtailment event.

2.1.2 Security of Supply Risk

The severity of security of supply risk is measured by the activated reserve:

$$R_{\text{call}} = \sum_{t=1}^T \text{Call}_t \Delta t \quad (3)$$

where $\Delta t = 0.25\text{h}$, and $Call_t$ is the reserve call amount in period t , calculated as the total increment of intraday redispatch relative to the day-ahead plan:

$$Call_t = \max(0, Res^{\text{req}} - Res_t^{\text{up}}) \quad (4)$$

where Res^{req} is the upward reserve requirement, taken as 20% of the daily maximum generation load; Res_t^{up} is the actual upward reserve capacity, obtained as the sum of additional output that online units can provide:

$$Res_t^{\text{up}} = \sum_{k=1}^{N_{th}} (p_k^{\text{max}} u_{k,t} - P_{k,t}) \quad (5)$$

where N_{th} is the total number of thermal units; p_k^{max} is the maximum output of unit k ; $u_{k,t}$ is the on/off status; $P_{k,t}$ is the output. When $Res_t^{\text{up}} < Res^{\text{req}}$, the dispatch interval is marked as a reserve shortage event.

2.2 Two-Stage Dispatch Simulation Framework

To generate risk labels, this paper adopts a two-stage framework consisting of day-ahead pre-dispatch and intraday redispatch. Taking a natural day as the unit, one dispatch interval is taken every 15 minutes, and the rolling solution is performed 96 times per day. First, the day-ahead pre-dispatch determines the unit startup/shutdown and reserve plan based on wind, PV, and load forecast values. Subsequently, the day-ahead startup/shutdown and reserve plan are fixed, and the actual wind and PV output values are substituted into the intraday redispatch model to obtain the output, power flow, and energy storage state. The two stages are connected in series to form a complete day-ahead-intraday dispatch simulation process.

2.2.1 Day-Ahead Unit Commitment Pre-Dispatch Model

The objective function minimizes the sum of generation cost, startup, shutdown cost, energy storage operation cost, curtailment penalty, and line flow slack penalty:

$$\min \sum_{t=1}^T \sum_{k=1}^{N_{th}} [c_k P_{k,t} + SU \cdot s_{k,t}^{\text{on}}] + c_{st} \sum_{x=1}^{N_{st}} (ch_{x,t} + dis_{x,t}) + C_{\text{cur}} \sum_{t=1}^T Curt_t + C_{\text{bal}} \sum_{l=1}^{N_l} \sum_{t=1}^T \delta_{l,t} \quad (6)$$

where c_k is the generation cost coefficient of unit k ; SU is the single startup cost; $s_{k,t}^{\text{on}}$ is the startup indicator variable of unit k in period t ; c_{st} is the energy storage charging/discharging cost; $ch_{x,t}$ and $dis_{x,t}$ are the charging and discharging power of energy storage x in period t , respectively; N_{st} is the number of energy storage units; C_{cur} is the curtailment penalty coefficient; C_{bal} is the line flow slack penalty coefficient; $\delta_{l,t}$ is the flow soft constraint slack of line l in period t ; N_l is the number of lines.

The model includes standard unit commitment constraints: nodal power balance, unit output limits and ramping, minimum up/down time, energy storage state of charge, renewable output limits, and line flow limits. These follow standard formulations and are not repeated here. The reserve constraints ensure sufficient upward and downward regulation capability:

$$\sum_{k=1}^{N_{th}} (p_k^{\max} u_{k,t} - P_{k,t}) \geq \rho_{up} \cdot \max_t D_t^{gen} \quad (7)$$

$$\sum_{k=1}^{N_{th}} (P_{k,t} - p_k^{\min} u_{k,t}) \geq \rho_{dn} \cdot D_t^{gen} \quad (8)$$

where $\rho_{up}=0.2$, $\rho_{dn}=0.02$, and D_t^{gen} is the generation load. Line flow constraints with a slack variable $\delta_{l,t}$ are:

$$-F_l^{\max} - \delta_{l,t} \leq \sum_{b=1}^{N_b} PTDF_{l,b} \cdot Pinj_{b,t} \leq F_l^{\max} + \delta_{l,t} \quad (9)$$

where $F_l^{\max}$ is the thermal stability limit of line l ; $PTDF_{l,b}$ is the power transfer distribution factor of line l to bus b . The slack variable is penalized in the objective function.

The day-ahead pre-dispatch yields the unit startup/shutdown plan $u_{k,t}$ and the day-ahead reserve plan. This plan ensures load supply while reserving sufficient flexible resources for intraday risks.

2.2.2 Intraday Redispatch Model

Once the day-ahead startup/shutdown plan is determined, the intraday redispatch is executed under actual conditions. This model fixes the day-ahead startup/shutdown plan $u_{k,t}$, replaces forecast values with actual wind, PV, and load data, and re-optimizes unit output, energy storage charging/discharging, curtailed energy, and load shedding. Unlike the day-ahead model, reserve requirements are not enforced, and reserve insufficiency is directly reflected as load shedding or system-level emergency slack. This allows forecast errors to propagate into actual risks.

The objective function extends the day-ahead model by incorporating load shedding and deep peak regulation penalties:

$$\min \sum_{t=1}^T \left[\sum_{k=1}^{N_{th}} c_k P_{k,t} + c_{st} \sum_{x=1}^{N_{st}} (ch_{x,t} + dis_{x,t}) + C_{cur} Curt_t \right. \\ \left. + C_{shed} Shed_t + C_{emg} emg_t + C_{bal} \sum_{l=1}^{N_l} \delta_{l,t} + C_{deep} \sum_{k=1}^{N_{th}} deep_{k,t} \right] \quad (10)$$

where C_{shed} is the load shedding penalty coefficient; C_{emg} is the emergency slack penalty coefficient; C_{deep} is the deep peak regulation penalty coefficient; $deep_{k,t}$ is the deep peak regulation downward adjustment.

The constraints of the intraday redispatch model are largely consistent with those of the day-ahead pre-dispatch model, including nodal power balance, unit output and ramping, minimum up/down time, energy storage, renewable output, and line-flow soft constraints. The main differences are as follows:

- (1) The startup/shutdown status is fixed. The unit startup/shutdown status $u_{k,t}$ is a constant from the day-ahead pre-dispatch solution and is no longer a decision variable.
- (2) Load shedding is allowed. A load-shedding variable $ploss_{b,t}$ is added to the nodal power balance constraint, and the load-shedding amount shall not exceed 5% of the original load:

$$ploss_{b,t} \leq 5\% \cdot P_{b,t}^{L,orig} \quad (11)$$

Where $ploss_{b,t}$ is the original load of bus b in period t .

- (3) Emergency slack is added to the system power balance:

$$\sum_{b=1}^{N_b} Pinj_{b,t} + emg_t = 0 \quad (12)$$

- (4) Deep peak regulation is allowed. The unit output lower limit can be decreased:

$$P_{k,t} \geq p_k^{\min} u_{k,t} - deep_{k,t} \quad (13)$$

$$deep_{k,t} \leq 0.5 \cdot p_k^{\min} u_{k,t} \quad (14)$$

Where $deep_{k,t}$ is the deep peak-regulation amount of unit k in period t , and the maximum downward adjustment is 50% of the minimum technical output.

The above intraday redispatch model ensures that intraday reserve insufficiency, load shedding, and emergency slack are reflected in the dispatch results, thereby providing a physical basis for the subsequent risk labels.

2.3 Historical Error Empirical Distribution

The two-stage dispatch simulation framework establishes a mapping from forecast scenarios to risk labels. However, the framework itself is deterministic. In actual operation, both renewable output and load are subject to forecast errors, and the system may operate under various actual conditions, during which curtailment risk and security of supply risk may arise. To fully characterize the impact of these uncertainties on system

operation, a large number of realistic error scenarios need to be generated based on historical data. Therefore, this paper directly extracts real forecast errors from historical data to construct an empirical distribution.

2.3.1 Error Scenario Set Construction

The renewable output data are divided by day into day-ahead forecast values and intraday actual values. The difference between the actual value and the forecast value is recorded as an error scenario:

$$e_{f,b,t} = A_{b,t} - F_{b,t} \quad (15)$$

where $e_{f,b,t}$ is the forecast error of bus b in period t ; $A_{b,t}$ is the actual value of bus b in period t ; $F_{b,t}$ is the day-ahead forecast value of bus b in period t . Three error scenario sets are constructed synchronously, corresponding to wind power, PV power, and load. Taking the wind power channel as an example, $A_{b,t}$ is the actual grid-connected power of the wind farm connected to the bus, and $F_{b,t}$ is the expected output of the wind farm in the day-ahead forecast.

2.3.2 Scenario Sampling

Since renewable output errors exhibit temporal autocorrelation, daily sampling is adopted. For each day to be assessed, K error scenarios are drawn from the corresponding stratum according to its season. The error field is superimposed on the day to be assessed:

$$W_{b,t}^{(k)} = \max(W_{b,t}^{fc} + ef_{b,d_{s,k}}, 0) \quad (16)$$

where $W_{b,t}^{fc}$ is the day-ahead forecast value; $ef_{b,d_{s,k}}$ is the error field of the source day $d_{s,k}$ corresponding to the k -th scenario. The whole-day error field of the source day is superimposed on the renewable forecast values of the day to be assessed to obtain the assumed actual values under the scenario. For buses that may have negative values after superposition, zero is taken to ensure that the available output is non-negative.

2.3.3 Scenario Aggregation

The intraday redispatch is performed on the K corrected forecast actual outputs to obtain the risk labels under each scenario. For the k -th scenario, the day-ahead startup/shutdown plan is fixed, and the following labels are output: $Curt^{(k)}$, $Shed^{(k)}$, and $Res^{up,(k)}$. Given the curtailment event threshold τ_{curt} and the security of supply event threshold τ_{call} , when the corresponding power reaches the threshold, the scenario is judged to have the corresponding event in period t . Then, the event occurrence proportion and the mean power over the K scenarios are counted to form the scenario event probability label and the expected power label, respectively:

$$P_{curt} = \frac{1}{K} \sum_{k=1}^K \mathbf{1}[Curt^{(k)} \geq 1000] \quad (17)$$

$$P_{\text{call}} = \frac{1}{K} \sum_{k=1}^K \mathbf{1} \left[Res^{\text{up},(k)} < Res^{\text{req}} \right] \quad (18)$$

$$E_{\text{curt}} = \frac{1}{K} \sum_{k=1}^K Curt^{(k)} \quad (19)$$

$$E_{\text{call}} = \frac{1}{K} \sum_{k=1}^K Call^{(k)} \quad (20)$$

$$Call^{(k)} = \max \left(0, Res^{\text{req}} - Res^{\text{up},(k)} \right) \quad (21)$$

where $Call^{(k)}$ is the reserve call amount of the k -th scenario; $\mathbf{1}[\cdot]$ is the indicator function, which takes 1 when the condition holds and 0 otherwise; $P_{\text{curt},t}$ and $P_{\text{call},t}$ denote the empirical occurrence probabilities of the corresponding events under the constructed scenario set. The expected power label is the arithmetic mean of the power over all scenarios. The probability label is the proportion of scenarios satisfying the condition to the total number K , i.e., a frequency estimate; the expected label is the arithmetic mean over the K scenarios.

3 Risk Perception Model Based on Improved MMoE

3.1 Multi-gate Mixture-of-Experts Model

The overall framework of the proposed method is shown in Fig. 1. Based on historical actual operation data, the framework consists of three parts: label acquisition, model training, and model inference.

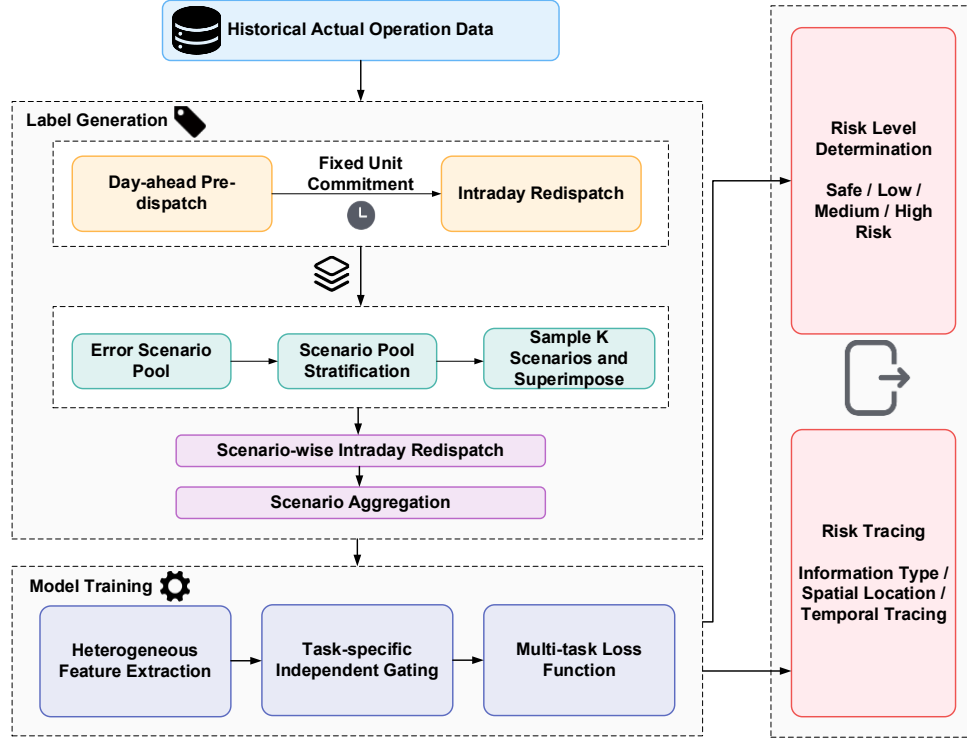


Fig. 1 Overall framework of the proposed risk assessment method

In the label acquisition step, an error scenario pool is first constructed from historical data and stratified. K error scenarios are then sampled, and their errors are superimposed on the day-ahead renewable forecast values to construct actual values. The day-ahead pre-dispatch and intraday redispatch are performed on the K superimposed scenarios, and the dispatch results are aggregated to obtain expected risk labels and probabilities. With the expected risk and probability as supervision, the MMoE model is trained and then used for online inference. The structure of the proposed MMoE model is shown in Fig. 2.

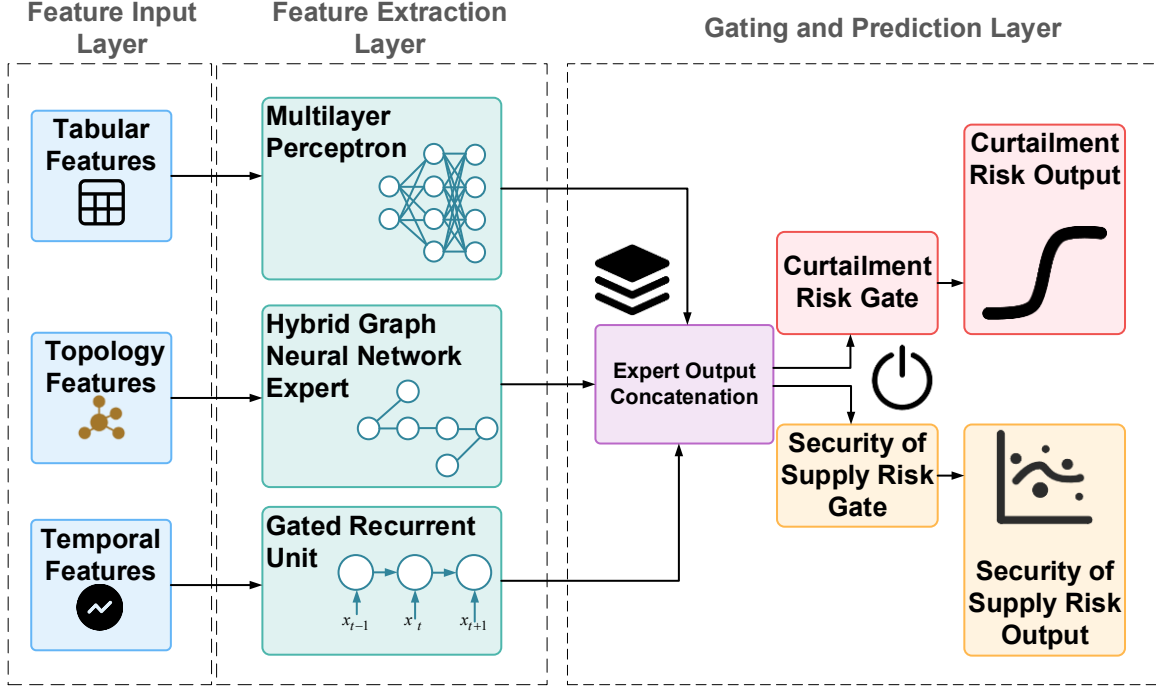


Fig .2 Multi-gate Mixture-of-Experts

3.1.1 MLP Expert

The multilayer perceptron (MLP) consists of an input layer, hidden layers, and an output layer. The layers are usually fully connected, and nonlinear activation functions are used to achieve mapping and fusion of complex features. The MLP expert processes source-side tabular features and schedule-side state features, including wind/PV/load forecast values, reserve capacity, and unit startup/shutdown status. After two fully connected layers, the hidden representation is:

$$\mathbf{h} = \sigma(\mathbf{W}_2 \sigma(\mathbf{W}_1 \mathbf{x} + \mathbf{b}_1) + \mathbf{b}_2) \quad (22)$$

where $\mathbf{x} \in \mathbb{R}^d$ is the input feature vector and $\mathbf{h}$ is the hidden representation.

3.1.2 Hybrid Graph Neural Network Expert

The hybrid graph expert combines GCN and GAT to capture grid-side topology features. The power grid is modeled as a graph $G = (V, E)$, where V is the set of bus nodes and E is the set of lines. The GCN layer updates node representations as:

$$\mathbf{H}^{(l+1)} = \sigma \left(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{H}^{(l)} \mathbf{W}^{(l)} \right) \quad (23)$$

where $\hat{A} = A_c + I$, $\hat{D}$ is the degree matrix, $H^{(l)}$ is the node representation of the l -th layer, $W^{(l)}$ is the learnable weight, and $\sigma(\cdot)$ is the activation function. The GAT layer further applies attention-based aggregation:

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(\mathbf{a}^T [Wh_i \parallel Wh_j]))}{\sum_{k \in N_i} \exp(\text{LeakyReLU}(\mathbf{a}^T [Wh_i \parallel Wh_k]))} \quad (24)$$

$$h_{i'} = \sigma \left(\sum_{j \in N_i} \alpha_{ij} Wh_j \right) \quad (25)$$

Where α_{ij} is the attention coefficient, N_i is the neighbor set of node i , $\mathbf{a}$ is the attention weight vector, W is the feature transformation matrix, and $\parallel$ denotes concatenation. This expert extracts line topology and network structure information, providing spatial support for security of supply risk prediction.

3.1.3 GRU Expert

The gated recurrent unit (GRU) is a typical recurrent neural network variant specifically designed for temporal data. In this module, the GRU expert is responsible for processing the temporal features of the source side over the previous 4 hours and capturing the dynamic variation trends of renewable output and load. The input sequence is $S = [s_1, s_2, \dots, s_T]$, where $T = 16$, and the feature dimension of each time step is 8, including wind power, PV power, load, and their differences and volatility.

3.1.4 Multi-gate Control

The outputs of the three experts are denoted as h_{MLP} , h_{GCN} , and h_{GRU} , and are concatenated before being fed into the gating network. Since curtailment risk and security of supply risk are physically negatively correlated, a shared gate would cause gradient conflict. To address this, an independent gate is assigned to each task. For task $m \in \{\text{curt}, \text{call}\}$, the expert weights are:

$$g_m = \text{Softmax}(W_{g,m} \cdot x + b_{g,m}) \quad (26)$$

where $W_{g,m}$ and $b_{g,m}$ are the gating parameters of task m . The system takes the Top-2 experts with the highest weights, normalizes them, and performs weighted fusion:

$$y_m = \sum_{i \in \text{Top02}} \frac{g_{m,i}}{\sum_{j \in \text{Top02}} g_{m,j}} \cdot E_i(x) \quad (27)$$

Each task selects the two experts with the highest weights and renormalizes their weights to achieve sparse task-specific fusion.

3.1.5 Model Output Layer

For each dispatch interval, four prediction quantities are output: expected curtailed energy E_{curt} , curtailment probability P_{curt} , expected activated reserve E_{call} , and reserve activation probability P_{call} . The numerical prediction branch employs the Mean Squared Error (MSE) loss, and the probability prediction branch employs the Binary Cross-Entropy (BCE) loss. The total loss is

$$L = L_{MSE} + \lambda L_{BCE} \quad (28)$$

where λ balances the two losses.

3.2 Risk Level Determination and Tracing

3.2.1 Risk Level Determination

To convert continuous power prediction results into event determination results, power thresholds are set for curtailment and reserve call, respectively. For task q , when the predicted expected power $\hat{\mu}_{q,t}$ reaches the corresponding threshold τ_q , the dispatch interval is determined to trigger the corresponding event, i.e., $\hat{z}_{q,t} = \mathbf{1}(\hat{\mu}_{q,t} \geq \tau_q)$. The determination results of the two tasks are combined into four joint states: no event, curtailment only, security of supply pressure only, and dual event.

3.2.2 Risk Tracing

The risk perception model needs not only to output risk levels but also to provide risk tracing information, which is crucial for the interpretability of actual dispatch. Unlike post-hoc explanation methods, the MMoE architecture naturally provides tracing information during forward propagation. Tracing information is extracted from three levels.

(1) Gating weights. The gating network outputs a three-dimensional weight vector for each task:

$$\mathbf{g}_m = [\mathbf{g}_{m,1}, \mathbf{g}_{m,2}, \mathbf{g}_{m,3}] \quad (29)$$

where $\mathbf{g}_{m,1}$, $\mathbf{g}_{m,2}$, and $\mathbf{g}_{m,3}$ correspond to the weights of the MLP expert, the GCN and GAT expert, and the GRU expert, respectively, and the sum of the three is 1, $\sum_{i=1}^3 \mathbf{g}_{m,i} = 1$.

The expert with the highest weight indicates the dominant information type driving the risk.

(2) Graph expert bus tracing. For line l , the edge attention weight is the average of the attention coefficients at both ends:

$$\beta_i = \frac{1}{2}(\alpha_{ij} + \alpha_{ji}) \quad (30)$$

where α_{ij} is the attention coefficient of node i to node j . The top K lines by β_i are selected as critical lines and are the main spatial sources of risk.

(3) Renewable fluctuation tracing. For time step t , the importance weight is the proportion of the hidden state norm:

$$\gamma_t = \frac{\|h_t\|_2}{\sum_{\tau=1}^T \|h_\tau\|_2} \quad (31)$$

where h_t is the GRU hidden state. The top K periods by γ_t are selected as critical periods and are the main temporal sources of risk. The three levels together form a structured tracing report, which is generated synchronously with the prediction without additional computation.

3.3 Model Training Method

3.3.1 Loss Function Design

The numerical prediction branches output a predicted mean μ and a log variance $\log \sigma^2$, using the Gaussian negative log-likelihood:

$$L_{\text{num}} = \frac{1}{N} \sum_{i=1}^N w_i \left[\frac{(y_i - \mu_i)^2}{2\sigma_i^2} + \frac{1}{2} \log \sigma_i^2 \right] \quad (32)$$

Where y_i is the true value, μ_i is the predicted mean, σ_i^2 is the predicted variance, and w_i is the sample weight. The probability prediction branches use the binary cross-entropy loss:

$$L_{\text{prob}} = -\frac{1}{N} \sum_{i=1}^N w_i [y_i \log \hat{p}_i + (1 - y_i) \log(1 - \hat{p}_i)] \quad (33)$$

To prevent the gating network from collapsing to a few experts, a load balancing loss is introduced:

$$L_{\text{bal}} = E \cdot \sum_{i=1}^E f_i \cdot g_i \quad (34)$$

where E is the number of experts, f_i is the selection frequency of expert i , and g_i is the average weight assigned to expert i . The total loss is:

$$L = L_{\text{curt}} + L_{\text{call}} + L_{\text{cev}} + L_{\text{rsv}} + \lambda_{\text{bal}} L_{\text{bal}} \quad (35)$$

where L_{curt} and L_{call} are the Gaussian NLL losses of the curtailment and security of supply branches, L_{cev} and L_{rsv} are the BCE losses of the probability branches, and λ_{bal} is the weight of the load balancing loss.

3.3.2 Output Probability Calibration

This paper fits a monotonically increasing function $f:[0,1] \rightarrow [0,1]$ to map the original probability to the calibrated probability, so that the calibrated probability is consistent with the actual occurrence frequency. First, on the validation set, an isotonic regression model is fitted with the original probability as input and the actual event occurrence as the label. Isotonic regression is solved by minimizing the squared error subject to a monotonicity constraint:

$$\min_f \sum_{i=1}^{N_{\text{val}}} (f(\hat{p}_i) - y_i)^2 \quad (36)$$

where $\hat{p}_i$ is the original probability on the validation set; y_i is the true label; N_{val} is the number of validation samples. During inference, the original probability is input into the calibration function to obtain the calibrated probability:

$$\hat{p}^{\text{cal}} = f(\hat{p}) \quad (37)$$

Through this method, the probability output by the model is calibrated to a credible level consistent with the actual occurrence frequency.

4 Case Study

4.1 Experimental Setup

4.1.1 Dataset Partition

This paper is based on full-year actual operation data from a provincial power grid, comprising 35,040 dispatch intervals. Each dispatch interval contains the load, wind power output, PV power output, unit startup/shutdown status, line flow, and other operational information at that moment. Risk labels are generated by the two-stage dispatch simulation method described in Section 1. The day-ahead pre-dispatch determines the unit startup/shutdown and reserve plan based on forecast values, and the intraday redispatch fixes the day-ahead startup/shutdown plan and substitutes actual values to obtain the curtailed energy, reserve call amount, and corresponding event labels for each dispatch interval.

The dataset is partitioned using a stratification-then-splitting strategy. First, the full-year dispatch intervals are stratified by season and peak-valley type. Then, within each stratum, the data are split into training, validation, and test sets in chronological order at the day

level in a ratio of 70%, 15%, and 15%. The resulting training set contains 24,534 dispatch intervals, and the validation and test sets each contain 5,253 dispatch intervals. To obtain a more discriminative evaluation, this paper takes five times the Monte Carlo standard error and defines the curtailment event in the experiments as curtailed energy $\geq 5\text{GW}$ and the security of supply event as reserve call amount $\geq 3\text{GW}$. The same thresholds are applied consistently to all models and all tables.

4.1.2 Evaluation Metrics

According to the determination rule, this paper performs binary classification evaluation for curtailment events and security of supply pressure events separately, and uses hit rate, determination accuracy, and false-alarm rate to measure the identification performance. For the k -th risk class, the hit rate is defined as the proportion of true dispatch intervals of that class that are correctly identified:

$$HR_q = \frac{TP_q}{TP_q + FN_q} \quad (38)$$

Where TP_q is the number of dispatch intervals with actual events and predicted events; FN_q is the number of dispatch intervals with actual events but predicted no events; FP_q is the number of dispatch intervals with no actual events but predicted events; TN_q is the number of dispatch intervals with no actual events and predicted no events. The determination accuracy is:

$$Acc_q = \frac{TP_q + TN_q}{TP_q + FN_q + FP_q + TN_q} \quad (39)$$

The hit rate of each risk class is calculated independently to reflect the model's ability to identify different risk types. The false-alarm rate is defined as $FP_q / (FP_q + TN_q)$.

4.2 Model Settings

To comprehensively evaluate the performance of the proposed method, the following models are selected for comparison:

1. GBDT: Gradient Boosting Decision Tree, serving as a baseline model for tabular data.
2. MGAT: Multi-head Graph Attention Network, serving as a baseline model for graph models.

The training hyperparameters are uniformly set as follows:

Table 1 Hyperparameters

Parameter	Value
Optimizer	Adam

Learning rate	1×10^{-3}
Weight decay	1×10^{-5}
Batch size	1024
Maximum epochs	60
Early stopping patience	10
Gradient clipping threshold	5.0
Number of experts	3
Gating hidden layer dimension	64
Graph convolution hidden layer dimension	32
GRU hidden layer dimension	32
Temporal window length	16

4.3 Prediction Results and Risk Determination

4.3.1 Prediction Results

To intuitively demonstrate the prediction performance of each model under different risk conditions, Fig. 3 presents a comparison of prediction results on a high-curtailment day and a security of supply peak day.

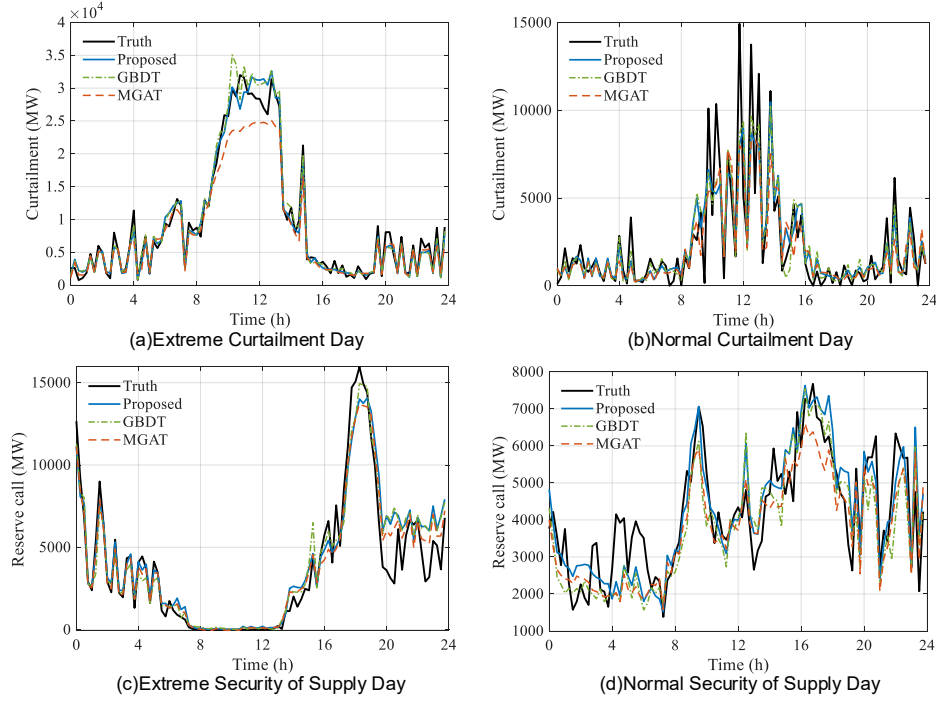


Fig.3 Comparison of prediction results of each model

Fig. 3(a) and (b) show the curtailed energy prediction results under extreme and normal curtailment day scenarios, respectively, while Fig. 3(c) and (d) show the activated reserve prediction results under extreme and normal security of supply day scenarios, respectively.

In the curtailment scenarios (a) and (b), all three models track the overall temporal variation of curtailed energy reasonably well. The prediction curves of the proposed method and GBDT are relatively close, and both maintain good consistency with the true values during the midday peak period on the extreme curtailment day (a). In contrast, MGAT exhibits significant underestimation near the extreme curtailment peak, with pronounced amplitude shrinkage, which is consistent with its relatively lower score in the curtailment task.

In the security of supply scenarios (c) and (d), the predicted values of all models during peak periods are generally lower than the true values, exhibiting systematic amplitude shrinkage. At the evening peak on the extreme security of supply day (c), the proposed method achieves the best tracking performance, while both GBDT and MGAT show varying degrees of underestimation. On the normal security of supply day (d), the prediction curve of GBDT fluctuates considerably, with abnormal spikes in some valley and transition periods, failing to accurately capture the dynamic changes of the true values. MGAT shows significant underestimation during multiple peak periods. In contrast, the prediction curve of the proposed method is smoother overall and outperforms the other two methods in tracking both peaks and valleys.

4.3.2 Probability Calibration

The probability outputs of the model are calibrated by isotonic regression on the validation set. To verify the effectiveness of the calibration, Fig. 4 presents the reliability diagrams for both curtailment risk and security of supply risk before and after calibration.

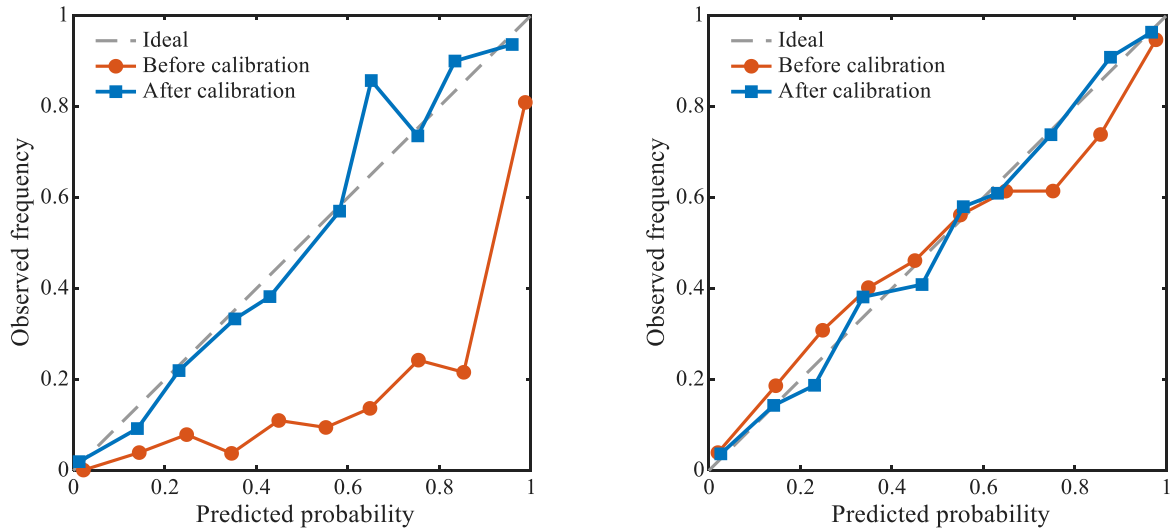


Fig.4 Reliability diagrams of probability calibration for the two risk types

In Fig. 4, the horizontal axis represents the predicted probability, and the vertical axis represents the observed frequency. The ideal diagonal indicates perfect calibration. Before calibration, the probability outputs exhibit systematic bias, particularly in the mid-range, where the observed frequency deviates notably from the predicted probability. After calibration, the curves for both risk types align closely with the ideal diagonal, indicating that the calibrated probabilities are consistent with the actual occurrence frequencies. This confirms that isotonic regression effectively corrects the bias in the probability outputs.

4.4 Risk Level Determination and Temporal Distribution

4.4.1 Risk Level Determination

Table 2 and Table 3 present the confusion matrices of the proposed method for the two types of risk events, and Table 4 compares its determination accuracy, hit rate and false-alarm rate with those of the two baselines on the test set.

The proposed method attains a determination accuracy of 92.7% for curtailment events and 84.2% for security of supply events, against trivial majority-class baselines of 74.2% and 55.2%, respectively. The corresponding hit rates are 85.5% and 88.5%, so that hit rates alone exceed 84% for both types. The two risks are, however, handled with opposite conservatism. On the curtailment side the model is restrained: a hit rate of 85.5% is obtained with a false-alarm rate of only 4.8%. On the security of supply side it over-reports: the hit rate rises to 88.5% but the false-alarm rate reaches 21.0%, which is the intended behaviour of a warning system that prioritises the avoidance of missed alarms. This asymmetry is consistent with the negative correlation between the two risks established in Section 2.

Against the other two models, the proposed method achieves the highest determination accuracy on both events. The improvements are 1.7 and 1.0 percentage points on curtailment events, and 3.9 and 4.9 percentage points on security of supply events, relative to GBDT and MGAT, respectively. MGAT shows a slightly higher hit rate on the security of supply side, but its false-alarm rate is 32.2% compared with 21.0% for the proposed method, and its determination accuracy is 4.9 percentage points lower. The proposed method therefore provides the most balanced performance in terms of determination accuracy and false-alarm rate.

Table 2 Confusion Matrix for Curtailment Event Determination

Actual \ Predicted	Predicted risk	Predicted no risk	Total
True risk	1159	196	1355
No risk	186	3712	3898
Total	1345	3908	5253

Table 3 Confusion Matrix for Security of Supply Event Determination

Actual \ Predicted	Predicted risk	Predicted no risk	Total
True risk	2567	333	2900
No risk	495	1858	2353
Total	3062	2191	5253

Table 4 Comparison of determination accuracy, hit rate and false-alarm rate on the test set

(a) Curtailment event (curtailed energy ≥ 5 GW)

Model	Accuracy (%)	Hit rate (%)	False alarm (%)
Proposed	92.7	85.5	4.8
GBDT	91.0	82.2	6.0
MGAT	91.8	84.1	5.6

(b) Security of supply event (activated reserve ≥ 3 GW)

Model	Accuracy (%)	Hit rate (%)	False alarm (%)
Proposed	84.2	88.5	21.0
GBDT	80.4	87.3	28.2
MGAT	79.3	88.7	32.2

4.4.2 Temporal Distribution of Risk Levels

To further illustrate the temporal characteristics of the four risk levels, Fig. 5 presents the risk level distribution on a typical curtailment day and a typical security of supply day.

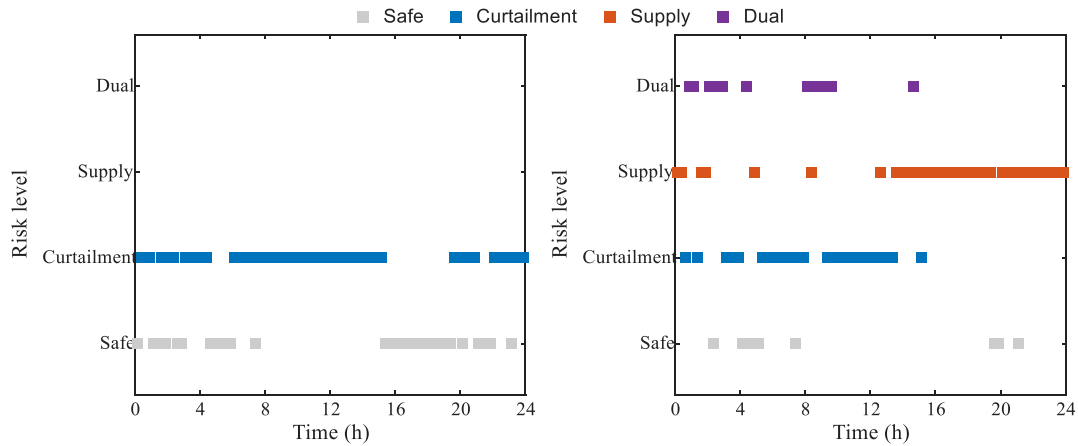


Fig. 5 Temporal distribution of risk levels on a typical curtailment day and a typical security of supply day

In Fig. 5, the left panel shows a typical curtailment day, where the safe level dominates during the night and early morning, while curtailment risk is concentrated around midday, corresponding to the period of high PV output and low net load. Security of supply risk

and dual risk rarely appear on this day. The right panel shows a typical security of supply day, where security of supply risk is concentrated during the evening peak, while curtailment risk and dual risk appear around midday. These temporal patterns are consistent with the theoretical analysis in Section 2.1. And reserve shortage events occur mostly during the evening peak. The complementary temporal characteristics of the two risk types are therefore verified by the model outputs.

4.5 Risk Tracing

To verify the risk tracing capability of the proposed method, a dispatch interval at 12:00 on a high-curtailment day is selected for demonstration. The curtailed energy, gating weights, bus graph attention weights, and bus curtailment rate distribution for this interval are presented in Figs. 6–8.

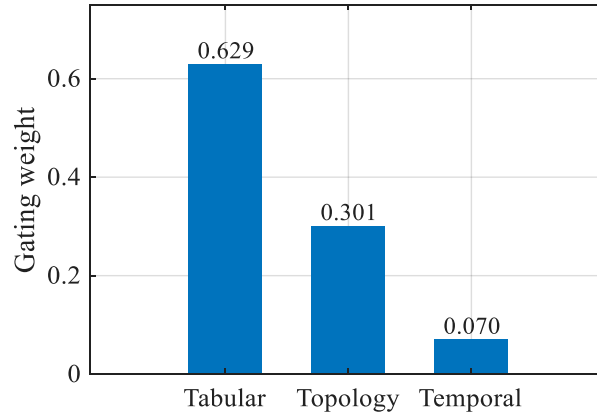


Fig .6 Gating weight distribution for the dispatch interval

Fig. 6 shows the gating weights of the three experts for this dispatch interval. The tabular expert has the highest weight of 0.629, followed by the topology expert at 0.301, and the temporal expert at only 0.070. This indicates that the curtailment risk of this dispatch interval is mainly driven by the source-side and schedule-side tabular features, namely wind, PV, and load forecast values as well as unit startup/shutdown status. The grid-side topology structure contributes to a lesser extent, and the temporal fluctuation has the smallest impact.

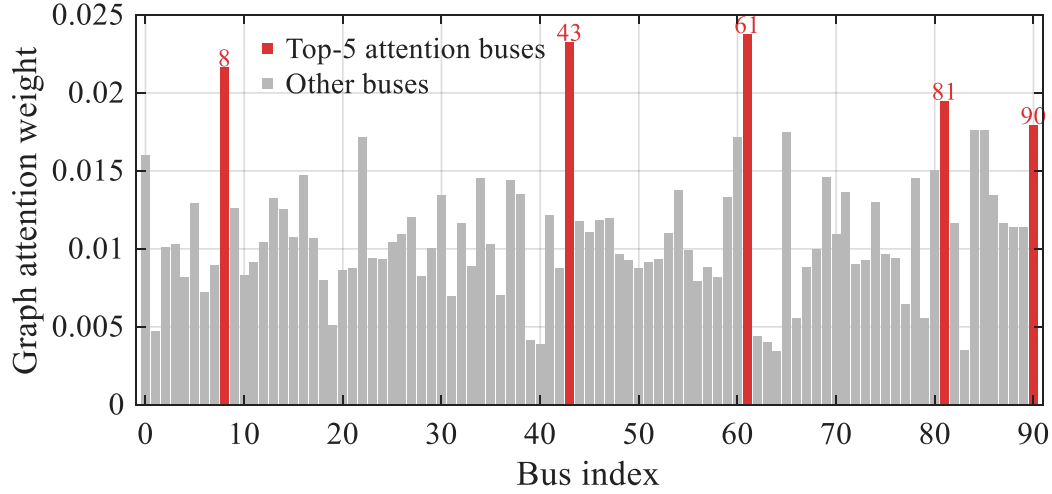


Fig .7 Graph attention weight distribution of buses

Fig. 7 shows the graph attention weights of all 91 buses for this dispatch interval. The top five buses with the highest weights are 61, 43, 8, 81, and 90 in descending order. The top five buses by model attention shown here are the five buses with the highest loading rates.

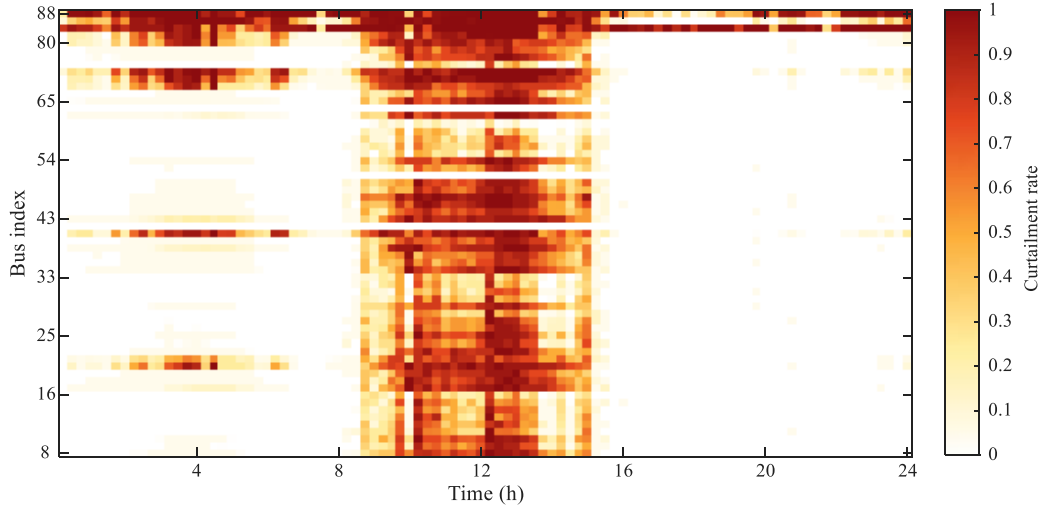


Fig .8 Curtailment rate distribution of renewable energy buses

Fig. 8 shows the distribution of curtailment rates of 61 renewable energy buses on that day in both temporal and spatial dimensions. The horizontal axis represents time, and the vertical axis represents bus number. Buses marked with an asterisk are those with the highest attention. From the temporal dimension, curtailment is concentrated between 9:00 and 15:00, which coincides with the period of high midday PV output. From the perspective of energy distribution, curtailment tends to be concentrated on certain buses.

The model outputs are summarized, and the risk tracing results are presented in Table 5.

Table 5 Risk Tracing Results for a Single Dispatch Interval

Bus	Curtailment rate	Loading rate of connected line
61	0.765	0.058
43	0.900	0.065
8	1.000	0.087
81	—	0.864
90	—	0.124

To verify the reliability of the spatial tracing results, Table 5 lists the operating states of the above five buses at the 12:00 dispatch interval. The model prediction results are compared with these states, and the operating states are basically consistent with the heatmap. Buses 61, 43, and 8 are buses where curtailment actually occurred. The model assigns the highest attention weights to these three buses, which is consistent with the actual spatial distribution of curtailment on that day. The second category includes buses 81 and 90, whose curtailment rates are zero and which are not the actual sources of curtailment on that day, but the loading rates of their connected lines are 0.864 and 0.124, respectively, with the line connected to bus 81 being in a heavily loaded state. This indicates that when determining spatial risk sources, the model not only focuses on buses where curtailment actually occurs but also assigns higher weights to buses with heavy power flow that may propagate risk, reflecting the graph attention mechanism's perception of topological importance.

In summary, the tracing results provided by the proposed method in the three dimensions of information type, spatial location, and temporal location are basically consistent with the actual operation conditions of that day. The above results indicate that the gating weights and graph attention coefficients within the MMoE architecture can provide dispatch operators with stable and reproducible risk tracing information without relying on post-hoc explanation methods.

5 Conclusion

This paper proposed a power grid operational risk assessment method based on a source-grid multi-dimensional feature fusion multi-gate mixture-of-experts model. The main results are summarized as follows.

(1) A two-stage dispatch simulation framework is proposed to generate risk labels from empirical forecast error distributions. Case studies show that the generated risk labels are consistent with actual dispatch constraints, and the dedicated temporal expert effectively extracts the temporal correlation features of renewable output and load.

(2) The improved MMoE model integrates source-side tabular features, grid-side topology features, and temporal features. Experimental results show that the proposed method achieves a determination accuracy of 92.7% for curtailment events and 84.2% for security of supply events, outperforming GBDT and MGAT.

(3)The gating weights and graph attention coefficients are directly used for risk tracing. The tracing results in the three dimensions of information type, spatial location, and temporal location are consistent with actual operation conditions, indicating that the model can provide stable and reproducible tracing information without post-hoc computation.

Future work will extend the proposed method to incorporate multi-type energy storage and demand-side response, and to explore online learning mechanisms for model updating in evolving power system conditions.

6 References

- [1] M. Du, X. Zhang, J. Zhang, Z. Wang, and V. Terzija, "Distributionally robust chance-constrained unit commitment for power systems considering wind power curtailment and load shedding levels," *IEEE Trans. Autom. Sci. Eng.*, vol. 23, pp. 2621–2635, 2026, doi: 10.1109/TASE.2026.3653716.
- [2] M. Z. Heidrich, M. A. F. Reis, C. K. Yamaguchi, L. O. Seman, and S. F. Stefenon, "Patch-based transformer with mixture of experts and conformal inference for curtailment time series forecasting," *IEEE Access*, vol. 14, pp. 64144–64160, 2026, doi: 10.1109/ACCESS.2026.3686958.
- [3] J. Li, H. Guo, W. Yang, Y. Xie, H. Zeng, and J. Li, "An improved probability efficient point method based on outer approximation for unit commitment of multi-area power systems," *Results Eng.*, vol. 28, p. 108347, 2025, doi: 10.1016/j.rineng.2025.108347.
- [4] Y. Yang, Y. Zhou, S. He, B. Hu, Y. Qu, T. Liu, and X. Guan, "A Markov tree model for cascading failure risk assessment in power grid with uncertain renewable energy generation," *IEEE Trans. Circuits Syst. I: Reg. Papers*, vol. 73, no. 2, pp. 1460–1473, 2026, doi: 10.1109/TCSI.2025.3601132.
- [5] G. Zhao, X. Luo, S. Yoo, et al., "GAN-based extreme conditional distribution estimation for renewable energy systems," in *Proc. 2024 Int. Conf. Commun. Control Comput. Technol. Smart Grids*, 2024, pp. 562–567.
- [6] C. Yarramsetty, T. Moger, D. Jena, and V. S. Rao, "Advances in composite power system reliability assessment: Trends and machine learning role," *IEEE Access*, vol. 13, pp. 80871–80888, 2025, doi: 10.1109/ACCESS.2025.3567464.
- [7] D. Pandit, R. K. Saket, A. Bera, and T. Nguyen, "Reliability analysis of power grids considering component failures of variable energy resources," *IEEE Trans. Ind. Appl.*, vol. 61, no. 6, pp. 8748–8759, 2025, doi: 10.1109/TIA.2025.3577149.
- [8] M. Draz, F. Pagel, and S. Albayrak, "Probabilistic risk assessment in power systems with high wind energy penetration," *IEEE Access*, vol. 12, pp. 140097–140111, 2024, doi: 10.1109/ACCESS.2024.3463882.
- [9] M. E. Za'ter, A. Sajadi, and B.-M. Hodge, "Semi-supervised multi-task learning based framework for power system security assessment," *Int. J. Electr. Power Energy Syst.*, vol. 170, p. 110910, 2025, doi: 10.1016/j.ijepes.2025.110910.
- [10] Z. Cao, Z. Li, L. Wang, et al., "Multi-region thermal–renewable coordinated dispatch considering chance-constrained reserve and supply security," in *Proc. 2026 IEEE Int. Symp. Circuits Syst. (ISCAS)*, 2026, pp. 4849–4853.
- [11] S. Gao, D. You, X. Chen, et al., "Operational risk assessment of power distribution systems: A physics-informed multi-task learning approach," *Int. J. Electr. Power Energy Syst.*, vol. 172, p. 111309, 2025, doi: 10.1016/j.ijepes.2025.111309.
- [12] Y. Zhang, X. Wu, G. Guo, et al., "Scenario generation of renewable energy power considering identification of meteorological risk factors," in *Proc. 8th IEEE Conf. Energy Internet Energy Syst. Integr. (EI2)*, 2024, pp. 1433–1437.

- [13] D. Veljković and Ž. Đurišić, “Thermal-dynamic critical contingency classification to reduce renewable energy curtailment in power systems,” *Int. J. Electr. Power Energy Syst.*, vol. 177, p. 111717, 2026, doi: 10.1016/j.ijepes.2026.111717.
- [14] A. Lorca and X. A. Sun, “Multistage robust unit commitment with dynamic uncertainty sets and energy storage,” *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 1678–1688, May 2017, doi: 10.1109/TPWRS.2016.2593422.
- [15] T. Ding, Z. Zeng, M. Qu, et al., “Two-stage chance-constrained stochastic thermal unit commitment for optimal provision of virtual inertia in wind-storage systems,” *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 3520–3530, 2021, doi: 10.1109/TPWRS.2021.3051523.
- [16] T. Jain and K. Verma, “Reliability enhancement with coordinated operation of wind power and battery energy storage using machine learning based unit commitment decision,” *J. Energy Storage*, vol. 111, 2025, doi: 10.1016/j.est.2025.115455.
- [17] H. Sheng and Y. Xu, “Risk-based dispatch of a three-phase unbalanced distribution network for maximizing renewable generation,” in *Proc. 2025 IEEE Power Energy Soc. General Meeting (PESGM)*, Austin, TX, USA, 2025, pp. 1–5.
- [18] G. Terrén-Serrano and M. Ludkovski, “Extreme day-ahead renewables scenario selection in power grid operations,” *Appl. Energy*, vol. 391, p. 125747, 2025, doi: 10.1016/j.apenergy.2025.125747.
- [19] Y. Liu, T. Ding, Y. Ding, et al., “Two-stage robust security-constrained unit commitment for wind power reserve provision optimisation under decision-dependent uncertainty,” *IET Gener. Transm. Distrib.*, vol. 20, no. 1, 2026, doi: 10.1049/gtd2.70251.
- [20] J. Lee, J. Cho, S. A. Rodrigazo, et al., “Feasibility study of solar power curtailment prediction using multivariate dynamic time warping-based machine learning: A case study of Jeju island,” *Electr. Power Syst. Res.*, vol. 253, p. 112517, 2026, doi: 10.1016/j.epr.2025.112517.
- [21] S. Li, L. Wu, T. Peng, et al., “RSynLLM: A risk-aware routing mixture-of-experts large language model for multi-energy load forecasting in large-scale distribution networks,” *Appl. Energy*, vol. 406, p. 127227, 2026, doi: 10.1016/j.apenergy.2025.127227.
- [22] R. Zhang, X. Dai, and H. Wang, “TGN-RiBC: A betweenness-aware temporal graph neural network for power grid risk prediction under dynamic unit commitment,” in *Proc. 51st Conf. Ind. Electron. Soc. (IECON)*, 2025, pp. 1–6.

Declarations

- **Funding:** This work was supported by the Science and Technology Program of State Grid Jiangsu Electric Power Co., Ltd. [grant number J2025136].
- **Author contribution:** The authors confirm contribution to the paper as follows: conceptualization, Shupeng Zhang and Xing Luo; methodology, Shupeng Zhang; software, Chenguang Qiu; validation, Chenguang Qiu and Chaojie Ding; formal analysis, Chenguang Qiu; investigation, Chaojie Ding; resources, Jie Li; data curation, Chaojie Ding; writing—original draft preparation, Shupeng Zhang; writing—review and editing, Jie Li; visualization, Chaojie Ding; supervision, Xing Luo; project administration, Jie Li; funding acquisition, Xing Luo. All authors reviewed and approved the final version of the manuscript.