modelling

Article

Spatiotemporal Coupled State Prediction Model for Local Power Grids Under Renewable Energy Disturbances

Zhixin Suo [1], Jingyang Zhou [2], Yukai Chen [1,3], Zihao Zhang [1,3], Liang Zhao [1], Shanshan Bai [1,3], Pengyu Wang [1] and Kangli Liu [2,*]

[1] National Energy Digital Grid Technology Research and Development Center (Southern Power Grid Digital Grid Research Institute Co., Ltd.), Guangzhou 510700, China
[2] School of Electrical Engineering, Southeast University, Nanjing 210096, China
[3] Guangdong Provincial Key Laboratory of Digital Grid Technology (Southern Power Grid Digital Grid Research Institute Co., Ltd.), Guangzhou 510700, China
* Correspondence: kangli@seu.edu.cn

Abstract

The modern power system is becoming increasingly complex, and the uncertainty in the operation of each link has intensified the possibility of risks emerging. Therefore, efficient risk prediction is of great significance for maintaining the reliable operation of the entire system. In this paper, to address the uncertainty and spatiotemporal coupling in local power grids with renewable integration, an integrated "state prediction–risk assessment–early warning" framework is proposed. A spatiotemporal graph neural network is used to predict node voltage, power, and phase angles under topological constraints, where physics-aware graph attention, disturbance-enhanced temporal modeling, and prediction-smoothing constraints are jointly incorporated to improve sensitivity to renewable fluctuations and ensure stable multi-step forecasting. Furthermore, voltage deviation, power fluctuation, and phase-angle variation are quantified to compute a composite risk index via normalized softmax weighting, with factor contributions enhancing interpretability. Test results on the IEEE 33-bus system under diverse disturbances show improved accuracy and stability over baselines, showing consistently lower MAE/RMSE than three baselines across all disturbance scenarios while pinpointing high-risk nodes and causes, highlighting good engineering potential.

Keywords: renewable energy disturbances; spatiotemporal coupled state; prediction model; risk assessment

check for updates

Academic Editor: Wei Gao

Received: 8 October 2025
Revised: 26 November 2025
Accepted: 28 November 2025
Published: 5 December 2025

Citation: Suo, Z.; Zhou, J.; Chen, Y.; Zhang, Z.; Zhao, L.; Bai, S.; Wang, P.; Liu, K. Spatiotemporal Coupled State Prediction Model for Local Power Grids Under Renewable Energy Disturbances. *Modelling* **2025**, *6*, 161. https://doi.org/10.3390/modelling6040161

1. Introduction

With the widespread integration of renewable energy into modern power systems, operational uncertainty has increased significantly [1,2]. In distribution networks with intermittent sources such as distributed wind and photovoltaics, output fluctuations frequently cause sharp variations in nodal power, voltage, and phase angle, threatening system stability [3]. Recent studies on microgrid infrastructures indicate that microgrid structure and renewable-energy coupling have notable impacts on system stability [4,5]. Disturbances exhibit both temporal dynamics and topological constraints, propagating along physical paths [6]. Their spatial effects are structurally dependent and non-Euclidean, making it difficult for traditional static or rule-based models to characterize them accurately and in real time [7]. This weakens the perception–warning–control chain of local grids and introduces risks of delayed dispatch [8]. To overcome these limitations, artificial intelligence methods have been explored: convolutional neural networks (CNNs) are effective at local feature extraction but struggle with complex topologies; recurrent neural networks (RNNs) and long short-term memory (LSTM) networks perform well in temporal sequence modeling but neglect spatial dependencies; and conventional machine learning approaches, though effective in specific cases, rely heavily on prior knowledge and show limited generalizability.

In recent years, graph-based methods have been used to model physical bus connections and support disturbance propagation and state prediction [9,10]. Existing studies include GNN-based event detection and localization [11], convolution–attention frameworks for transient frequency stability prediction [12,13], spatiotemporal graph convolution for frequency response forecasting [14], and the combination of deep graph learning with spatiotemporal recurrence for fault diagnosis and localization [15,16]. However, challenges remain in renewable-dominated local grids: some methods suffer from high computational costs [17,18], while data scarcity and topological uncertainty undermine robustness. For example, the structure-aware graph learning approach relies on complete PMU measurements, which are impractical in local grids [7], whereas the robust temporal graph convolution proposed alleviates topological noise but lacks explicit modeling of rapid disturbances [8].

These limitations underscore the necessity for a unified modeling paradigm capable of simultaneously capturing physics-informed spatial dependencies, disturbance-driven temporal dynamics, and associated operational risks within an interpretable framework. Existing spatial-only or temporal-only methods are inherently insufficient to characterize such tightly coupled spatiotemporal interactions. By jointly embedding physical topology and disturbance-induced temporal evolution, a more comprehensive framework can significantly enhance predictive accuracy and robustness under renewable generation variability while also enabling real-time, risk-aware operational decision-making. To address these issues, this paper proposes a Temporal Dynamic Weighted Graph Neural Network (TDWGNN). The method leverages grid topology and nodal time-series data, incorporating a dynamic weighting mechanism and recurrent units to jointly capture disturbance propagation paths and temporal evolution. A disturbance-aware enhancement term is introduced to strengthen adaptability to rapid fluctuations, and a prediction-smoothing constraint is incorporated to mitigate multi-step error accumulation and ensure stable trajectory evolution. These additions help improve model training and long-term prediction behavior. Building on this, an integrated framework of "prediction–risk assessment" is established, where indicators such as voltage deviation, power fluctuation, and phase-angle variation rate enable quantitative risk evaluation and graded early warning. The framework provides interpretable node-level indicators that enhance its engineering applicability in renewable-rich distribution networks, thereby offering practical guidance for the secure operation of distribution networks with high penetration of distributed renewable energy. Moreover, the proposed framework is extensible and can be adapted to hybrid renewable–storage systems by incorporating storage interaction states into node features, enabling flexible modeling of systems with dynamic charging–discharging behavior. Compared with existing studies, the main contributions of this work can be summarized as follows:

1. A TDWGNN framework is developed that embeds physical topology into spatiotemporal learning to better reflect distribution-network structural characteristics.
2. The disturbance-aware temporal module is designed to enhance responsiveness to sudden renewable fluctuations beyond standard recurrent or ST-GNN models.
3. The prediction-smoothing constraint is introduced to reduce error accumulation and suppress oscillations, improving multi-horizon forecasting stability.

4. A unified prediction–risk assessment pipeline is established, linking predicted states to quantitative node-level risk indicators for interpretable early warning.

The remainder of this paper is organized as follows. Section 2 introduces the modeling of graph-node dependencies based on distribution-network physical topology. Section 3 presents the proposed TDWGNN, including topology embedding, disturbance-aware temporal modeling, and prediction-smoothing constraints, as well as the risk assessment framework. Section 4 reports the experimental setup and results on the IEEE 33-bus system. Finally, Section 5 concludes the paper and discusses potential future work.

2. Graph-Node-Dependency Modeling Based on Physical Topology

2.1. Graph Structure Modeling

The physical topology and node-to-node connections of the local distribution grid are shown in Figure 1.

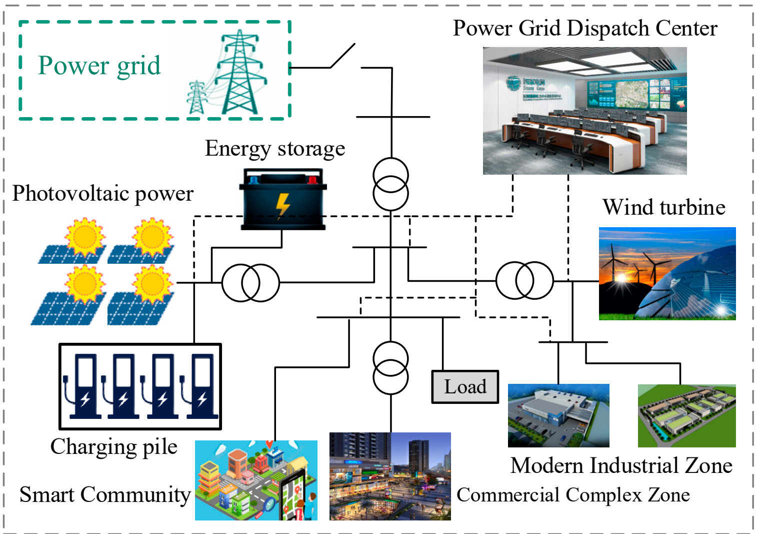


Figure 1. Schematic Diagram of Local Power Grid.

In power system modeling, a distribution network can be naturally abstracted as a graph, where nodes represent buses or equipment connection points, and edges correspond to physical transmission lines. This study models the distribution network as a weighted undirected graph $G = (V, E, A)$, where V denotes the set of nodes in the system, E represents the set of edges indicating direct physical connections between nodes, and $A \in \mathbb{R}^{N \times N}$ is the adjacency matrix characterizing both the connectivity and the strength of interactions between nodes. To enhance the physical accuracy of modeling, a branch-parameter-based weighting mechanism is introduced, where line impedance is adopted as the metric for edge weights. Specifically, for any two connected nodes i and j, the corresponding edge weight A_{ij} is defined as:

$$A_{ij} = \begin{cases} \frac{1}{|Z_{ij}|}, & \text{if } (v_i, v_j) \in E \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

where $Z_{ij} = R_{ij} + jX_{ij}$ denotes the complex impedance of the branch between nodes. This weighting scheme preserves the sparsity of the graph structure while retaining the key characteristics of the original circuit topology. To address numerical stability and training convergence issues in graph neural networks, self-loops are first added to the adjacency matrix to ensure that each node retains its own information during propagation:

$$\widetilde{A} = A + I \tag{2}$$

where I denotes the identity matrix.

Subsequently, a symmetric normalization strategy is applied to stabilize feature propagation and prevent gradient vanishing:

$$\hat{A} = D^{-\frac{1}{2}} \widetilde{A} D^{-\frac{1}{2}} \tag{3}$$

where D is the degree matrix, defined as $D_{ii} = \sum_j \widetilde{A}_{ij}$.

This normalization strategy helps prevent gradient vanishing or feature dilution during feature propagation, while $\hat{A}$ serves as the fundamental operator for information transmission in the graph neural network.

2.2. Dynamic Adjacency Dependency Modeling Based on Graph Attention Mechanism

In the context of frequent renewable energy disturbances, heterogeneity in operating states, supply–load characteristics, and topological positions leads to disturbance propagation that is inherently directional and asymmetric. To capture such properties, a Graph Attention Mechanism (GAT) is employed to enhance the modeling of node dependencies. By learning attention weights between adjacent nodes, the GAT dynamically adjusts the aggregation strength during feature propagation, thereby allowing each node to place greater emphasis on neighbors with stronger correlations when updating its representation. The original feature vectors associated with node i and its neighbor j be denoted as $h_i, h_j \in \mathbb{R}^F$. The model first performs a linear transformation and concatenation of these features, and then computes the unnormalized attention score through a feedforward neural network:

$$e_{ij} = \text{LeakyReLU}\left(\alpha^T \left[Wh_i \middle|\middle| Wh_j\right]\right) \tag{4}$$

Here, $W \in \mathbb{R}^{F' \times F}$ is a shared linear transformation matrix, $\alpha \in \mathbb{R}^{2F'}$ denotes a learnable attention vector, F represents the dimensionality of the original input features of each node, F' is the output dimensionality of each attention head, and | | denotes the vector concatenation operator. The attention coefficients are normalized using the softmax function to yield the final attention weights for feature aggregation:

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k \in \mathcal{N}(i)} \exp(e_{ik})} \tag{5}$$

where $\mathcal{N}(i)$ denotes the set of neighboring nodes of node i. To further enhance the expressive capacity and stability of the model, a multi-head attention mechanism is employed. Specifically, K independent attention heads are executed in parallel within the same layer, each equipped with its own linear transformation and attention parameters, thereby generating node representations from multiple perspectives. For intermediate layers, the outputs of these attention heads are concatenated into a single extended vector:

$$h_i' = \mathop{||}_{k=1}^{K} \sigma\left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij}^{(k)} W^{(k)} h_j\right) \tag{6}$$

Considering that the concatenated representation has a dimensionality of $K \cdot F'$, directly feeding it into downstream tasks may lead to an excessive number of parameters and weaken the fusion of information across different attention heads. To address this issue, a linear mapping layer is introduced after multi-head attention concatenation to unify the dimensionality and enhance semantic integration. The transformation is formulated as:

$$h_i'' = W_{out} \cdot \left(\overset{K}{\underset{k=1}{||}} h_i^{(k)} \right) + b_{out} \tag{7}$$

Here, $W_{out} \in \mathbb{R}^{F'' \times (K \cdot F')}$ denotes the projection matrix, b_{out} is the bias term, and F'' is the target dimensionality. This layer not only controls model complexity but also enables the adaptive integration of multi-head attention features, thereby providing a unified node representation for subsequent tasks. Considering that each edge in the power grid graph carries explicit physical meaning, the branch-impedance-based physical edge weights are further fused with the data-driven attention mechanism. Specifically, the inverse form of branch impedance is incorporated as a prior adjustment term in the attention weights, thereby enhancing the physical consistency of disturbance propagation modeling:

$$\widetilde{\alpha}_{ij} = a_{ij} \cdot f\left(|Z_{ij}|^{-1} \right) \tag{8}$$

Here, $f(\cdot)$ denotes the edge-weight normalization function, which ensures consistency between the incorporated physical information and the distribution of attention weights. This strategy balances the flexibility of graph neural networks with the structural and physical constraints of power systems, thereby improving the interpretability and generalization capability of the model in capturing disturbance response patterns. The above algorithm is shown in Algorithm 1.

Algorithm 1 Graph-Node-Dependency Modeling with Physics-Aware Attention

Require: Node set V, edge set E with branch impedance Z_{ij}, input features $\{h_i\}$; number of heads K
Ensure: Updated node features $\{h_i{}'\}$

Graph construction: Initialize adjacency matrix $A \leftarrow 0$
for all $(i, j) \in E$ **do**
  $A_{ij} \leftarrow 1/|Z_{ij}|$; $A_{ji} \leftarrow A_{ij}$
end for
for all $i \in V$ **do**
  $A_{ii} \leftarrow 1$
end for
Normalize adjacency: A
Attention-based aggregation:
for all $i \in V$ **do**
for k = 1 to K **do**
  $Wh_i \leftarrow$ linear transform of h_i by $W^{(k)}$
for all $j \in \mathcal{N}(i)$ **do**
    $Wh_j \leftarrow$ linear transform of h_j by $W^{(k)}$
    $e_{ij}^{(k)} \leftarrow \mathrm{LeakyReLU}\left(\left(a^{(k)} \right)^T [Wh_i || Wh_j] \right)$
end for
  $a_{ij}^{(k)} \leftarrow \mathrm{softmax}_{j \in \mathcal{N}(i)} \left(e_{ij}^{(k)} \right)$
  $\widetilde{a}_{ij}^{(k)} \leftarrow a_{ij}^{(k)} \cdot A_{ij}$
  $z_i^{(k)} \leftarrow \sigma\left(\sum_{j \in \mathcal{N}(i)} \widetilde{a}_{ij}^{(k)} Wh_j \right)$

Algorithm 1 *Cont.*

20: **end for**
21: $h_i^{(mh)} \leftarrow \text{concat}\left(z_i^{(1)}, \ldots, z_i^{(K)}\right)$
22: $h_i{}' \leftarrow \sigma\left(W_{out} h_i^{(mh)} + b_{out}\right)$
23: **end for**
24: **return** $\{h_i{}'\}$

2.3. Construction of Node State Features

On the basis of the constructed grid topology and edge-weight modeling, a stable and learnable state representation is established for each node as the input to the graph neural network. The original state feature vector of node i at time t is denoted by $x_i(t) \in \mathbb{R}^d$, where d denotes the feature dimension of a single node at each time step. To enhance representational capacity and ensure training stability, all state variables are normalized to a unified range, avoiding gradient imbalance and feature dominance caused by scale differences. A fixed-length temporal window is further introduced to model the state sequence of each node over T consecutive steps. With the current prediction time denoted as t, the input of node i is given by:

$$X_i = [x_i(t-T+1),\ x_i(t-T+2) \ldots, x_i(t)] \in \mathbb{R}^{T \times d} \tag{9}$$

3. Disturbance-Aware Spatiotemporal Coupled State Prediction Model

3.1. Architecture of a Temporal State Prediction Model

Building on the graph representation established in Section 2, a spatiotemporally coupled state prediction model is proposed, which integrates the spatial perception capability of graph neural networks with the temporal modeling mechanism of recurrent neural structures, thereby enabling rolling state prediction across multiple time scales. At each time step $t' \in [t-T+1, t]$, the feature vectors of all nodes $\{x_i(t')\}$ are taken as input and processed by a GAT, which performs neighborhood feature aggregation based on grid topology and the attention mechanism, thereby producing structure-aware node embeddings:

$$z_i(t') = \text{GAT}\left(x_i(t'), \mathcal{N}(i), \widetilde{A}\right) \in \mathbb{R}^{d_z} \tag{10}$$

Here, d_z denotes the embedding dimension, resulting in a graph embedding sequence for each node over T time steps:

$$Z_i = [z_i(t-T+1), z_i(t-T+2) \ldots, z_i(t)] \in \mathbb{R}^{T \times d_z} \tag{11}$$

Here, Z_i denotes the spatial embedding of X_i obtained through the Graph Attention Network, and d_z represents the embedding dimension. This process preserves the physical interpretability of the original node state features while enhancing the representation of disturbance propagation in the network topology via the attention mechanism. Once the spatial representations of nodes over multiple time steps are obtained, they are fed into a Gated Recurrent Unit (GRU) to capture the temporal evolution of disturbances. The update process of the GRU unit is formulated as follows:

$$\begin{aligned}
r_i(t') &= \sigma(W_r z_i(t') + U_r h_i(t'-1) + b_r) \\
u_i(t') &= \sigma(W_u z_i(t') + U_u h_i(t'-1) + b_u) \\
\widetilde{h}_i(t') &= \tanh(W_h z_i(t') + U_h(r_i(t') \odot h_i(t'-1)) + b_h) \\
h_i(t') &= (1 - u_i(t')) \odot h_i(t'-1) + u_i(t') \odot \widetilde{h}_i(t')
\end{aligned} \tag{12}$$

Here, $r_i(t')$ denotes the reset gate, which controls the degree of information forgetting; $u_i(t')$ represents the update gate, which governs the fusion of new and past information; $\widetilde{h}_i(t')$ is the candidate hidden state. W, U, and b are learnable parameters associated with the gating operations. The final hidden state $h_i(t)$ of the GRU encapsulates the historical disturbance information of each node within the temporal window and serves as the core representation for state prediction.

3.2. Temporal Response Enhancement Mechanism Under Disturbance Awareness

To enhance the model's sensitivity and representational capacity for disturbance evolution, a disturbance-driven enhancement strategy is incorporated into temporal modeling, which is reflected in the following two aspects:

(1) To address the sharp variations in node states caused by disturbances along the temporal axis, this study introduces the node feature variation rate Δz. For a given node, let its spatial embedding features at two consecutive time steps be $z_i(t)$ and $z_i(t-1)$; the corresponding change in its graph embedding state is defined as:

$$\Delta z_i(t) = z_i(t) - z_i(t-1) \tag{13}$$

This differential term reflects the direct driving intensity of disturbances on the node state at time t and is used to assist in updating the temporal state representation. To ensure dimensional consistency, the embedding difference $\Delta z_i(t)$ is linearly projected into the GRU hidden space prior to its fusion with the hidden state through a learnable matrix W_z. This operation enables element-wise addition while preserving the physical interpretability of disturbance modulation.

To enhance the sensitivity of the GRU to sudden disturbances, it is incorporated into the hidden state update process, thereby constructing a disturbance-adjusted enhanced representation:

$$s_i(t) = h_i(t) + \gamma \cdot \Delta z_i(t) \tag{14}$$

Here, γ is a trainable disturbance-response coefficient that regulates the model's sensitivity to sudden changes. With this structure, the model preserves long-term trend memory while exhibiting greater dynamic adaptability to local disturbances.

(2) In multi-step prediction with a rolling update strategy, early-stage errors may gradually accumulate, causing oscillations in the output trajectory or even deviations from physical evolution laws. To suppress such non-physical oscillations, a prediction-smoothing constraint is required to limit abrupt variations between consecutive time steps:

$$\mathcal{L}_{smooth} = \sum_{h=1}^{H-1} \|\hat{y}_i(t+h+1) - \hat{y}_i(t+h)\|_2^2 \tag{15}$$

This smoothing regularization term is incorporated into the training objective function. By minimizing state variations between consecutive prediction steps, it suppresses error amplification and non-physical oscillations induced by disturbances, thereby improving the smoothness and controllability of the predicted trajectories. Finally, the disturbance-enhanced state $s_i(t)$ is used as the input to the output layer to predict the state of node i at the next time step:

$$\hat{y}_i(t+1) = W_o \cdot s_i(t) + b_o \tag{16}$$

The output is $\left[\hat{V}_i, \hat{\theta}_i, \hat{P}_i, \hat{Q}_i\right]_{t+1}$, corresponding to the predicted values of voltage magnitude, phase angle, active power, and reactive power, respectively. Using a rolling update strategy, the current prediction is fed back as the input for the next step, thereby generating the prediction sequence for the subsequent H time steps through iteration:

$$\hat{Y}_i = [\hat{y}_i(t+1), \hat{y}_i(t+2), \ldots, \hat{y}_i(t+H)] \in \mathbb{R}^{H \times d} \tag{17}$$

The prediction process is shown in Figure 2.

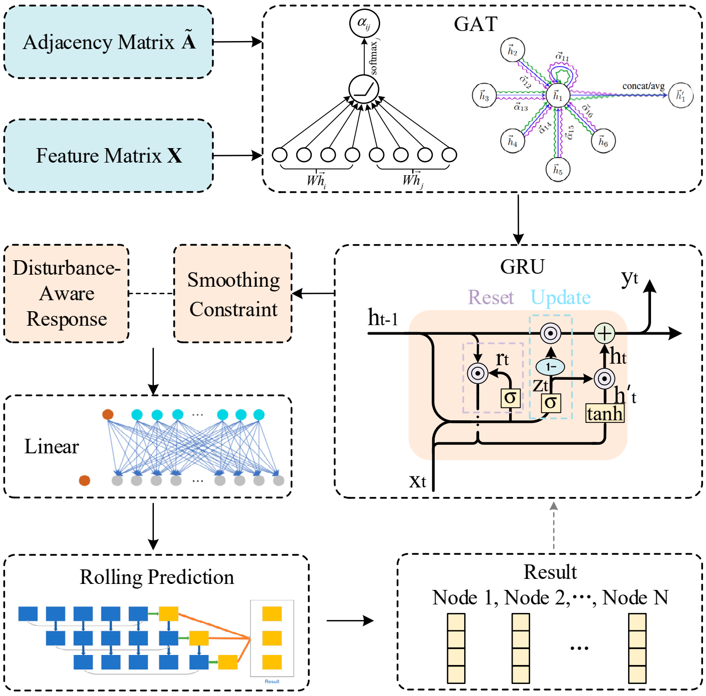


Figure 2. The Prediction Process of TDWGNN.

3.3. Risk Assessment Model

A risk assessment method based on the spatiotemporal coupled prediction model is proposed to rapidly evaluate grid operating conditions and support dispatching decisions. Predicted nodal states are used to quantify system risk, providing a basis for scheduling, early warning, and emergency response. Risk indicators are constructed from predicted voltage, current, active power, and reactive power, and compared with thresholds to determine stability and nodal risk levels. Based on the predicted sequences of grid nodes within the future time window $[t+1, t+H]$, the following three categories of risk factors are constructed:

(1) Voltage deviation risk: This factor measures the extent to which the voltage of a node deviates from its rated range within the prediction horizon. Let the predicted voltage of node i at future step H be $\hat{V}_i(t+h)$; the voltage deviation risk is then defined as:

$$R_i^{(v)} = \frac{1}{H} \sum_{h=1}^{H} \left(\frac{\max\left(\hat{V}_i(t+h) - V_i^{\max}, 0\right) + \max\left(V_i^{\min} - \hat{V}_i(t+h), 0\right)}{V_i^{\max} - V_i^{\min}} \right) \tag{18}$$

Here, $V_i^{\max}$ and $V_i^{\min}$ denote the permissible upper and lower voltage limits of the node.

(2) Power fluctuation risk: This factor considers both the intensity and abruptness of power fluctuations under disturbances. By adopting a coupled modeling approach of stability and mutation behavior, the power fluctuation risk is defined as:

$$\begin{aligned} R_i^{(P)} &= \lambda \cdot \frac{\operatorname{std}\left(\hat{P}_i(t+1:t+H)\right)}{P_i^{\mathrm{ref}}} \\ &+ (1-\lambda) \cdot \frac{1}{H-1} \sum_{h=2}^{H} \phi\left(\frac{\left|\hat{P}_i(t+h) - \hat{P}_i(t+h-1)\right|}{\delta_P} \right) \end{aligned} \tag{19}$$

where std$(\cdot)$ denotes the standard deviation of the predicted power, P_i^{ref} is the rated reference power, $\phi(x) = \frac{x^2}{1+x^2}$ represents a saturation activation function that suppresses excessively large slope values, δ_P is the safety fluctuation threshold, and $\lambda \in [0,1]$ is the weighting factor balancing stability and abruptness. This indicator enables the identification of both long-term volatility and short-term abrupt risks, making it more consistent with the uncertainties introduced by renewable energy disturbances in real operating scenarios.

(3) Phase-Angle Variation Risk: The variation of phase angles is closely related to system frequency. In renewable-dominated grids with insufficient inertia and weak frequency regulation capability, disturbances often cause sharp frequency fluctuations at nodes, which may further lead to severe consequences such as loss of synchronism or generator tripping. In scenarios where frequency prediction sequences are not directly available, the phase-angle variation rate is adopted as a proxy indicator of frequency dynamics to capture potential synchronization stability risks:

$$R_i^{(\theta)} = \frac{1}{H-1}\sum_{h=2}^{H}\left|\frac{\hat{\theta}_i(t+h) - \hat{\theta}_i(t+h-1)}{\Delta t}\right| \tag{20}$$

Here, $\hat{\theta}_i(t+h)$ denotes the predicted phase angle of the node, Δt represents the time step, and H is the length of the prediction window.

3.4. Comprehensive Risk Assessment Method

Considering the significant differences in dimension and numerical scale among voltage deviation, power fluctuation, and phase-angle variation rate, directly applying weighted summation may bias the results toward numerically dominant factors and obscure other critical risk sources. To achieve a unified evaluation of these three risk categories, all indicators are standardized to a common statistical scale, thereby enhancing the comparability and stability of their integrated representation. The three original risk factors of node i are denoted as a vector $R_i = \left[R_i^k, R_i^p, R_i^\theta\right]^T$, which is normalized using the max-normalization method:

$$\overline{R}_i^k = \frac{R_i^k}{\max_j^k R_j^k}, \quad k \in \{v, p, \theta\} \tag{21}$$

This normalization maps the risk values of all nodes into the range [0, 1], preserving their relative magnitudes. Here, $\max_j^k R_j^k$ denotes the maximum value of the k-th risk factor across all nodes within the current time window. A trainable global weight vector $w = \left[w^v, w^p, w^\theta\right]^T$ is then introduced and processed through the softmax function:

$$\alpha_i^k = \text{softmax}\left(\frac{e^{w^k}}{e^{w^v} + e^{w^p} + e^{w^\theta}}\right), \quad k \in \{v, p, \theta\} \tag{22}$$

The comprehensive risk value is finally obtained as:

$$R_i = \sum_{k \in \{v,p,\theta\}} \alpha_i^k \cdot R_i^k \tag{23}$$

4. Case Studies

To verify the adaptability and accuracy of the proposed state prediction and risk assessment model under renewable-dominated disturbance scenarios, the IEEE 33-bus distribution system is employed as the test platform. Renewable energy sources are integrated at selected nodes, and multiple representative disturbance scenarios are constructed to simulate the impacts of wind and solar output variations on nodal states, thereby generating a multidimensional time-series dataset for model training and evaluation.

4.1. Test System and Scenario Construction

The IEEE 33-bus system, a medium-scale benchmark widely used in distributed generation and smart distribution network studies, is characterized by a radial topology, significant power-flow coupling, and multi-branch meshed configurations. It consists of 33 nodes and 32 branches, with a nominal voltage of 12.66 kV and a total load of approximately 3.7 MW/2.3 MVar. Its original configuration is a radial distribution network with a single power source, making it suitable for testing voltage fluctuations, power distribution, and disturbance propagation. To construct representative renewable disturbance scenarios, 0.5 MW and 0.8 MW photovoltaic units are integrated at nodes 17 and 22, respectively, while 0.6 MW and 0.7 MW wind units are connected at nodes 25 and 32. Energy storage stations are additionally deployed at nodes 7 and 25 to analyze the impact of wind–solar complementary generation on system states. The selected nodes differ in load levels and topological positions, thereby facilitating heterogeneous disturbance propagation within the network.

As the focus of this study is on modeling and predicting local grid disturbance responses, renewable generation models are not analyzed in detail. Instead, physically consistent photovoltaic and wind models are adopted to simulate output fluctuations under typical operating conditions, providing disturbance-driven inputs for state prediction.

4.1.1. Photovoltaic Output Model

Photovoltaic generation is primarily driven by solar irradiance. To capture the periodic variation of irradiance over the course of a day, a sinusoidal function is employed to construct the baseline output curve, with a Gaussian disturbance term superimposed to emulate weather effects:

$$P_{PV}(t) = \eta_{pv} \cdot \max\left(0, \sin\left(\pi \cdot \frac{t - t_s}{t_e - t_s}\right)\right) \cdot \left(1 + \delta_{pv}(t)\right) \tag{24}$$

Here, η_{pv} denotes the conversion efficiency of the photovoltaic modules, $[t_s, t_e]$ represents the effective generation time window, and $\delta_{pv}(t) \sim \mathcal{N}\left(0, \sigma^2\right)$ is the irradiance disturbance term used to characterize output fluctuations caused by clouds, haze, and other weather conditions.

4.1.2. Wind Power Output Model

Wind power output exhibits a nonlinear relationship with wind speed. In this study, a three-segment power curve is employed to simulate the operating characteristics of wind turbines, while perturbed wind speeds are introduced to capture the effects of gusts and wind shear:

$$P_W(t) = \begin{cases} 0, & v(t) < v_{ci} \text{ or } v(t) \geq v_{co} \\ P_r \cdot \left(\frac{v(t) - v_{ci}}{v_r - v_{ci}}\right)^3, & v_{ci} \leq v(t) < v_r \\ P_r, & v_r \leq v(t) < v_{co} \end{cases} \tag{25}$$

Here, $v(t) = \overline{v} + \delta_w(t)$ denotes the wind speed sequence with disturbances; v_{ci}, v_r and v_{co} represent the cut-in, rated, and cut-out wind speeds, respectively; P_r is the rated output power of the wind turbine; and $\delta_w(t) \sim \mathcal{N}\left(0, \sigma^2\right)$ denotes the wind speed disturbance term.

To comprehensively capture the diverse disturbance patterns that local power grids may encounter in practice, three representative scenarios are designed: gradual disturbances, abrupt disturbances, and load disturbances. Gradual disturbances simulate conditions such as steadily increasing solar irradiance on clear days or smoothly rising wind speeds, under which renewable outputs exhibit continuous and trend-like variations with longer durations. Abrupt disturbances represent sudden events such as cloud occlusion or sharp wind fluctuations, leading to rapid and pronounced output variations with distinct non-stationary characteristics. Load disturbances are modeled by superimposing or reducing 10–20% of random variations on the baseline load curve to reflect the impact of demand-side uncertainty on system states.

4.2. Data Preprocessing and Graph Model Construction

After constructing disturbance-driven operating scenarios, the raw operational data must be organized into an input structure compatible with graph neural networks to enable effective modeling of system dynamics. This study focuses on short- to medium-term state prediction, where training samples are generated using a sliding time-window mechanism to form spatiotemporal graph inputs. Considering that disturbances such as wind speed fluctuations and cloud occlusion typically occur on a minute scale, the sampling frequency is set to one step per interval to enhance the capture of short-term dynamics. The prediction task is defined as forecasting the nodal states at time $t + H$ based on historical sequences up to time t. A window that is too short may omit disturbance propagation paths, whereas an excessively long window introduces redundancy and increases the risk of overfitting. Balancing disturbance response times, system inertia, and model complexity, the window length is set to $T = 6$, with a prediction horizon of $H = 4$. To verify the robustness of this parameter choice, comparative tests with different configurations were examined. Adjusting the input window to $T \in [4, 8]$ or the prediction horizon to $H \in [2, 6]$ yielded similar overall performance trends. Shorter windows limited the ability to capture propagation dynamics, while longer ones increased redundancy and the likelihood of overfitting; similarly, excessively large prediction horizons noticeably increased future-state uncertainty. These observations indicate that the chosen values offer a well-balanced trade-off between prediction accuracy, dynamic coverage, and computational stability.

The prediction targets encompass the state variables of all nodes at the specified time, with both inputs and outputs represented in tensor form. As shown in (9), the input sample is structured as $X_t \in \mathbb{R}^{N\times T\times d}$, which corresponds to $N = 33$ nodes in the system over the past $T = 6$ time steps, each step containing $d = 4$ state features: voltage magnitude, active power, reactive power, and voltage phase angle. This representation provides a comprehensive description of the dynamic behavior of nodes during disturbance propagation. As indicated in (17), the prediction target is $Y_{t+H} \in \mathbb{R}^{N\times d}$, representing the state distribution of all nodes at the future time step $t + H$. To construct stable and efficient training samples, a sliding sampling strategy is applied to each disturbance simulation sequence, with the window stride set to 1, thereby ensuring fine-grained coverage of the entire disturbance process.

Graph structure modeling adopts a static topology embedding, where the IEEE 33-bus distribution network is represented by an adjacency matrix $A \in \mathbb{R}^{N\times N}$. Its construction is based on the branch-impedance weighting mechanism described in Section 2, reflecting the physical accessibility of disturbance propagation. As illustrated in Figure 3, the adjacency matrix remains fixed across all samples and serves as a structural prior for information propagation in the graph neural network. The graph attention mechanism is then employed to adaptively learn the attention weights of neighboring nodes according to the current system states and disturbance characteristics, thereby dynamically adjusting the dependency strength among nodes.

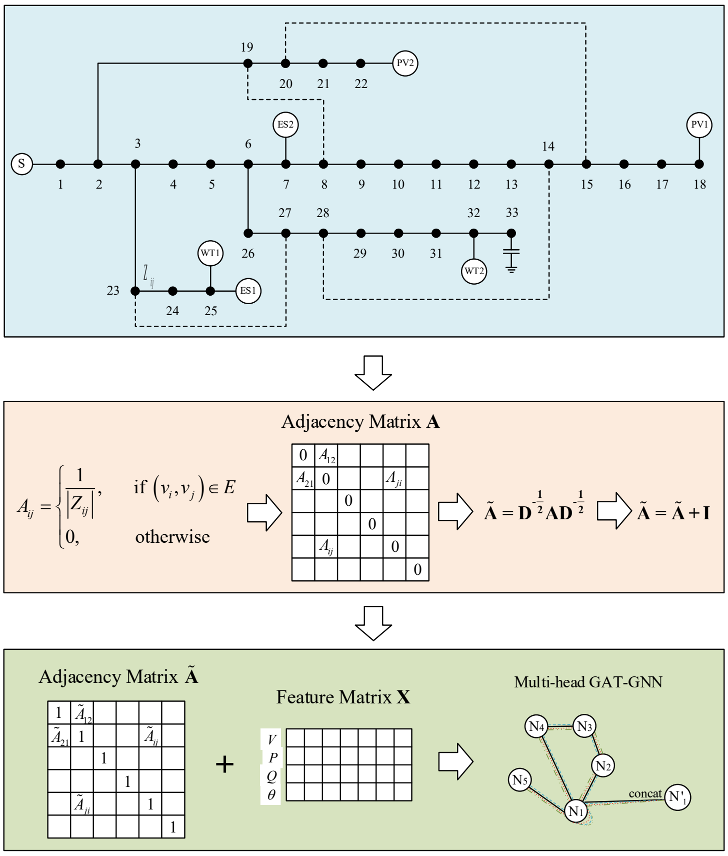


Figure 3. Construction of the adjacency matrix and modeling process based on the graph attention mechanism.

4.3. Temporal Network and Model Training Configuration

Building upon the constructed input samples and graph structure, a GRU is incorporated to establish a joint framework for state evolution, simulating the dynamic changes of system states under disturbance conditions. The core idea is to first extract the spatial dependencies of each node at every time step using a graph neural network, and then feed the resulting graph representations across consecutive time steps into a temporal model to capture the disturbance-driven dynamics of system states, as illustrated in Figure 4. In the temporal modeling module, a single-layer GRU encodes the state feature sequences of each node within the time window and outputs their latent dynamic representations under the current disturbance context. These representations serve as the inputs for predicting the future states of each node, followed by a linear projection layer to generate the final predicted values. The GRU hidden state dimension is set to 64, with the tanh activation function, and a dropout rate of 0.2 is applied during training to mitigate overfitting. The loss function is primarily based on mean squared error (MSE), augmented with disturbance-enhancement and state-smoothing terms, and is formulated as follows:

$$\mathcal{L} = \mathcal{L}_{\text{MSE}} + \lambda_1 \mathcal{L}_{\text{perturb}} + \lambda_2 \mathcal{L}_{\text{psmooth}} \tag{26}$$

Here, λ_1 and λ_2 are used to regulate the contributions of the regularization terms to the training objective. To improve numerical stability during training, all nodal features are standardized along the temporal dimension using Z-score normalization, where each state variable is independently normalized based on the mean and standard deviation of the training set. This prevents gradient imbalance caused by differences in feature scales. During the testing phase, the normalization parameters are kept fixed to ensure consistency of the input distribution. In addition, to account for real-world issues such as missing measurements and sampling interruptions, a feature masking mechanism is incorporated during training. In each training batch, certain feature dimensions of selected nodes are randomly masked to enhance the robustness and fault tolerance of the model. After generation by the sliding-window approach, data samples are randomly shuffled and partitioned into training, validation, and test sets with a ratio of 7:2:1. During this process, the distribution of the three disturbance types (gradual, abrupt, and load) is kept balanced across all subsets to ensure stable generalization of the model under different disturbance scenarios.

The disturbance-driven raw system operation data are thus organized into structured and standardized graph-based time-series samples, providing a high-quality data foundation for model training. Model training is performed using the Adam optimizer with a batch size of 64. To mitigate the adverse effects of error accumulation in state evolution modeling, all prediction tasks are conducted in a rolling single-step fashion, where multi-step-ahead forecasts are obtained through consecutive recursive single-step predictions to ensure stability. The algorithms are implemented using the PyTorch Geometric (PyG)2.3.0 and PyTorch 1.12.1 frameworks. All experiments were repeated multiple times to ensure the stability and reliability of the results, and the subsequent results represent the averaged outcomes across these repeated runs.

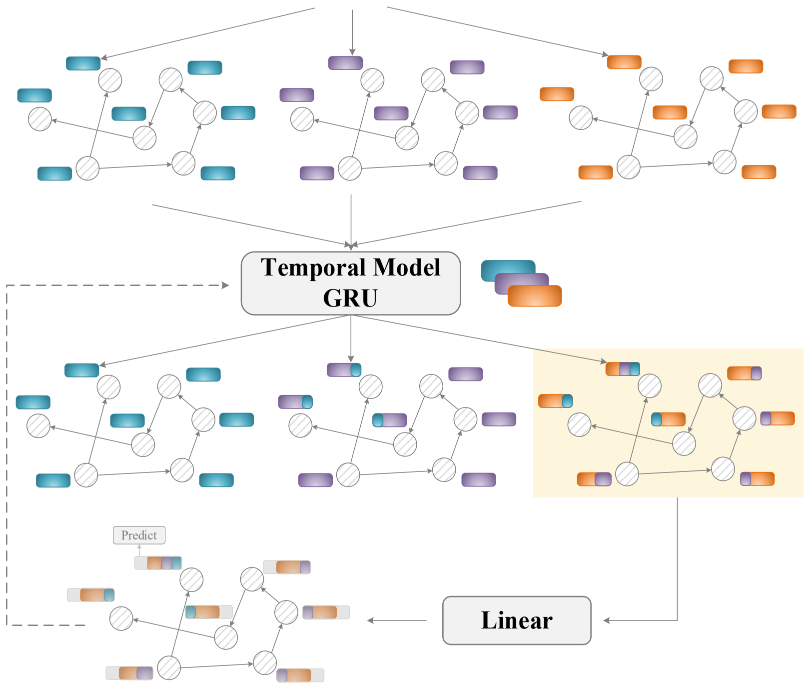


Figure 4. Temporal state modeling framework based on GRU.

4.4. Results Analysis

To systematically evaluate the performance of the proposed TDWGNN under disturbance scenarios, a comprehensive experimental scheme is designed, incorporating multi-model comparisons, metric-based evaluation, and diverse disturbance coverage. By adopting a unified data source, identical training procedures, and consistent evaluation metrics, fairness in model comparison and reliability in generalization assessment are ensured. Within the same framework, GCN, GCN + GRU, and GAT+GCN are selected as baseline models, with mean absolute error (MAE) and root mean square error (RMSE) used as evaluation metrics. Figure 5 illustrates the variation of the loss function for each model during the training phase. It can be observed that all four models converge rapidly in the early stages of training and reach relative stability after approximately 20 epochs. The GCN model exhibits large training fluctuations and relatively poor prediction performance. Incorporating temporal modeling (GCN + GRU) and graph attention (GAT + GCN) improves the results to some extent, yet their stability and final loss remain inferior. In contrast, TDWGNN demonstrates superior fitting ability and stronger resistance to overfitting, with the loss consistently maintained at a lower level.

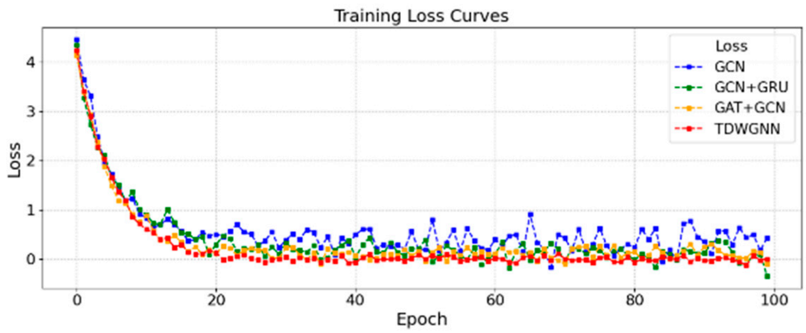


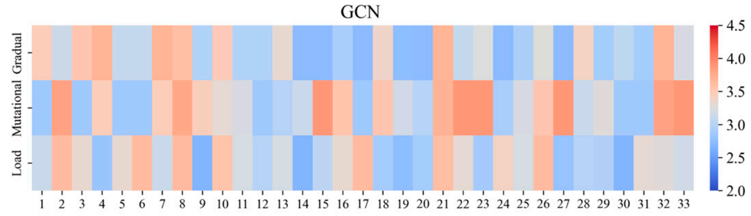
Figure 5. Loss Function Curves.

Table 1 reports the error performance of the four models under three representative scenarios: gradual disturbances, abrupt disturbances, and load disturbances. The results show that the conventional GCN exhibits relatively high errors across all three scenarios. With the introduction of temporal modeling (GCN + GRU), both MAE and RMSE decrease, indicating the positive role of temporal information in state prediction. Incorporating the graph attention mechanism (GAT + GCN) enhances the model's ability to capture nodal heterogeneity, leading to further error reduction. The proposed TDWGNN, which integrates graph structure, disturbance information, and temporal dynamics, achieves the best performance across all disturbance types and demonstrates superior generalization capability.

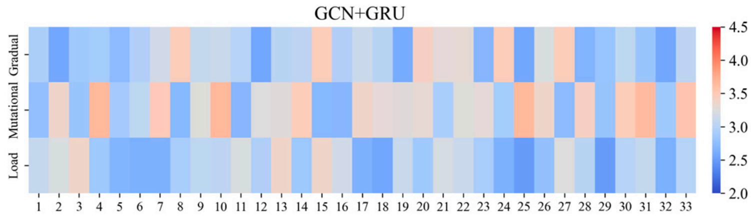
Table 1. Comparison of prediction errors of different models under three disturbance scenarios.

Algorithm	Evaluation Metric	Disturbance		
		Gradual	Abrupt	Load
GCN	MAE	3.41	3.57	3.35
	RMSE	4.22	4.31	4.10
GCN + GRU	MAE	3.02	3.18	2.94
	RMSE	3.77	3.83	3.65
GAT + GCN	MAE	2.78	2.96	2.73
	RMSE	3.42	3.54	3.38
TDWGNN	MAE	2.27	2.42	2.12
	RMSE	3.03	3.18	2.97

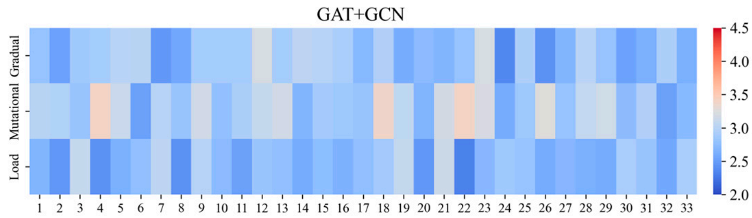
To further reveal model performance differences at the nodal level, Figure 6 presents the node-wise MAE heatmaps of different models under the three disturbance scenarios. The horizontal axis denotes node indices, while the vertical axis corresponds to disturbance types. It can be observed that the GCN exhibits significant error fluctuations across multiple nodes, particularly under abrupt disturbances. In contrast, TDWGNN yields smoother error distributions with smaller variations, demonstrating more stable cross-node modeling capability.



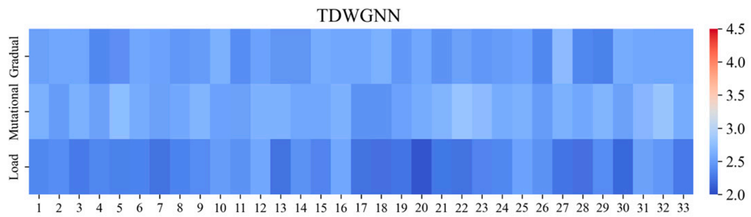
(**a**)



(**b**)



(**c**)



(**d**)

Figure 6. Node-wise MAE heatmaps under three disturbance scenarios. (**a**) GCN (**b**) GCN + GRU (**c**) GAT + GCN (**d**) TDWGNN.

4.5. Risk Assessment Levels

To visualize the model's capability in risk identification under renewable disturbances, the comprehensive risk assessment framework defined in Section 3 is applied to compute nodal risk values over the prediction horizon, which are then categorized into corresponding risk levels. The risk levels are classified into four categories: Level I (safe), Level II (mild), Level III (moderate), and Level IV (high risk), corresponding to risk intervals of [0, 0.25), [0.25, 0.5), [0.5, 0.75), and [0.75, 1.0], respectively. Taking the TDWGNN model under the abrupt disturbance scenario as an example, Table 2 reports the average risk levels and corresponding comprehensive risk values of each node over four consecutive time steps after the disturbance. The values in parentheses represent the contribution ratios of the three factors—voltage, power, and phase angle—denoted as $c^v = w^v R^v / R$, $c^P = w^P R^P / R$, $c^\theta = w^\theta R^\theta / R$, and $c^v + c^P + c^\theta = 1$. These values are derived from the comprehensive risk formulation in Section 3. D, where the normalized contributions of individual risk indicators and their weights are used.

Table 2. Node risk values, factor contributions, and levels under disturbance scenario.

Node	R_{total}	(R^v, R^P, R^θ)	Level	Node	R_{total}	(R^v, R^P, R^θ)	Level	Node	R_{total}	(R^v, R^P, R^θ)	Level
1	0.12	(0.63,0.25,0.12)	I	12	0.37	(0.24,0.51,0.25)	II	23	0.02	(0.17,0.09,0.74)	I
2	0.07	(0.14,0.11,0.75)	I	13	0.74	(0.49,0.25,0.26)	III	24	0.21	(0.08,0.67,0.25)	I
3	0.07	(0.09,0.57,0.34)	I	14	0.56	(0.40,0.20,0.40)	III	25	0.83	(0.33,0.41,0.26)	IV
4	0.01	(0.62,0.30,0.08)	I	15	0.42	(0.16,0.49,0.35)	II	26	0.79	(0.38,0.28,0.34)	IV
5	0.15	(0.20,0.13,0.67)	I	16	0.19	(0.63,0.21,0.16)	I	27	0.76	(0.31,0.52,0.17)	IV
6	0.62	(0.24,0.38,0.38)	III	17	0.06	(0.25,0.17,0.58)	I	28	0.89	(0.32,0.39,0.29)	IV
7	0.51	(0.39,0.40,0.21)	III	18	0.18	(0.12,0.64,0.24)	I	29	0.67	(0.36,0.32,0.32)	III
8	0.32	(0.45,0.24,0.31)	II	19	0.59	(0.41,0.36,0.23)	III	30	0.50	(0.30,0.35,0.35)	II
9	0.48	(0.17,0.55,0.28)	II	20	0.91	(0.36,0.37,0.27)	IV	31	0.63	(0.33,0.39,0.28)	III
10	0.31	(0.41,0.40,0.19)	II	21	0.88	(0.31,0.39,0.30)	IV	32	0.81	(0.26,0.36,0.38)	IV
11	0.29	(0.35,0.26,0.39)	II	22	0.88	(0.39,0.30,0.31)	IV	33	0.91	(0.30,0.35,0.35)	IV

The table shows that different nodes exhibit pronounced risk heterogeneity under spatiotemporal disturbance propagation. A combined analysis of network topology and renewable integration locations indicates that nodes located near disturbance sources or at the network boundary—such as nodes 25–28, 32–33, and 20–22—are subject to higher dynamic response stress. These nodes generally display higher average risk levels and comprehensive risk values, mostly within Levels III (moderate) and IV (high risk). In contrast, nodes situated in the central backbone or supported by multiple neighbors (nodes 1–5 and 16–18) experience lower overall risk, predominantly falling within Levels I and II.

It is noteworthy that nodal risk levels are not solely determined by topological location but are also influenced by factors such as disturbance propagation direction, temporal accumulation effects, and the model's ability to capture historical states. For instance, node 22 is not located closest to the disturbance source, yet its connections to PV2 and multiple branches significantly amplify its risk response when disturbances accumulate. The above analysis validates the feasibility and practicality of the comprehensive risk assessment method in spatiotemporal prediction scenarios. On the one hand, it enables further quantification of system operational security beyond prediction accuracy; on the other hand, it offers node-level references for the formulation of risk control strategies and scheduling, thereby supporting the development of secure and resilient operational mechanisms for next-generation local power grids.

5. Conclusions

This paper proposes a disturbance-driven spatiotemporal graph prediction model (TDWGNN) that integrates the structural dependency modeling of graph neural networks with the temporal dynamics of gated recurrent units, enabling accurate prediction of system state evolution under diverse disturbance scenarios. Experimental results demonstrate that TDWGNN consistently outperforms existing methods in terms of prediction accuracy, convergence rate, and generalization capability. A comprehensive risk assessment method is further introduced, which incorporates nodal topological features to quantitatively identify high-risk regions, thereby providing interpretable and practically relevant guidance for secure grid operation. By coupling multi-horizon prediction with node-level risk quantification, the proposed approach offers a more actionable perspective for real-time operational awareness in renewable-rich distribution networks. The study overcomes the limitations of static approaches in capturing disturbance propagation and establishes an integrated framework for prediction, assessment, and early warning, delivering both methodological innovation and practical engineering value for resilient and secure operation of future power systems under high uncertainty.

Despite these strengths, the present study has certain limitations. The experimental evaluation is conducted on a single benchmark system and therefore does not yet capture the wider diversity found in real distribution networks. In addition, the risk assessment module adopts a relatively simple linear formulation to ensure real-time applicability and interpretability, leaving room for further extension when addressing more complex operating environments. Future work will extend the model to larger-scale or real-world distribution networks as data availability permits, and will explore more flexible risk modeling approaches and explainability techniques to further enhance its practical applicability across diverse operational scenarios. In parallel, developing lightweight or edge-deployable variants of TDWGNN may facilitate integration into real-time monitoring platforms, enabling fast, interpretable, and disturbance-aware analytics for next-generation digital distribution grids.

Author Contributions: Conceptualization, J.Z. and K.L.; Funding acquisition, Z.S.; Methodology, J.Z., Y.C., Z.Z., L.Z., S.B., P.W. and K.L.; Supervision, Z.S. and S.B.; Validation, J.Z.; Writing—original draft, J.Z.; Writing—review and editing, Z.S., Y.C., Z.Z., L.Z., S.B., P.W. and K.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Research Project from Southern Power Grid Digital Grid Research Institute (8516009544).

Data Availability Statement: The data are contained within the article.

Conflicts of Interest: Authors Zhixin Suo, Yukai Chen, Zihao Zhang, Liang Zhao, Shanshan Bai and Pengyu Wang were employed by the company Southern Power Grid Digital Grid Research Institute Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Hu, Z.Y.; Xu, Z.; Sun, X.Z.; Ruan, J.Q.; Yang, X.Y.; Qian, T.; Shi, W.Z. Meteorological-Electrical Integrated Real-Time Resilience Assessment for Power Systems Based on Deep Learning Methods. *IEEE Trans. Smart Grid* **2025**, *16*, 3700–3713. [CrossRef]
2. Li, S.J.; Hu, R.C.; Chen, G.L.; Chen, L.L.; Li, H.; Jiang, H.G.; Xue, Y.; Kang, J.W.; Zhang, J.; Gao, W.Z. Efficient Net Load Forecasting in Large-scale Power Distribution Systems via Dual-branch Experts Fusion Memory Network. *IEEE Trans. Power Syst.* **2025**, 1–12, *Early Access*. [CrossRef]
3. Jiang, Y.Z. Data-Driven Probabilistic Fault Location of Electric Power Distribution Systems Incorporating Data Uncertainties. *IEEE Trans. Smart Grid* **2021**, *12*, 4522–4534. [CrossRef]
4. Zadeh, S.B.I. Environmental Benefits of Reducing Greenhouse Gas Emissions from Smart Ports via Implementation of Smart Energy Infrastructure. *GMSARN Int. J.* **2024**, *18*, 431–439.
5. Zadeh, S.B.I.; Soltani, H.R.; Ghoneim, N.I. Revamping Seaport Operations with Renewable Energy: A Sustainable Approach to Reducing Carbon Footprint. *GMSARN Int. J.* **2024**, *18*, 315–324.
6. Ye, Y.J.; Wu, Y.Z.; Hu, J.X.; Hu, H.; Qian, S.Q.; Zhang, X.; Wang, Q.; Strbac, G. Physics-Guided Safe Policy Learning with Enhanced Perception for Real-Time Dynamic Security Constrained Optimal Power Flow. *J. Mod. Power Syst. Clean Energy* **2025**, *13*, 1507–1519. [CrossRef]
7. Zhao, T.Q.; Yue, M.; Wang, J.H. Structure-Informed Graph Learning of Networked Dependencies for Online Prediction of Power System Transient Dynamics. *IEEE Trans. Power Syst.* **2022**, *37*, 4885–4895. [CrossRef]
8. Haghshenas, S.H.; Naeini, M. Resilient Temporal Graph Convolutional Network for Smart Grid State Estimation Under Topology Inaccuracies. *IEEE Open Access J. Power Energy* **2025**, *12*, 529–540. [CrossRef]
9. Dolatyabi, P.; Khodayar, M. Graph Neural Networks and Their Applications in Power Systems: A Review. In Proceedings of the 2025 IEEE International Conference on Electro Information Technology (eIT), Valparaiso, IN, USA, 29–31 May 2025.
10. Nandanoori, S.P.; Guan, S.; Kundu, S.; Pal, S.; Agarwal, K.; Wu, Y.H.; Choudhury, S. Graph Neural Network and Koopman Models for Learning Networked Dynamics: A Comparative Study on Power Grid Transients Prediction. *IEEE Access* **2022**, *10*, 32337–32349. [CrossRef]
11. Ahmed, A.; Basumallik, S.; Gholami, A.; Sadanandan, S.K.; Namaki, M.H.N.; Srivastava, A.K.; Wu, Y.H. Spatio-Temporal Deep Graph Network for Event Detection, Localization, and Classification in Cyber-Physical Electric Distribution System. *IEEE Trans. Ind. Inform.* **2024**, *20*, 2397–2407. [CrossRef]

12. Wang, T.; Zhu, L.P.; Wen, W.J.; Hou, J. Interpretable Transient Frequency Stability Prediction Based on Graph Convolutional Learning. In Proceedings of the 2024 IEEE 8th Conference on Energy Internet and Energy System Integration (EI2), Shenyang, China, 29 November–2 December 2024.
13. Feng, K.; Zhou, H.; Qin, W.; Han, S.; Rong, N.; Du, D.; He, Y.; Yang, C. A Frequency Stability Prediction Method Using Spatial Temporal Graph Convolutional Networks and Self-attention Mechanism. In Proceedings of the 2024 6th International Conference on Electrical Engineering and Control Technologies (CEECT), Shenzhen, China, 20–22 December 2024.
14. Zhou, Z.; Li, H.; Zheng, J.; Xiang, X.; Yao, R.; Tan, H. Power System Frequency Response Prediction with Spatial-Temporal Graph Convolutional Networks. In Proceedings of the 2025 IEEE International Conference on Electrical Energy Conversion Systems and Control (IEECSC), Chongqing, China, 23–25 May 2025.
15. Hu, J.; Hu, W.; Chen, J.; Cao, D.; Zhang, Z.; Liu, Z.; Chen, Z.; Blaabjerg, F. Fault Location and Classification for Distribution Systems Based on Deep Graph Learning Methods. *J. Mod. Power Syst. Clean Energy* **2023**, *11*, 35–51. [CrossRef]
16. Nguyen, B.L.H.; Vu, T.V.; Nguyen, T.T.; Panwar, M.; Hovsapian, R. Spatial-Temporal Recurrent Graph Neural Networks for Fault Diagnostics in Power Distribution Systems. *IEEE Access* **2023**, *11*, 46039–46050. [CrossRef]
17. Mahto, D.K.; Saini, V.K.; Mathur, A.; Kumar, R.; Yadav, R. GAT- DNet: High Fidelity Graph Attention Network for Distribution Optimal Power Flow Pursuit. In Proceedings of the 2023 9th IEEE India International Conference on Power Electronics (IICPE), Sonipat, India, 28–30 November 2023.
18. Mahto, D.K.; Bukya, M.; Kumar, R.; Mathur, A.; Saini, V.K. GAT-ADNet: Leveraging Graph Attention Network for Optimal Power Flow in Active Distribution Network With High Renewables. *IEEE Access* **2024**, *12*, 185728–185739. [CrossRef]