energies

MDPI

Article

Data–Physics Fusion-Driven Dynamic Partitioning of Active Distribution Networks for Fast Coordinated Power Control

Zhi Zhou [1], Siyang He [1], Rui He [1], Quanhai Yang [1], Zhenglin Zhong [1], Yubin Liu [2,*], Tao Yu [2] and Zixi Mo [2]

[1] Duyun Power Supply Bureau, Guizhou Power Grid Co., Ltd., Duyun 558000, China; 18786382520@163.com (Z.Z.); hll6@163.com (S.H.); 18084230327@163.com (R.H.); alettattt@163.com (Q.Y.); 15121392855@163.com (Z.Z.)

[2] School of Electric Power Engineering, South China University of Technology, Guangzhou 510641, China; taoyu1@scut.edu.cn (T.Y.); zixikenmo@163.com (Z.M.)

* Correspondence: 202510183241@mail.scut.edu.cn

Abstract

High penetrations of distributed energy resources make active distribution networks strongly time-varying, nonlinear, and spatially coupled, which limits the online applicability of centralized voltage/reactive-power optimization. This paper proposes a data–physics fusion dynamic partitioning method for fast power coordination. A physics-based rolling partition baseline is first developed by integrating node operating behavior, voltage/reactive sensitivity, electrical distance, and feeder topology, providing an interpretable and efficient partitioning scheme for normal operating conditions. For high-volatility and strongly coupled scenarios, a heterogeneous dynamic graph and a heterogeneous spatio-temporal graph attention network are introduced to learn control-oriented latent node embeddings. Physical regularization, boundary-coupling penalties, and temporal smoothing constraints are further embedded into soft clustering to reduce cross-partition coupling and partition fluctuation. Tests on the IEEE 33-bus, IEEE 123-bus, and practical Feeder Z systems show that the dynamic partition closely approximates global OPF results, achieving normalized costs of 1.00017 and 1.00099 on the two IEEE systems with 74.3% and 83.2% time reductions. It further reduces the Feeder Z fixed-partition cost gap by 88.0%, while HST-GAT lowers boundary P/Q exchanges by 1.55%/6.57% under volatile conditions.

Keywords: active distribution network; dynamic partitioning; data–physics fusion; HST-GAT; power coordination; VVC-OPF

Check for updates

Academic Editor: Andrea Mariscotti

Received: 25 May 2026
Revised: 19 June 2026
Accepted: 26 June 2026
Published: 29 June 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

With the large-scale integration of distributed photovoltaics, wind generation, energy storage, and flexible loads, distribution networks are evolving from passive one-way supply systems into active distribution networks (ADNs) with deeply coupled source–network–load–storage interactions [1–4], with recent studies further highlighting the economic and operational challenges of high-penetration renewable distribution networks [5]. Compared with traditional distribution networks, ADNs are more strongly affected by renewable output fluctuations, stochastic load changes, and network reconfiguration [6]. Node voltages, power-flow distributions, and local voltage-control requirements therefore show significant time variation, nonlinearity, and spatial coupling [7–9]. Under these conditions, using a detailed global model directly for online analysis and coordinated control often leads to high computational dimension, insufficient real-time performance, and difficult model

maintenance [10]. Therefore, constructing control partitions that can evolve with operating conditions has become a fundamental problem in online optimization and hierarchical control of ADNs.

This work does not aim at improving clustering accuracy in a generic machine-learning sense. Instead, it investigates what kind of dynamic partition best supports partition-level voltage/reactive-power coordination and voltage-constrained Volt/Var-control OPF (VVC-OPF) solving, where VVC-OPF is based on linearized voltage/reactive sensitivity relations, reactive-power regulation limits, and voltage-deviation/violation penalties. From the viewpoint of control decomposition, a useful partition should satisfy three requirements. First, nodes within the same region should exhibit strong voltage/reactive coupling consistency so that the partition carries clear physical meaning. Second, the partition should evolve smoothly over time so that frequent boundary changes do not disrupt hierarchical dispatch. Third, inter-region coupling should be weak, and controllable resources should be adequate within each region so that the partitioned optimization can approximate the centralized solution with bounded coordination error.

Existing control-oriented partitioning and graph-partitioning studies provide important foundations, but they do not fully satisfy the above requirements. Electrical distance, voltage/reactive-power sensitivity, modularity-based, and spectral-clustering methods are widely used to identify electrically coherent regions for voltage control and hierarchical dispatch [11–16]. For example, reference [11] constructs network partitions according to computational complexity, power balance, and electrical distance for distributed reactive-power optimization; reference [12] develops a combined central-local voltage-control strategy with DRO-based tuning of DG local control curves; reference [13] proposes a cluster-optimized regulation method considering heterogeneous DPV control modes; and reference [14] extracts representative PV scenarios through graph clustering with shared weights and adaptive fusion. These studies are physically meaningful and useful for voltage regulation, renewable uncertainty modeling, or distributed optimization. However, most of them are derived from fixed or representative operating points, or are designed for scenario reduction and device-level regulation. Therefore, their partition definitions may become suboptimal when PV output, load distribution, and local topology vary rapidly.

Graph-cut and controlled-islanding methods use related graph-partitioning tools but pursue a different objective. They determine separation boundaries using generator coherency, real-time WAMS criteria, clustering algorithms, and minimum electrical distance under severe disturbances [17,18]. These methods provide useful insights into electrical-distance metrics and post-split feasibility checking. However, their main purpose is emergency islanding and frequency/transient-stability preservation, rather than constructing daily control regions for fast voltage/reactive-power coordination under normal and high-volatility ADN operating conditions.

Hierarchical and multi-time-scale voltage/var control methods further demonstrate the value of regional decomposition. Decentralized voltage management methods use network sensitivity matrices and regional control centers [19]; hybrid time-scale methods incorporate PV uncertainty into partition design [20]; and hierarchical multi-stage schemes coordinate large-scale EV charging with volt/var control [21]. These studies show that partition-based VVC can reduce computational burden and improve regulation efficiency. Nevertheless, many of them rely on predefined or slowly updated regions, or focus on a specific resource class. They do not fully learn dynamic partitions from operating behavior, mechanism-response, and topology features in a unified framework.

Data-driven clustering and graph-learning methods offer stronger representation capability for complex power-system data. Data-driven clustering and aggregation methods can capture device, load, or scenario similarity in distribution networks [22–25], but without

physical constraints the resulting clusters may be statistically similar while electrically unreasonable. Graph-learning-based methods can further exploit non-Euclidean topology information. For instance, reference [26] proposes a GAT-SAC-based fault-restoration method for ADNs, showing that graph attention can improve topology perception and generalization under network changes. Reference [27] studies dynamic network partitioning for large-scale microgrid clusters and confirms that adaptive region updating is important when operating modes vary over time. Reference [28] reviews GNN applications in power systems, including fault diagnosis, time-series prediction, power-flow calculation, and data generation. Reference [29] introduces a GAT-based multi-agent reinforcement learning method with a physical-assisted mechanism for voltage regulation under topology variations. Reference [30] embeds topology, electrical characteristics, source-load distribution, and reactive-voltage sensitivity into a graph-attention-based multi-agent RL framework for VVC. These studies improve topology perception, state representation, adaptive region updating, and policy learning. However, their main focus remains restoration, switching optimization, state perception, microgrid-level region management, or control-policy generation, rather than interpretable dynamic control-region construction linked with physical correction and partition-level VVC-OPF validation. Table 1 provides a structured comparison.

Table 1. Comparison between the proposed method and existing partitioning-related methods for ADNs.

Method Category	Representative Studies	Main Principle	Main Limitation for Fast Coordinated Control	Difference of This Work
Electrical-distance/sensitivity-based clustering	[11–14]	Uses electrical distance, voltage sensitivity, or electrical modularity to form electrically coherent regions	Usually based on fixed or representative operating points; limited adaptability to renewable-driven temporal variation	Uses rolling operating features, V–Q sensitivity, electrical distance, and topology jointly
Spectral-clustering-based partitioning	[16]	Builds graph similarity matrix and extracts partition labels through spectral embedding and clustering	Often optimizes graph similarity, edge-computing, or communication objectives rather than OPF decomposition quality	Adds physical correction, temporal smoothing, controllable-resource checking, and OPF validation
Source–load–storage cluster scheduling	[15]	Divides clusters according to electrical modularity and active-power balance, then optimizes cluster scheduling	Emphasizes economic scheduling and power balance; partition update and OPF decomposition error are not the core focus	Constructs control-oriented partitions and evaluates them through partition-level OPF
Graph-cut/controlled islanding	[17,18]	Determines boundary cuts according to coherency, electrical distance, WAMS criteria, or stability constraints	Mainly designed for emergency islanding and transient/frequency stability, not routine ADN voltage/reactive-power coordination	Focuses on dynamic control regions for normal and high-volatility operating scenarios
Hierarchical/multi-time-scale VVC	[19–21]	Uses regional control centers, time-scale decomposition, and local voltage/var control	Often relies on predefined or slowly updated regions and specific resource models	Learns dynamic partitions from operating, mechanism-response, and topology features

Table 1. *Cont.*

Method Category	Representative Studies	Main Principle	Main Limitation for Fast Coordinated Control	Difference of This Work
GNN/GAT-based ADN methods	[26–30]	Uses graph learning to enhance topology perception and decision-making	Mainly used for restoration, state perception, or control-policy learning	Uses HST-GAT as a high-volatility enhancement for dynamic partitioning with physical constraints
Proposed method	This work	Data–physics fusion rolling dynamic partitioning with HST-GAT enhancement	Additional parameter and implementation complexity	Closes the loop among feature fusion, dynamic clustering, physical correction, resource adequacy, temporal smoothness, and OPF validation

The remaining research gap is threefold. First, existing electrical-distance, sensitivity-based, modularity-based, and spectral-clustering methods provide useful physical or graph-theoretic partitioning criteria, but they are often derived from fixed or representative operating points and may not remain suitable when renewable generation, load demand, and local topology vary rapidly. Second, graph-cut and controlled-islanding methods provide effective tools for emergency network separation, but their objectives are mainly frequency stability, transient stability, and post-islanding feasibility, rather than routine partition-level voltage/reactive-power coordination in ADNs. Third, hierarchical voltage/var control and graph-learning-based methods have improved regional control and topology perception, but many studies still rely on predefined regions or focus on control-policy learning, state perception, or restoration decisions. Therefore, existing studies have not sufficiently closed the loop among dynamic clustering, physical correction, controllable-resource adequacy, temporal smoothness, and OPF-based validation. In other words, it is insufficient to show that nodes can be clustered; the resulting partition must demonstrably reduce control-relevant boundary coupling and support fast and accurate coordinated voltage/reactive-power control.

To address these limitations, this study proposes a data–physics fusion-driven dynamic partitioning method for ADNs. Unlike generic clustering methods that mainly pursue topological compactness, electrical similarity, communication efficiency, or latent-space representation accuracy, the proposed method is designed to construct control-oriented dynamic partitions for fast partition-level voltage/reactive-power coordination. Here, dynamic clustering refers to the algorithmic process of latent representation learning and time-varying assignment updating, whereas dynamic partitioning refers to the engineering partition used for ADN operation. The two are connected through physical correction and partition-level OPF validation. The main contributions are as follows.

1. A control-oriented multi-source feature system is constructed by jointly organizing operating behavior, mechanism response, and network structure. Compared with conventional electrical-distance clustering, sensitivity-based clustering, spectral clustering, and modularity-based partitioning, the proposed physics-based rolling partition baseline simultaneously considers intra-region electrical consistency, boundary coupling strength, controllable-resource adequacy, and temporal partition smoothness. Therefore, it produces interpretable dynamic control regions rather than static similarity clusters.

2. An HST-GAT enhancement module is developed for high-volatility and strongly coupled scenarios. Different from existing GNN/GAT-based ADN methods that mainly focus on fault restoration, topology perception, surrogate modeling, or control-policy learning, the proposed module learns control-oriented latent node embeddings for dynamic partitioning. It constructs relation-specific subgraphs from topology, voltage/reactive sensitivity, and electrical distance, and integrates physical consistency, boundary decoupling, controllable-resource adequacy, and temporal smoothing into the clustering loss so that the learned partitions remain electrically interpretable and implementable.
3. A partition-level voltage/reactive-power coordination and OPF validation framework is established on three systems of increasing complexity: a modified IEEE 33-bus system, the IEEE 123-bus system, and a practical Feeder Z system with non-consecutive node numbering and switch status records. Unlike studies that evaluate partitioning only through clustering compactness, graph metrics, communication efficiency, or scheduling cost, this work evaluates whether the resulting partitions reduce control-decomposition error, computational burden, boundary coupling, and voltage-violation risk. The results clarify both the benefit and the applicability boundary of HST-GAT relative to the physics-based rolling partition baseline.

The rest of this paper is organized as follows. Section 2 formulates the multi-feature dynamic partitioning framework for ADNs. Section 3 presents the HST-GAT based data–physics fusion model. Section 4 describes the experimental systems, data sources, and scenarios. Section 5 reports and discusses the results. Section 6 concludes the paper.

2. Multi-Feature Fusion Dynamic Partitioning Framework for ADNs

Figure 1 summarizes the proposed data–physics fusion dynamic partitioning framework. The workflow starts from three groups of input information: operating behavior, mechanism response, and network structure. These data are transformed into multi-source features, including voltage and power fluctuation features, voltage/reactive sensitivity, electrical distance, and feeder-topology relations. The partitioning engine then applies either the physics-based rolling partition baseline or the HST-GAT enhanced representation module. The obtained partitions are further corrected through connectivity repair, boundary-coupling checking, temporal smoothing, and resource-adequacy checking before being used for partition-level VVC-OPF validation.

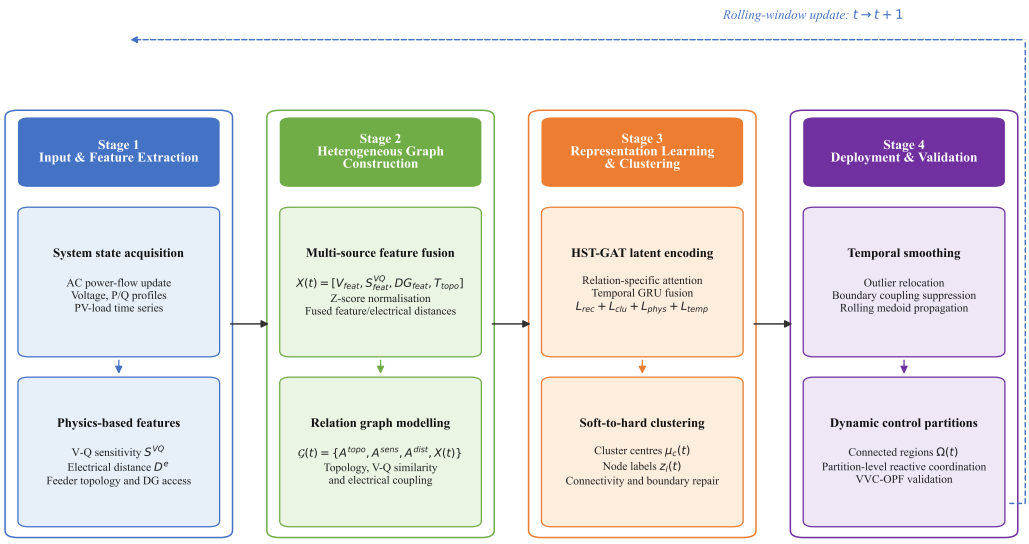

Figure 1. Data–physics fusion dynamic partitioning framework for fast coordinated control.

2.1. Multi-Dimensional Feature Matrix Extraction

The proposed partitioning method is built on a multi-source feature system. The features are not simply stacked. Instead, they are organized according to three roles: operating behavior, mechanism response, and network structure.

Operating behavior features describe the dynamic response of each node under renewable and load fluctuations. For node i in time window t, the voltage series, active-power injection, and reactive-power injection are denoted as $V_i(\tau)$, $P_i(\tau)$, and $Q_i(\tau)$, respectively, where τ denotes the sampling index within the rolling window. Typical statistical features include the voltage mean, standard deviation, maximum deviation, ramping rate, and fluctuation energy:

$$x_i^{\mathrm{op}}(t) = \left[\bar{V}_i, \sigma(V_i), \max_{\tau} |V_i(\tau) - 1|, \sigma(P_i), \sigma(Q_i), r_i^P, r_i^Q\right]. \tag{1}$$

In Equation (1), $\bar{V}_i$ is the mean voltage magnitude of node i within the rolling window, $\sigma(V_i)$ is the voltage standard deviation, $\max_\tau |V_i(\tau) - 1|$ is the maximum voltage deviation from 1.0 p.u., $\sigma(P_i)$ and $\sigma(Q_i)$ are the active- and reactive-power fluctuation levels, and r_i^P and r_i^Q denote the active- and reactive-power ramping rates, respectively.

For renewable-rich nodes, the output fluctuation index is further calculated as

$$\phi_i(t) = \frac{1}{T-1} \sum_{\tau=2}^{T} |P_i(\tau) - P_i(\tau - 1)|, \tag{2}$$

where T is the length of the rolling window. The sampling interval Δt is fixed according to the system measurement resolution (e.g., 15 min sampling). Each feature is constructed from historical measurements within the window $[t - T + 1, t]$, enabling the model to capture short-term temporal variations of renewable generation and load demand. The output fluctuation index $\phi_i(t)$ is introduced to capture short-term renewable or DG output variations, which more directly reflects rapid ramping behavior caused by PV intermittency or abrupt load changes than mean and variance features alone. Voltage power-quality related features are also used to characterize voltage fluctuation and violation risk:

$$\begin{aligned} \psi_i(t) = \big[&\sigma(V_i), \max_{\tau} |V_i(\tau) - 1|, N_i^{\mathrm{viol}}, \\ &\sum_{\tau} \max(0, V_i(\tau) - V^{\max}) + \sum_{\tau} \max(0, V^{\min} - V_i(\tau))\big]. \end{aligned} \tag{3}$$

In Equation (3), N_i^{viol} denotes the number of samples in which the voltage of node i violates the allowable range, $V^{\max}$ and $V^{\min}$ are the upper and lower voltage limits, respectively, and the last two summation terms represent the cumulative upper- and lower-voltage violation magnitudes within the rolling window.

Before feature fusion, all heterogeneous features are normalized to eliminate scale differences among voltage, power, sensitivity, and electrical-distance variables. For each feature dimension, a standard z-score normalization is applied:

$$\tilde{x} = \frac{x - \mu}{\sigma + \epsilon}, \tag{4}$$

where μ and σ are the mean and standard deviation of the feature over the available samples, and ϵ is a small constant to avoid numerical instability. This preprocessing ensures that no single feature group dominates the clustering process due to magnitude differences.

Mechanism-response features are derived from power-flow equations. Around an operating point, the linearized relationship between nodal injection variation and voltage variation can be written as

$$\begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} = \mathbf{J} \begin{bmatrix} \Delta \theta \\ \Delta V \end{bmatrix}, \tag{5}$$

where $\mathbf{J}$ is the Jacobian matrix of the power-flow equations. The voltage/reactive sensitivity matrix is extracted as

$$\mathbf{S}^{VQ} = \frac{\partial V}{\partial Q}, \tag{6}$$

and is used to measure the control-relevant coupling between nodes.

The mechanism-response features are derived from the linearized power-flow model around the current operating point. The Jacobian matrix captures the first-order sensitivity structure of the network, and the V–Q sensitivity submatrix $\mathbf{S}^{VQ}$ quantifies the physical coupling strength between bus reactive-power injections and voltage magnitudes.

To avoid scale ambiguity between sensitivity and impedance quantities, the electrical distance is constructed from two independently normalized components. The sensitivity-based distance between nodes i and j is defined as

$$d_{ij}^{\text{sens}}(t) = \left\| \mathbf{S}_{i,:}^{VQ}(t) - \mathbf{S}_{j,:}^{VQ}(t) \right\|_2, \tag{7}$$

where $\mathbf{S}_{i,:}^{VQ}(t)$ and $\mathbf{S}_{j,:}^{VQ}(t)$ are the i-th and j-th rows of the voltage/reactive sensitivity matrix.

The impedance-based distance is defined from the reduced Thevenin impedance matrix $\mathbf{Z}_{\text{th}} = \mathbf{Y}_{\text{red}}^{-1}$, where $\mathbf{Y}_{\text{red}}$ is the bus admittance matrix with the slack bus removed. The impedance distance is

$$d_{ij}^{\text{imp}} = \left| Z_{\text{th},ii} + Z_{\text{th},jj} - Z_{\text{th},ij} - Z_{\text{th},ji} \right|, \tag{8}$$

which for symmetric admittance matrices reduces to $\left| Z_{\text{th},ii} + Z_{\text{th},jj} - 2Z_{\text{th},ij} \right|$. Both distance components are then normalized before fusion:

$$\bar{d}_{ij}^{\text{sens}}(t) = \frac{d_{ij}^{\text{sens}}(t)}{\max_{p<q} d_{pq}^{\text{sens}}(t) + \epsilon_d}, \quad \bar{d}_{ij}^{\text{imp}} = \frac{d_{ij}^{\text{imp}}}{\max_{p<q} d_{pq}^{\text{imp}} + \epsilon_d}, \tag{9}$$

where ϵ_d is a small positive constant to avoid division by zero. The final electrical distance is defined as the convex combination

$$d_{ij}^{e}(t) = (1 - \gamma_z)\bar{d}_{ij}^{\text{sens}}(t) + \gamma_z \bar{d}_{ij}^{\text{imp}}, \quad \gamma_z \in [0, 1]. \tag{10}$$

The coefficient γ_z controls the relative contribution of the impedance-based component. In this study, $\gamma_z = 0$ is used as the default value because the voltage/reactive sensitivity distance directly reflects operating-point-dependent control coupling and is already available from the voltage-control model. The optional impedance term is retained for completeness. A sensitivity analysis over γ_z is provided in Section 5.

Network-structure features are extracted from the feeder topology. Let $\mathbf{A}$ be the adjacency matrix and $\mathcal{E}$ be the branch set. Topological relation, sensitivity relation, and electrical-distance relation are jointly used to construct a multi-relation graph:

$$\mathcal{G}(t) = \left(\mathcal{V}, \mathcal{E}^{\text{top}}, \mathcal{E}^{\text{sen}}(t), \mathcal{E}^{\text{dist}}(t) \right). \tag{11}$$

where $\mathcal{V}$ denotes the node set, $\mathcal{E}^{\text{top}}$ denotes topology edges, $\mathcal{E}^{\text{sen}}$ denotes voltage/reactive sensitivity edges, and $\mathcal{E}^{\text{dist}}$ denotes electrical-distance edges.

The final feature vector is

$$x_i(t) = \left[x_i^{\text{op}}(t), x_i^{\text{mech}}(t), x_i^{\text{top}}(t)\right], \tag{12}$$

and the feature matrix for all nodes is denoted as

$$\mathbf{X}(t) = [x_1(t), x_2(t), \ldots, x_N(t)]^\top. \tag{13}$$

For missing or corrupted measurements, linear interpolation is applied when data gaps are short (no more than a few consecutive samples). For longer missing intervals, the last available valid observation is retained together with a binary mask indicating missingness. A median filtering step is applied before feature extraction to reduce high-frequency measurement noise.

To balance the contribution of heterogeneous feature groups, a group-wise weighting strategy is introduced. Operating behavior features, mechanism-response features, and topology-related features are first normalized separately within each group, and then concatenated using predefined weighting coefficients α_{op}, α_{mech}, and α_{top}. In the baseline configuration, equal weights are adopted ($\alpha_{\text{op}} = \alpha_{\text{mech}} = \alpha_{\text{top}} = 1$), and a sensitivity analysis on these weights is conducted in Section 5 to verify robustness against different weighting choices.

The above multi-source feature system is designed to jointly capture temporal variability, physical coupling, and network topology, enabling a control-oriented clustering process rather than purely statistical similarity grouping.

2.2. Dynamic Clustering Problem for Multi-Dimensional Features

Let the dynamic partition at time window t be

$$\Omega(t) = \{\Omega_1(t), \Omega_2(t), \ldots, \Omega_K(t)\}, \tag{14}$$

where K is the number of regions and $\Omega_c(t)$ is the node set of region c. The assignment variable is defined as

$$z_i(t) \in \{1, 2, \ldots, K\}. \tag{15}$$

A desirable control-oriented partition should satisfy compactness in the latent space, electrical consistency inside regions, weak boundary coupling, topology connectivity, resource adequacy, and temporal smoothness. Accordingly, the general objective is formulated as

$$\min_{\{z_i(t)\}} J(t) = \lambda_{\text{emb}} J_{\text{emb}}(t) + \lambda_e J_{\text{elec}}(t) + \lambda_b J_{\text{bound}}(t) + \lambda_t J_{\text{temp}}(t), \tag{16}$$

where J_{emb} is the embedding compactness term, J_{elec} penalizes large intra-region electrical distance, J_{bound} penalizes cross-region coupling, and J_{temp} suppresses unnecessary label changes. The partition granularity is controlled by the prescribed number of partitions K and the connectivity correction step, rather than by an additional independently tuned penalty coefficient.

The electrical consistency term is

$$J_{\text{elec}}(t) = \sum_{c=1}^{K} \frac{1}{|\Omega_c(t)|^2} \sum_{i,j \in \Omega_c(t)} d_{ij}^e(t). \tag{17}$$

where K is the number of partitions, $|\Omega_c(t)|$ is the number of nodes in partition c, and $d^e_{ij}(t)$ is the electrical distance defined in Section 2.1.

The boundary coupling term is

$$J_{\text{bound}}(t) = \sum_{(i,j)\in\mathcal{E}} \mathbb{I}[z_i(t) \neq z_j(t)] \Big(\omega_s |S^{VQ}_{ij}(t)| + \omega_p |P_{ij}(t)| + \omega_q |Q_{ij}(t)|\Big), \tag{18}$$

where ω_s, ω_p, and ω_q are weighting coefficients that balance the contributions of sensitivity-based coupling, active-power flow, and reactive-power flow to the boundary penalty, respectively.

Temporal smoothness is written as

$$J_{\text{temp}}(t) = \sum_{i=1}^{N} \rho_i(t) \mathbb{I}[z_i(t) \neq z_i(t-1)], \tag{19}$$

where $\rho_i(t)$ is a state-dependent switching penalty. It is large when the local operating state changes only slightly, and small when a real operating transition justifies repartitioning.

In summary, the dynamic clustering problem is cast as a multi-term constrained optimization that simultaneously accounts for latent-space compactness, intra-region electrical consistency, boundary coupling strength, and temporal partition smoothness. Unlike generic clustering that treats all features equally, this formulation explicitly encodes the control-oriented requirements identified earlier: the electrical and boundary terms ensure that the resulting partition carries clear physical meaning for reactive-power coordination, the temporal term suppresses unnecessary label fluctuation across rolling windows, and the partition granularity is enforced through the prescribed cluster number K together with the connectivity correction. The penalty coefficients $\lambda_{\text{emb}}, \lambda_e, \lambda_b, \lambda_t$ together with the state-dependent switching weight $\rho_i(t)$ provide tunable knobs that allow the partition to adapt to different operating regimes. The numerical values and tuning of all parameters are described in Section 4, and a sensitivity analysis is provided in Section 5.

2.3. Partition-Driven Reactive-Power Coordination Architecture

The dynamic partition is used to support partition-level reactive-power coordination. After partitioning, each region is treated as a control unit. Let Δq_i be the reactive-power adjustment of controllable resource i. A partition-level control problem can be written as

$$\min_{\Delta q} \sum_{i\in\mathcal{C}} c_i(\Delta q_i)^2 + \eta_v \sum_{i\in\mathcal{V}} (V_i - 1)^2 + \eta_b J_{\text{bound}}, \tag{20}$$

subject to

$$V^{\min} \leq V_i \leq V^{\max}, \quad \forall i, \tag{21}$$

$$\underline{q}_i \leq \Delta q_i \leq \overline{q}_i, \quad \forall i \in \mathcal{C}, \tag{22}$$

$$\Delta V = \mathbf{S}^{VQ} \Delta Q. \tag{23}$$

Here, $\mathcal{C}$ is the set of controllable resources. The global VVC-OPF solves the full system directly. The fixed-partition VVC-OPF uses unchanged regions across all time windows. The dynamic-partition VVC-OPF updates regions according to the current operating state. The purpose of clustering is therefore not merely to group similar nodes but to construct regions that reduce control-relevant coupling and improve the economy–efficiency tradeoff of partitioned optimization.

To ensure that the obtained partitions can serve as control-feasible units, a resource-adequacy check is introduced after dynamic clustering and physical correction. For each controllable resource $r \in \mathcal{C}$, let b_r denote its connected bus, and $Q^{\min}_r$, $Q^{\max}_r$, and $Q_r(t)$

denote its lower limit, upper limit, and current reactive-power operating point, respectively. The controllable-resource set in partition c is defined as

$$\mathcal{C}_c(t) = \{r \in \mathcal{C} \mid b_r \in \Omega_c(t)\}. \tag{24}$$

The available upward and downward reactive-power regulation capacities are

$$Q_c^{+,\mathrm{ava}}(t) = \sum_{r \in \mathcal{C}_c(t)} (Q_r^{\max} - Q_r(t)), \quad Q_c^{-,\mathrm{ava}}(t) = \sum_{r \in \mathcal{C}_c(t)} \left(Q_r(t) - Q_r^{\min}\right). \tag{25}$$

A partition with $|\mathcal{C}_c(t)| = 0$ is identified as a resource-insufficient electrical partition and is not used as an independent VVC-OPF control region. For such partitions, a control-region mapping is applied before VVC-OPF validation. Let $\mathcal{R}(t) = \{c \mid |\mathcal{C}_c(t)| > 0\}$ be the set of resource-containing partitions and $\mathcal{Z}(t) = \{c \mid |\mathcal{C}_c(t)| = 0\}$ be the set of zero-resource partitions. For each $c \in \mathcal{Z}(t)$, the supporting control region is selected as

$$\pi(c) = \arg \min_{m \in \mathcal{R}(t)} \bar{d}^{e}_{c,m}(t), \tag{26}$$

where $\bar{d}^{e}_{c,m}(t)$ is the average electrical distance across the boundary between partition c and resource-containing partition m. The control-feasible region is then formed as

$$\widetilde{\Omega}_m(t) = \Omega_m(t) \cup \left(\bigcup_{c \in \mathcal{Z}(t), \pi(c) = m} \Omega_c(t) \right). \tag{27}$$

Therefore, zero-resource partitions are treated as boundary-supported electrical sub-regions rather than independent control regions. Their voltage constraints and boundary exchanges are still included in the partition-level VVC-OPF, but the required reactive-power regulation is provided by the nearest electrically coupled resource-containing region.

3. Data–Physics Fusion Dynamic Clustering Model Based on Heterogeneous Spatio-Temporal Graph Attention

Figure 2 illustrates the workflow of the HST-GAT enhanced dynamic partitioning module. For each rolling window, normalized node features and relation-specific graph structures are first constructed from topology, voltage/reactive sensitivity, and electrical-distance information. Relation-specific graph attention layers then extract node embeddings under each relation. These embeddings are fused through learnable relation weights and updated by a GRU across rolling windows. The resulting latent embeddings are used for soft assignment and k-medoids clustering. Finally, physical correction is applied to enforce connectivity, reduce boundary coupling, and improve control implementability.

3.1. Heterogeneous Dynamic Graph Construction

At time window t, each node in the graph corresponds to a physical bus of the ADN, carrying the feature vector $x_i(t)$ defined in Section 2.1. To capture the multi-faceted coupling among buses, three relations are introduced to construct a heterogeneous dynamic graph [31]:

1. Topology edges, which inherit the physical branch connectivity of the feeder and preserve the fundamental structure of power delivery paths;
2. Sensitivity edges, which link node pairs whose voltage/reactive coupling exceeds a threshold, thereby encoding the control-relevant interaction strength;
3. Electrical-distance edges, which connect nodes that are close in the electrical-distance space defined by (6), promoting compactness in the partition.

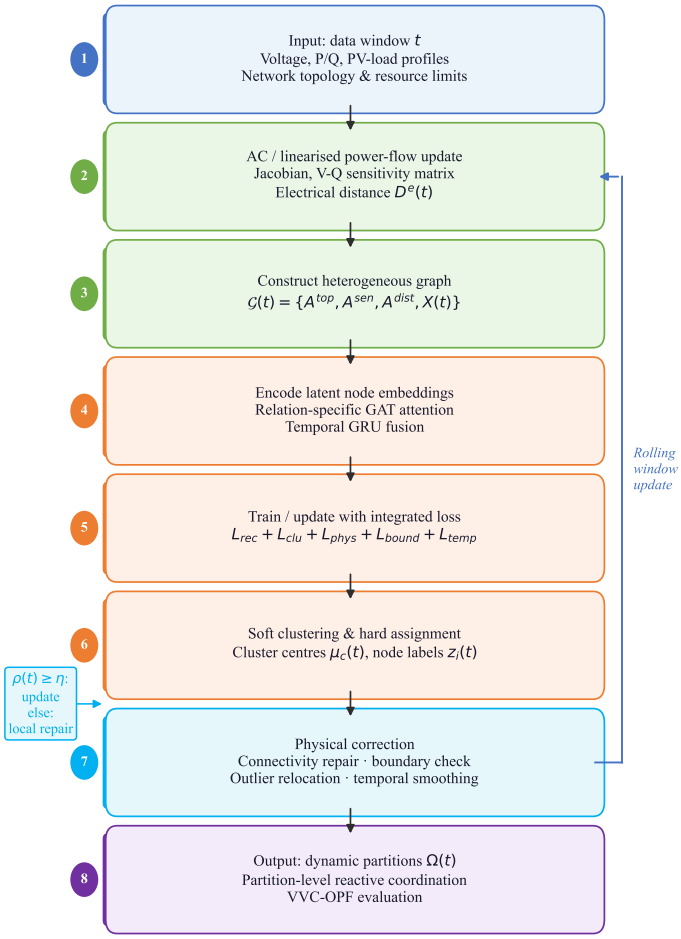

Figure 2. Algorithm flow of the HST-GAT enhanced dynamic partitioning method.

For relation type r, the adjacency matrix is denoted as $\mathbf{A}^{(r)}(t)$. The heterogeneous graph is

$$\mathcal{G}(t) = \left(\mathcal{V}, \{\mathcal{E}^{(r)}(t)\}_{r=1}^{R}, \mathbf{X}(t)\right). \tag{28}$$

The use of heterogeneous relations is motivated by the fact that topology, sensitivity, and electrical proximity describe different but related aspects of coupling. Topology preserves physical connectivity, sensitivity reflects control response, and electrical distance supports compact and connected partitions. However, these relations may also contain redundant information. Therefore, the paper treats heterogeneous graph learning as a scenario-dependent enhancement rather than an unconditional replacement of the physics-based baseline.

3.2. Heterogeneous Spatio-Temporal Graph Attention Encoder

For relation r, the attention coefficient between nodes i and j is

$$\alpha_{ij}^{(r)}(t) = \frac{\exp\left(\mathrm{LeakyReLU}\left(\mathbf{a}_r^\top [\mathbf{W}_r h_i(t) \| \mathbf{W}_r h_j(t)]\right)\right)}{\sum_{k \in \mathcal{N}_i^{(r)}} \exp\left(\mathrm{LeakyReLU}\left(\mathbf{a}_r^\top [\mathbf{W}_r h_i(t) \| \mathbf{W}_r h_k(t)]\right)\right)}. \tag{29}$$

The relation-specific embedding is then

$$\tilde{h}_i^{(r)}(t) = \sigma\left(\sum_{j \in \mathcal{N}_i^{(r)}} \alpha_{ij}^{(r)}(t) \mathbf{W}_r h_j(t)\right). \tag{30}$$

Relation fusion is carried out by

$$h_i^g(t) = \sum_{r=1}^{R} \omega_r(t) \tilde{h}_i^{(r)}(t), \tag{31}$$

where $\omega_r(t)$ is the relation weight. A gated recurrent unit (GRU) cell is used to transfer hidden states across rolling windows. It updates the current hidden state using the fused graph-attention representation and the hidden state from the previous window, thereby helping the HST-GAT module capture temporal evolution in operating conditions while maintaining compact model complexity.

$$h_i^s(t) = \mathrm{GRU}\Big(h_i^g(t), h_i^s(t-1)\Big). \tag{32}$$

The final latent node representation is

$$e_i(t) = \mathbf{W}_o h_i^s(t). \tag{33}$$

3.3. Latent-Space Clustering and Physical Correction

Soft assignment is obtained from latent-space distance:

$$p_{ic}(t) = \frac{\exp(-\|e_i(t) - \mu_c(t)\|_2^2/\tau_c)}{\sum_{k=1}^{K} \exp(-\|e_i(t) - \mu_k(t)\|_2^2/\tau_c)}, \tag{34}$$

where $\mu_c(t)$ is the cluster center and τ_c is the temperature coefficient. The initial hard label is

$$\hat{z}_i(t) = \arg\max_c p_{ic}(t). \tag{35}$$

The initial label may still violate engineering requirements. Therefore, physical correction is applied after latent clustering. The correction includes connectivity repair, high electrical-distance outlier relocation, boundary-coupling reduction, and temporal rollback when label changes are not justified by operating-state changes. The corrected label is

$$z_i(t) = \arg\min_c \Big[d_{ic}^{\mathrm{emb}}(t) + \alpha_e d_{ic}^e(t) + \alpha_b b_{ic}(t) + \alpha_t \mathbb{I}(c \neq z_i(t-1))\Big]. \tag{36}$$

This stage is the key link between data-driven representation learning and physical implementability. It ensures that the learned partition remains connected, electrically reasonable, and sufficiently stable for partition-level control.

3.4. Loss Function and Training Strategy

The training objective combines representation learning, clustering compactness, physical consistency, boundary decoupling, and temporal smoothness [32]:

$$L(t) = L_{\mathrm{rec}}(t) + \lambda_c L_{\mathrm{clu}}(t) + \lambda_p L_{\mathrm{phys}}(t) + \lambda_b L_{\mathrm{bound}}(t) + \lambda_t L_{\mathrm{temp}}(t). \tag{37}$$

The reconstruction term is

$$L_{\mathrm{rec}}(t) = \|\hat{\mathbf{X}}(t) - \mathbf{X}(t)\|_F^2. \tag{38}$$

The clustering compactness term is

$$L_{\mathrm{clu}}(t) = \sum_{i=1}^{N} \sum_{c=1}^{K} p_{ic}(t) \|e_i(t) - \mu_c(t)\|_2^2. \tag{39}$$

The physical consistency term is

$$L_{\text{phys}}(t) = \sum_{c=1}^{K} \sum_{i,j \in \Omega_c(t)} p_{ic}(t) p_{jc}(t) d_{ij}^{e}(t). \tag{40}$$

The boundary penalty is

$$L_{\text{bound}}(t) = \sum_{(i,j) \in \mathcal{E}} \sum_{c \neq c'} p_{ic}(t) p_{jc'}(t) \Big(|P_{ij}(t)| + |Q_{ij}(t)| + \kappa |S_{ij}^{VQ}(t)| \Big). \tag{41}$$

The temporal term is

$$L_{\text{temp}}(t) = \sum_{i=1}^{N} \| p_i(t) - p_i(t-1) \|_2^2. \tag{42}$$

The HST-GAT training loss weights, k-medoids distance weights, and related clustering parameters are reported in Section 4.

During online operation, the physics-based baseline can be used directly under normal scenarios, whereas HST-GAT is invoked under high-volatility and high-coupling conditions.

3.5. Dynamic Updating Mechanism

The implemented method adopts a periodic rolling update with temporal smoothing, rather than a purely event-triggered repartitioning strategy. The partition is updated once for each rolling window, and unnecessary label oscillations are suppressed by the smoothing term J_{temp} in the assignment cost. Since the fused distance is normalized to $[0, 1]$ through the preprocessing described in Section 2.1, the smoothing weight can be interpreted as an effective update threshold. The dynamic updating parameters and a comparison with event-triggered updating are reported in Section 4.

4. Case Studies and Simulation Setup

All experiments are implemented in Python (PyTorch-based implementation for HST-GAT components) and Python with Gurobi-based OPF solver. All experiments are conducted on a single machine with Intel Core i5-14600K CPU, without GPU acceleration or parallel computing. All runtime results are reported as mean values over multiple independent runs and rolling windows. Each configuration is executed repeatedly under identical solver settings, and the reported variance reflects run-to-run numerical fluctuation of the OPF solver and clustering initialization. Although absolute runtimes for small systems are in the millisecond range, repeated execution is necessary due to solver numerical variability; therefore all timing results are statistically averaged to ensure fair comparison.

4.1. Test Systems and Data Description

Three test systems are used: a modified IEEE 33-bus system, an IEEE 123-bus system, and a practical feeder system (Feeder Z). The IEEE 33-bus system is used to test the effect of dynamic partitioning on a medium-scale network. The IEEE 123-bus system is used to test a larger network and the high-volatility HST-GAT enhancement. The Feeder Z topology is used to verify the adaptability of the proposed workflow to practical feeder data, non-consecutive node indices, distributed generation access, and switch status records.

The load and solar profiles used in this study are actual measured data collected from the practical distribution network. The load dataset contains measurements from 123 buses over 35,040 time steps at 15 min resolution. The solar dataset contains six measured PV-site irradiance profiles over the same time horizon. These profiles include clear-sky, partly cloudy, and intermittent cloud-passage conditions, thereby reflecting

realistic daily, seasonal, and site-dependent renewable variability rather than manually smoothed synthetic profiles.

Table 2 summarizes the tested systems, including their network scale, controllable-resource configuration, and validation purpose.

Table 2. Test systems and data sources used in the experiments.

System	Nodes	Branches	Controllable Resources	Main Role
IEEE 33	33	32	8	Small-scale normal-scenario validation
IEEE 123	123	122	14	Medium-scale and high-volatility validation
Feeder Z	102	101	5	Practical topology-format adaptability validation

4.2. Scenario Definition

Four daily scenarios, illustrated in Figure 3, are designed for evaluation. S1 and S2 represent normal operating conditions and are used to compare global VVC-OPF, fixed-partition VVC-OPF, and dynamic-partition VVC-OPF. S3 and S4 represent high-volatility conditions, including PV intermittency and sharp load ramping, and are used to assess whether learned representations bring additional benefits over the physics-based dynamic partition baseline.

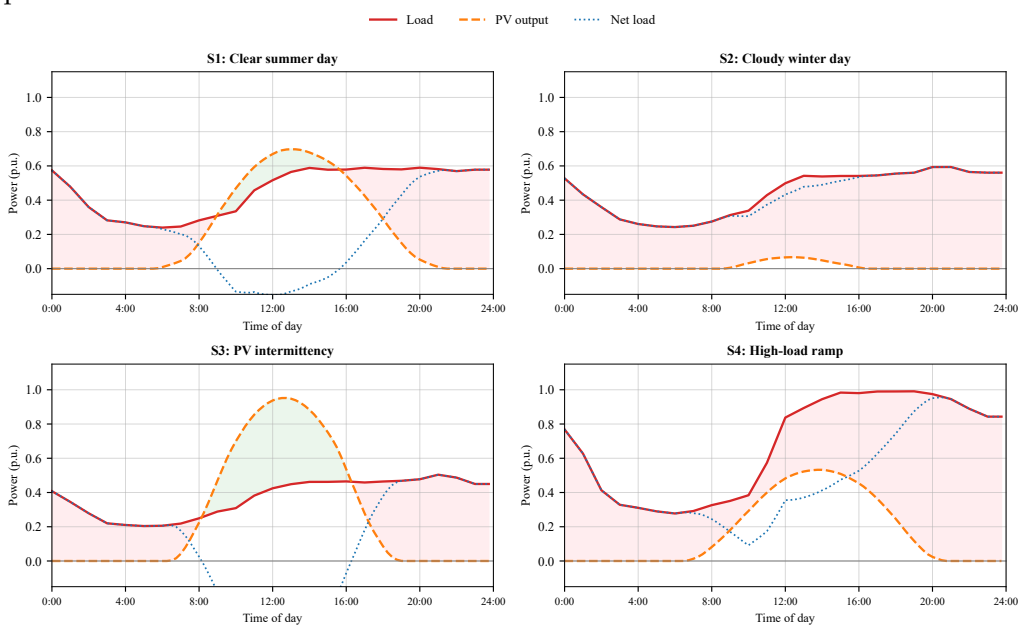

Figure 3. Representative 24 h load, PV output, and net-load profiles for the four test scenarios.

The compared methods are summarized in Table 3.

Table 3. Overview of compared methods.

Method	Description
Global VVC-OPF	Solves the whole network directly
Fixed-partition VVC-OPF	Solves partitioned problems using unchanged regions
Physics-based VVC-OPF	Rolling dynamic-partition VVC-OPF based on physical heuristics
MLP-based partition	Dynamic partition based on a Multi-Layer Perceptron model
Single-relation GAT	Dynamic partition based on single-relation Graph Attention Networks
HST-GAT	Dynamic partition based on the proposed Heterogeneous Spatio-temporal GAT

4.3. Parameter Settings

The main parameters used in the dynamic clustering objective and physical correction are summarized in Table 4. The same penalty-weight configuration is used for all test systems. Only K is changed according to network scale and the expected partition granularity. Specifically, $K = 3$, $K = 6$, and $K = 4$ are used for the modified IEEE 33-bus system, IEEE 123-bus system, and Feeder Z system, respectively. The loss weights $\lambda_{\text{emb}} = 0.30$, $\lambda_e = 0.429$, $\lambda_b = 0.231$, and $\lambda_t = 0.14$, as well as the smooth weight and connectivity quantile, are kept unchanged across all systems. These values are calibrated on the IEEE 123-bus system and then directly transferred to the other test systems without case-specific retuning.

Table 4. Main parameters used in the dynamic clustering objective and physical correction.

Test System	K	λ_{emb}	λ_e	λ_b	λ_t	Smooth Weight	Conn. Quantile
Modified IEEE 33-bus	3	0.30	0.429	0.231	0.14	0.10	0.80
IEEE 123-bus	6	0.30	0.429	0.231	0.14	0.10	0.80
Feeder Z, 102-node	4	0.30	0.429	0.231	0.14	0.10	0.80

The parameters are determined through a staged procedure. First, the number of partitions K is selected according to system scale and then checked through a compactness–coupling trade-off. Second, the loss weights are calibrated on the IEEE 123-bus system, which is used as the main validation feeder due to its larger scale and stronger operating diversity. The same loss-weight configuration is then directly transferred to the modified IEEE 33-bus and Feeder Z systems. The physical loss is assigned the dominant role, with $\lambda_e + \lambda_b = 0.660$, where 65% is assigned to intra-region electrical consistency and 35% to boundary coupling, yielding $\lambda_e = 0.429$ and $\lambda_b = 0.231$. Third, the temporal smoothness weight λ_t is selected as the smallest value that removes spurious label oscillations without suppressing necessary operating-state-driven partition updates. Finally, the connectivity quantile is selected to correct only the nodes with relatively large electrical distance to their assigned medoids, thereby enforcing physical connectivity while avoiding excessive reassignment.

The HST-GAT training objective is defined as $\mathcal{L} = \mathcal{L}_{\text{rec}} + 0.30\mathcal{L}_{\text{clu}} + 0.429\mathcal{L}_{\text{elec}} + 0.231\mathcal{L}_{\text{bound}} + 0.14\mathcal{L}_{\text{temp}}$. The soft-assignment temperature is set to 0.8. After HST-GAT training, the final k-medoids distance is computed as $d_{\text{hybrid}} = 0.32d_{\text{emb}} + 0.38d_{\text{base}} + 0.04d_{\text{temp}} + 0.26d_{\text{elec}}$. The k-medoids smoothing weight for the HST-GAT enhanced partition is 0.135, the connectivity correction quantile is 0.80, and the maximum number of k-medoids iterations is 20. Since the fused distance is normalized to $[0, 1]$, the smoothing weight can be interpreted as an effective update threshold:

$$\eta_{\text{eff}} = \lambda_{\text{smooth}} \max d_{\text{fused}} = 0.10. \tag{43}$$

The dynamic updating parameters are: rolling-window length $T_w = 96$, maximum number of rolling windows $W_{\max} = 8$, smoothing weight $\lambda_{\text{smooth}} = 0.10$, and effective update threshold $\eta_{\text{eff}} = 0.10$.

The practical configuration of the HST-GAT module is summarized in Table 5. The model is trained in an unsupervised full-batch manner for each rolling window. Since no ground-truth partition labels are used, a conventional supervised training/validation split is not applicable. The learned embeddings are evaluated through downstream partition-quality and VVC-OPF metrics.

Table 5. HST-GAT model configuration and training settings.

Item	Setting
Input dimension	28
Hidden dimension	16
Latent embedding dimension	6
GAT structure	1 heterogeneous graph attention layer
Attention heads	1 head per relation, 3 heads in total
Relation types	Topology, voltage/reactive sensitivity, electrical distance
Temporal module	Single-layer GRU
GRU input/hidden dimension	16/16
Activation function	tanh
Optimizer	Adam, $\beta_1 = 0.9$, $\beta_2 = 0.999$
Learning rate/weight decay	$3 \times 10^{-3}/1 \times 10^{-4}$
Batch size	Full graph per rolling window
Training epochs	20
Hardware	Intel Core i5-14600K CPU
Average training time	Approximately 0.09 s per rolling window

The 28-dimensional input vector consists of 7 voltage features, 5 DG features, 10 temporal features, and 6 sensitivity-statistic features. For each relation, the query, key, and value projections are implemented as bias-free linear mappings from 16 to 16 dimensions. The top-K neighbor number is set to 8, the attention scaling factor is $1/\sqrt{16}$, and non-neighbor entries in the attention mask are filled with -10^9. The relation prior weights are 0.20, 0.45, and 0.35 for topology, sensitivity, and electrical-distance relations, respectively. Relation-specific embeddings are fused through learnable relation logits followed by softmax, and a residual term with coefficient 0.20 is added to the fused hidden representation. The decoder is implemented as Linear $6 \rightarrow 16$, tanh, and Linear $16 \rightarrow 28$.

The value of λ_{smooth} was selected by sweeping it from 0 to 0.30; 0.05 was sufficient to remove spurious label oscillations, and 0.10 was adopted as a conservative setting. In an event-triggered comparison, the trigger-based updating can skip approximately 75–100% of windows in stable scenarios, but because one repartitioning step costs only about 0.06 s in the tested systems, the absolute time saving is limited. Therefore, periodic rolling updating is retained as the default implementation.

5. Results and Discussion

5.1. Control Performance Under Normal Operating Scenarios

Table 6 compares the global VVC-OPF, fixed-partition VVC-OPF, and dynamic-partition VVC-OPF under normal operating scenarios, where the normalized cost reflects control-decomposition accuracy and the runtime reflects online computational burden.

Table 6. VVC-OPF comparison under normal operating scenarios.

System	Method	Norm. Cost	Time/s (Mean ± Std)
IEEE 33	Global	1.00000	0.4600 ± 0.3992
IEEE 33	Fixed-partition	1.00145	$0.03788 \pm 2.27 \times 10^{-5}$
IEEE 33	Dynamic-partition	1.00017	$0.05280 \pm 1.88 \times 10^{-5}$
IEEE 123	Global	1.00000	0.24226 ± 0.12813
IEEE 123	Fixed-partition	1.01148	$0.05097 \pm 9.23 \times 10^{-6}$
IEEE 123	Dynamic-partition	1.00099	$0.06004 \pm 4.68 \times 10^{-5}$

All runtimes are averaged over multiple independent runs and rolling windows (IEEE 33: 3 windows, IEEE 123: 4 windows). The results show that fixed partitioning has

the shortest solution time but introduces a larger control-decomposition error. Dynamic partitioning preserves most of the computational advantage of partitioned optimization while keeping the objective close to the global solution. In the IEEE 123-bus case, the normalized cost of dynamic partitioning is 1.000985, much lower than 1.011484 of fixed partitioning, while the solution time is only about 24.8% of the global VVC-OPF.

Figure 4 confirms that dynamic partitioning lies closer to the global solution in cost while incurring a much smaller computational burden. This supports the core motivation of using dynamic partitions as control units.

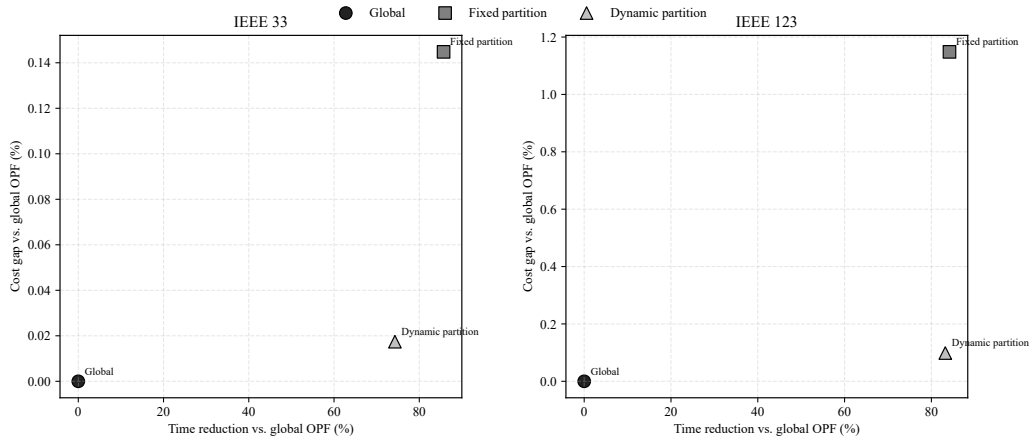

Figure 4. Cost-time tradeoff of global, fixed-partition, and dynamic-partition VVC-OPF.

5.2. Performance Enhancement Under High-Volatility and High-Coupling Scenarios

Under high-volatility scenarios, the original dynamic partition achieves a normalized cost of 1.00699 (feature extraction 0.01535 s, updating 0.01535 s, OPF solving 0.00603 s), while the HST-GAT enhanced method achieves a normalized cost of 1.00670 (feature extraction 0.01480 s, representation learning 0.00252 s, updating 0.03632 s, OPF solving 0.00755 s).

All runtime results are reported as full wall-clock execution time, including feature extraction, representation learning, clustering, correction, and OPF solving. The original method requires 0.080 s in total (0.036 s partitioning and 0.044 s OPF), while the HST-GAT method requires 0.209 s in total (0.150 s partitioning and 0.059 s OPF). HST-GAT increases the total computational cost due to additional graph learning overhead but improves boundary decoupling under high-volatility conditions. Figure 5 shows that HST-GAT reduces boundary active and reactive exchanges in most high-volatility scenarios.

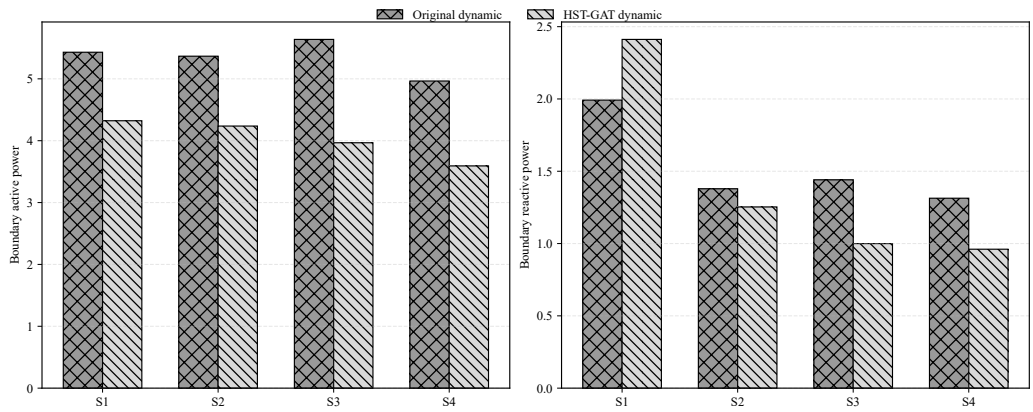

Figure 5. Boundary active and reactive exchange (p.u.) comparison under high-volatility scenarios.

Figure 6 compares the label change rates of the two partitions across the four high-volatility scenarios, providing a direct measure of temporal partition stability. In scenarios

S1 and S3, both methods exhibit zero label change, indicating that the partition structure remains fully stable when operating conditions follow relatively predictable patterns. In scenarios S2 and S4, which feature stronger renewable intermittency and load fluctuations, the HST-GAT partition shows higher label switching activity, with rates of 15.4% and 5.7%, respectively. This increased switching reflects the graph-learning model's responsiveness to fine-grained spatio-temporal variations that the physics-based baseline does not capture. Despite this, the label change rates remain contained within acceptable bounds, confirming that the temporal smoothing penalty embedded in the loss function effectively prevents excessive partition oscillation while allowing necessary boundary adaptation.

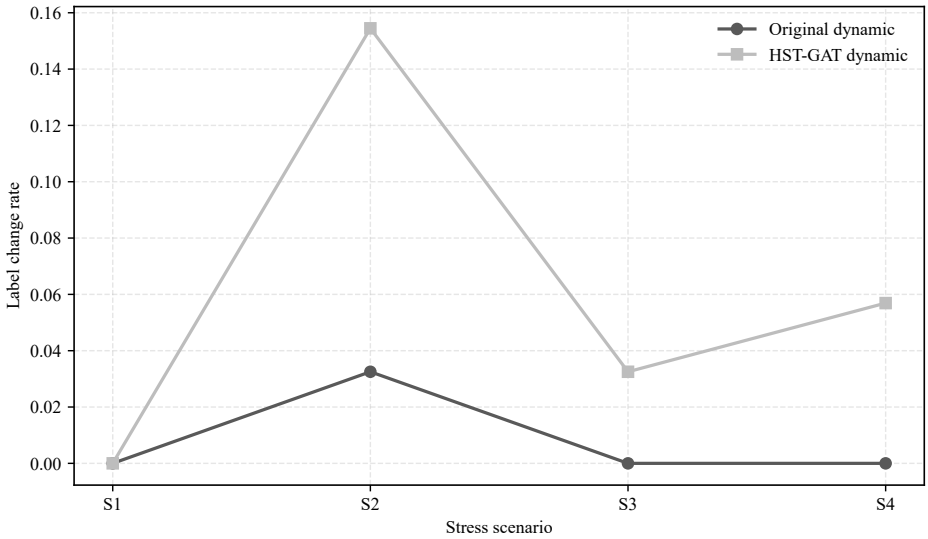

Figure 6. Label change rate comparison under high-volatility scenarios.

Table 7 provides a scenario-level breakdown of two complementary partition-quality indicators. Across all four scenarios, the original dynamic partition maintains marginally tighter intra-cluster electrical distances, with mean values ranging from 0.01298 to 0.01348, confirming that the physics-based features produce electrically compact clusters. The HST-GAT partition shows slightly higher electrical distances of 0.01473 to 0.01887, indicating a controlled relaxation of electrical homogeneity. This relaxation is accompanied by scenario-dependent changes in label stability: in S2, the HST-GAT label change rate reaches 15.4%, substantially higher than the baseline's 3.3%, whereas in S1 both methods remain fully stable. This pattern reveals that the graph-learning model selectively permits boundary adjustments only when the operating state exhibits strong temporal variation, while maintaining partition stability under quiescent conditions. The scenario-level granularity of Table 7 thus clarifies the conditional nature of the HST-GAT enhancement and its underlying trade-off between physical compactness and boundary decoupling.

Table 7. Scenario-level partition-quality comparison under high-volatility scenarios.

Scenario	Method	Mean Intra-Cluster Electrical Distance	Label Change Rate
S1	Original dynamic partition	0.01298	0.00000
S1	HST-GAT dynamic partition	0.01887	0.00000
S2	Original dynamic partition	0.01348	0.03252
S2	HST-GAT dynamic partition	0.01601	0.15447
S3	Original dynamic partition	0.01348	0.00000
S3	HST-GAT dynamic partition	0.01551	0.03252
S4	Original dynamic partition	0.01348	0.00000
S4	HST-GAT dynamic partition	0.01473	0.05691

Figure 7 tracks the evolution of four partition-quality indicators across the four high-volatility scenarios. Two differentiating behaviors are evident. First, the HST-GAT partition consistently achieves lower boundary reactive exchange than the original dynamic partition, confirming that the learned representations provide stronger boundary decoupling under high-coupling conditions. Second, the intra-cluster electrical distance of the HST-GAT partition is slightly higher in several windows, indicating that the latent-space clustering trades some electrical compactness for reduced cross-boundary coupling. The label change rate remains low for both methods in most windows, with HST-GAT showing modestly higher switching activity in Scenarios S2 and S4 where operating conditions fluctuate more rapidly. Overall, the normalized cost traces of the two methods track each other closely, demonstrating that the HST-GAT enhancement preserves the control economy of the physics-based baseline while selectively improving boundary decoupling.

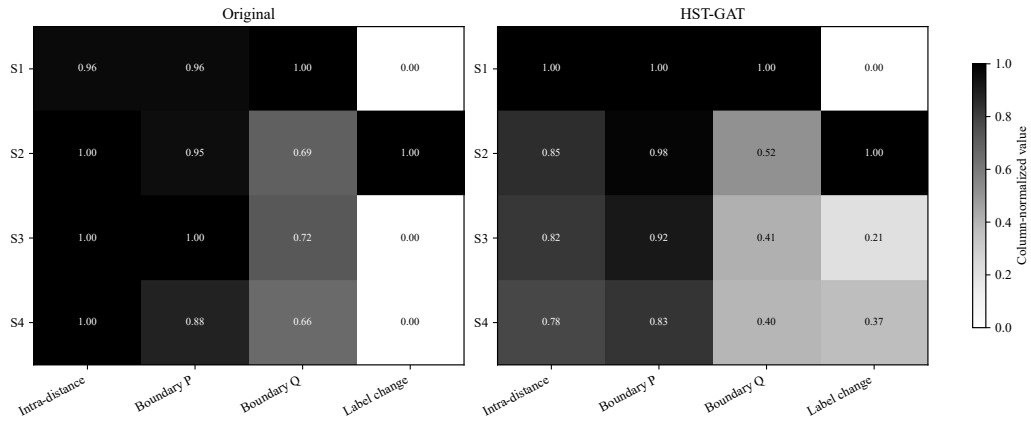

Figure 7. Evolution of partition-quality indicators under high-volatility scenarios.

Figure 8 visualizes the temporal evolution of partition labels for all 123 buses across a full day under two contrasting operating conditions. In both sub-figures, each horizontal color band represents a group of buses that remain in the same partition. Three observations can be drawn from this figure.

The partition labels exhibit strong temporal persistence, with wide continuous color bands indicating that the smoothing penalty suppresses unnecessary boundary oscillations and maintains multi-hour stability. Boundary changes are confined to transition periods, such as PV ramp-up and ramp-down, where $\rho(t)$ captures shifts in voltage/reactive sensitivity and triggers only local reassignment. The consistent coarse patterns across daily scenarios further show that the physics-based features produce reproducible, topology-driven partitions with adaptive boundary refinement under volatile operating conditions.

5.3. Model Comparison

All methods are evaluated under identical partitioning windows. Each reported value is averaged over $n = 10$ or $n = 30$ runs depending on stability requirements, as summarized in Table 8. The original dynamic partition remains the strongest baseline in terms of electrical compactness and label stability. The MLP achieves the lowest average normalized cost in this comparison group but incurs higher computational burden and weaker physical compactness. HST-GAT delivers a more balanced result by reducing boundary reactive exchange while maintaining a moderate solution time.

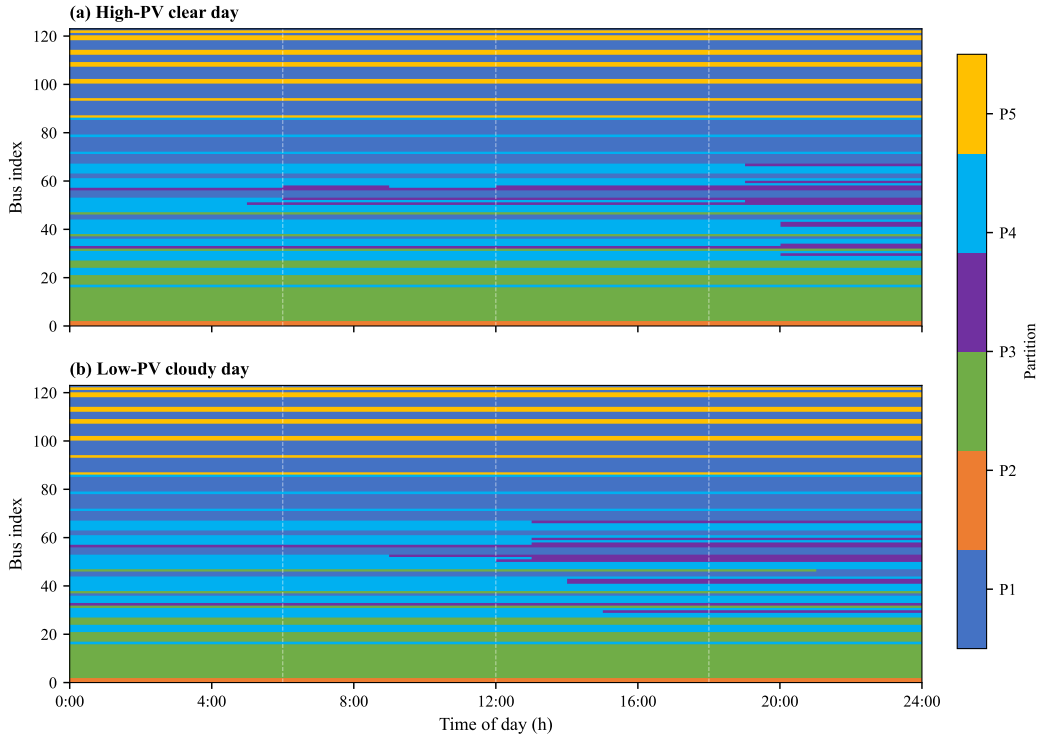

Figure 8. Dynamic partition evolution on the IEEE 123-bus system over 24 h: (**a**) high-PV clear day; (**b**) low-PV cloudy day. Colors represent partition labels; white dashed lines mark 6:00, 12:00, and 18:00.

Table 8. Representation-level comparison under high-volatility scenarios.

Method	Norm. Cost	Time/s	Mean Intra-Cluster Electrical Distance	Boundary P (p.u.)	Boundary Q (p.u.)	Label Change Rate
Original	1.00490	0.05225	0.01337	4.06065	1.63164	0.00813
MLP	1.00461	0.04270	0.01903	3.62343	1.77015	0.07317
Single-relation GAT	1.00476	0.07418	0.01638	3.65654	1.89450	0.11992
HST-GAT	1.00479	0.04627	0.01608	3.99771	1.52450	0.02236

5.4. Ablation Study

Table 9 demonstrates that physical constraints and temporal smoothing are essential components. Removing physical constraints increases the normalized cost from 1.00479 to 1.00647 and raises boundary active exchange from 3.998 to 5.803 p.u., confirming that physical regularization is necessary for maintaining control-oriented partition quality. Removing temporal smoothing leads to a substantially higher label change rate (from 0.022 to 0.100), indicating that the temporal term effectively suppresses unnecessary partition oscillation. Removing heterogeneous relations reduces the intra-cluster electrical distance from 0.01608 to 0.01262 and the label change rate from 0.022 to 0.000 but increases boundary active exchange from 3.998 to 4.044 p.u., indicating a trade-off between electrical compactness and boundary decoupling that the multi-relation fusion mechanism helps to balance.

The HST-GAT module introduces additional representation-learning cost. For normal or weakly fluctuating operation, the physics-based dynamic partition is preferable because of its low update burden, high stability, and clear physical interpretation. For high-volatility and high-coupling scenarios, HST-GAT can be used as an enhancement module to improve boundary decoupling and control-oriented cost.

Table 9. Ablation results of the compact HST-GAT model.

Variant	Norm. Cost	Time/s	Mean Intra-Cluster Electrical Distance	Boundary P (p.u.)	Boundary Q (p.u.)	Label Change Rate
HST-GAT	1.00479	0.04612	0.01608	3.99771	1.52450	0.02236
Without heterogeneous relations	1.00461	0.09206	0.01262	4.04371	1.15058	0.00000
Without physical constraints	1.00647	0.08626	0.01313	5.80252	1.40008	0.00000
Without temporal smoothing	1.00481	0.04940	0.01557	3.92075	1.55047	0.09959

5.5. Sensitivity Analysis of Electrical-Distance Weighting Coefficient

To examine the influence of the impedance weighting coefficient γ_z on the resulting partition, γ_z is varied from 0 to 1 on the IEEE 123-bus system, and the results are summarized in Table 10. In this test, each γ_z value is evaluated independently without previous-label inheritance or temporal-smoothing regularization so that the measured partition changes are caused only by the electrical-distance fusion weight. The resulting partitions are compared with the $\gamma_z = 0$ baseline. The label change rate is calculated after label matching, and the normalized mutual information (NMI) is also reported to measure structural consistency between partitions.

Table 10. Sensitivity of partitioning results to γ_z.

γ_z	Label Change Rate	NMI	Avg. Intra-Cluster Electrical Distance
0.0	0.0%	1.000	0.1333
0.5	0.0%	1.000	0.1370
0.6	18.2%	0.878	0.1445
1.0	27.6%	0.881	0.1479

The results show that the partition is strictly unchanged when $\gamma_z \leq 0.5$, with a label change rate of 0.0% and NMI of 1.000. When γ_z increases beyond 0.5, part of the boundary-node assignment changes. For $\gamma_z = 0.6$ and $\gamma_z = 1.0$, the label change rates are 18.2% and 27.6%, respectively. However, the corresponding NMI values remain high, 0.878 and 0.881, indicating that the overall partition structure is still largely preserved. Meanwhile, the average intra-cluster electrical distance increases from 0.1333 to 0.1479, corresponding to an 11.0% increase. These results indicate that the partitioning result is stable in the recommended range $\gamma_z \leq 0.5$ and remains structurally similar even when the impedance term dominates. Therefore, $\gamma_z = 0$ is adopted in the reported experiments for computational efficiency and direct consistency with voltage/reactive-power control coupling.

5.6. Hyperparameter Sensitivity Analysis

To verify that the conclusions are not caused by manually favorable parameter choices, a one-at-a-time hyperparameter sensitivity analysis is conducted on the IEEE 123-bus system under representative scenarios. The tested parameters include the number of partitions K, the k-medoids smooth weight, the connectivity quantile, and the fused-distance weights. The results are summarized in Table 11.

The sensitivity analysis shows that the selected parameters do not artificially create the reported conclusions. The number of partitions K mainly controls the physical trade-off between intra-region compactness and inter-region coupling. For the IEEE 123-bus system, increasing K from 4 to 8 reduces the intra-cluster electrical distance from 0.1600 to 0.1269 but increases boundary active exchange from 3.826 MW to 7.012 MW. Thus, $K = 6$ is selected as a balanced setting. The smooth weight and connectivity quantile have negligible influence on the main partition structure, with a label change rate no larger than 0.81% and NMI no lower than 0.98. Although the fused-distance weights are more influential, the resulting

partitions remain structurally similar under representative perturbations as indicated by NMI values no lower than 0.82. Therefore, the selected parameter set represents a stable configuration rather than a manually favorable choice.

Table 11. Summary of hyperparameter sensitivity analysis.

Parameter	Tested Range	Main Observation	Conclusion
K	4–8	Larger K improves intra-region compactness but increases boundary coupling	K controls the expected compactness–coupling trade-off
Smooth weight	0–0.30	Label change rate $\leq 0.81\%$, NMI ≥ 0.98	Low sensitivity
Connectivity quantile	0.60–0.95	Boundary metrics change slightly and NMI remains high	Low sensitivity
Fused-distance weights	representative perturbations	Most sensitive parameter group, but NMI remains ≥ 0.82	Main structure remains preserved

5.7. Resource-Adequacy Analysis

Table 12 reports the controllable-resource distribution in the representative baseline operating window. The table lists the local controllable-resource buses, available upward/downward reactive-power capacities, and resource-to-reactive-load ratio for each electrical partition. Partitions without local controllable resources are explicitly marked with their control-handling strategy.

Table 12. Controllable-resource distribution and handling of zero-resource partitions.

System	Part.	Resource Buses/Types	$Q^{+,\mathrm{ava}}$ (p.u.)	$Q^{-,\mathrm{ava}}$ (p.u.)	Res./Load (p.u.)	Control Handling
IEEE 33	1	Bus 14, GFM-DG	0.0010	0.0010	0.0019	Independent
IEEE 33	2	Buses 3, 6, 20, 24, 27, GFM-DG	0.0155	0.0155	0.0175	Independent
IEEE 33	3	Buses 31, 33, GFM-DG	0.0053	0.0053	0.0059	Independent
IEEE 123	1	none	0.0000	0.0000	0.0000	Boundary-supported
IEEE 123	2	none	0.0000	0.0000	0.0000	Boundary-supported
IEEE 123	3	Buses 1, 3, 6, 14, GFM-DG	0.0078	0.0078	0.0299	Independent
IEEE 123	4	none	0.0000	0.0000	0.0000	Boundary-supported
IEEE 123	5	Buses 20, 24, 27, 31, 33, GFM-DG	0.0149	0.0149	0.0826	Independent
IEEE 123	6	none	0.0000	0.0000	0.0000	Boundary-supported

The results show that all three partitions in the modified IEEE 33-bus system contain at least one controllable GFM-DG resource. In the IEEE 123-bus system, controllable resources are concentrated in partitions 3 and 5, while partitions 1, 2, 4, and 6 do not contain local reactive-power resources. Therefore, not every electrical partition is assumed to operate as an independent control region. Zero-resource partitions are identified and mapped to the nearest resource-containing control region before VVC-OPF validation. This treatment preserves the electrical interpretability of the dynamic partition while ensuring that the final control decomposition remains implementable.

5.8. Adaptability to Practical Topology

The Feeder Z case verifies whether the proposed workflow can handle engineering data formats, as summarized in Table 13. The case includes non-consecutive node numbering, branch topology, switch status, and distributed generation information. These features reflect common difficulties in real distribution-system data preparation.

Table 13. VVC-OPF comparison on the Feeder Z topology.

Method	Norm. Cost	Time/s (Mean ± Std)	Violation Sum (p.u.)	Partition Time/s	Solve/s
Global	1.00000	0.25509 ± 0.10599	27.27043	0.00000	0.13805
Fixed-partition	1.31815	$0.05068 \pm 2.42 \times 10^{-5}$	34.96951	0.00298	0.00514
Dynamic-partition	1.03818	$0.08153 \pm 2.52 \times 10^{-5}$	28.81025	0.01066	0.01117

The fixed partition has the shortest solution time but causes a much larger objective deviation and violation sum. Dynamic partitioning reduces the normalized cost from 1.318153 to 1.038178 and keeps the solution time far below that of the global VVC-OPF. The comparison suggests that the proposed dynamic partitioning workflow can reduce the control-decomposition error caused by fixed partitions while retaining a clear computational advantage.

Figure 9 shows the simplified topology and the dynamic partition labels at the last representative time window. The Feeder Z case is based on a practical 102-node feeder topology with non-consecutive bus numbering, switch status records, actual DG capacity information, and measured load profiles. The complete dynamic partitioning and VVC-OPF validation pipeline is successfully executed on this feeder, providing further evidence that the proposed method can be applied to practical feeder-format data.

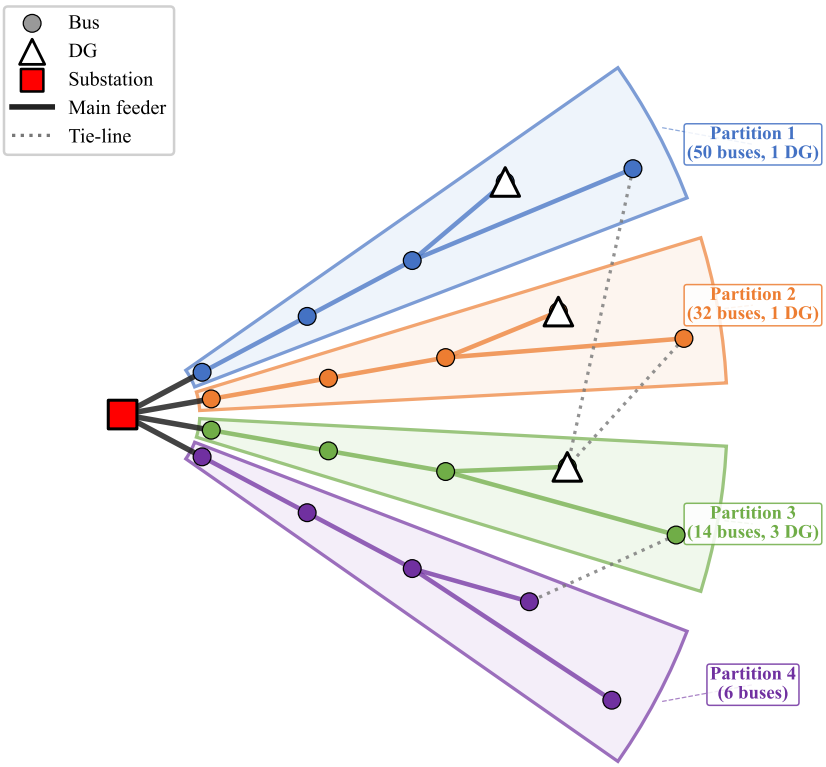

Figure 9. Simplified schematic of the Feeder Z dynamic partition at a representative time window.

5.9. Full-Year Validation

To further evaluate the generalization capability of the proposed method beyond the four representative scenarios, a full-year validation is conducted using 364 consecutive daily operating windows. Each daily window contains 96 time points at 15 min resolution. The results are summarized in Table 14.

Table 14. Full-year validation results over 364 daily operating windows.

Metric	Mean	Std.	P50	P95
Intra-cluster electrical distance	0.01360	0.00070	0.01350	0.01490
Boundary active exchange (MW)	4.421	0.796	4.053	5.809
Boundary reactive exchange (MVAr)	1.181	0.359	1.057	2.048
Boundary edges	6.76	1.23	7.00	9.00
Label change rate	0.052	0.101	0.000	0.244
Partition time (s)	0.0306	0.0078	0.0296	0.0407

The full-year results show that the proposed partitioning method maintains stable electrical compactness across diverse annual operating conditions. The coefficient of variation of the intra-cluster electrical distance is approximately 5.1%, indicating that the electrical consistency of the partitions is preserved over the annual cycle. The median label change rate is zero, while the 95th percentile is 0.244, showing that most daily partitions remain stable and that larger changes mainly occur under more pronounced seasonal or renewable-output transitions. The partition time remains below 0.041 s at the 95th percentile, confirming that the method remains computationally feasible for online rolling implementation.

6. Conclusions

This paper proposed a data–physics fusion-driven dynamic partitioning method for active distribution networks oriented to fast power coordination. The method is designed to construct dynamic control partitions rather than to pursue generic clustering accuracy. A physics-based rolling dynamic partition baseline is first built from operating behavior, voltage/reactive sensitivity, electrical distance, and feeder topology. For high-volatility and high-coupling scenarios, an HST-GAT enhancement module is introduced to learn control-relevant latent node embeddings. Physical correction, boundary-coupling penalties, and temporal smoothing are used to ensure electrical consistency, implementability, and dynamic continuity.

Results on the modified IEEE 33-bus and IEEE 123-bus systems show that dynamic partitioning approaches the global VVC-OPF cost while greatly reducing computational burden. Under high-volatility scenarios, HST-GAT further reduces boundary active and reactive exchanges and improves the average normalized cost, although its additional computational overhead is non-negligible. Therefore, HST-GAT should be regarded as a scenario-dependent enhancement over a strong physics-based rolling partition baseline. The Feeder Z case further verifies that the data-processing, topology-reconstruction, dynamic-partitioning, and OPF verification workflow can be transferred to practical distribution-system data formats.

Future work will focus on three aspects. First, event-triggered repartitioning under topology reconfiguration, fault isolation, and service restoration should be further modeled. Second, stronger classical baselines such as sensitivity-based spectral clustering and graph-cut partitioning should be included. Third, online deployment should consider communication delay, missing measurements, uncertainty in sensitivity estimation, and interactions with distributed controllers.

Author Contributions: Conceptualization, Z.Z. (Zhi Zhou), S.H. and T.Y.; methodology, Z.Z. (Zhi Zhou), Z.M. and T.Y.; investigation, Z.Z. (Zhi Zhou), S.H., R.H., Y.L. and Z.M.; resources, Z.Z. (Zhi Zhou), S.H., R.H., Q.Y. and Z.Z. (Zhenglin Zhong); writing—original draft preparation, Y.L. and Z.M.; writing—review and editing, Z.Z. (Zhi Zhou), T.Y. and Y.L.; visualization, Y.L. and Z.M.; supervision, T.Y. and Z.Z. (Zhi Zhou); project administration, Z.Z. (Zhi Zhou) and T.Y.; funding

acquisition, Z.Z. (Zhi Zhou) and T.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the China Southern Power Grid Corporation Technology Project (GZKJXM20240546).

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

Conflicts of Interest: Authors Zhi Zhou, Siyang He, Rui He, Quanhai Yang, and Zhenglin Zhong were employed by the company Duyun Power Supply Bureau, Guizhou Power Grid Co., Ltd., a subsidiary of China Southern Power Grid Corporation. The remaining authors (Yubin Liu, Tao Yu, and Zixi Mo) declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest. The funder had the following involvement with the study: through the employment of authors Zhi Zhou, Siyang He, Rui He, Quanhai Yang, and Zhenglin Zhong, who contributed to conceptualization, resources, investigation, writing—review and editing, supervision, project administration, and funding acquisition as detailed in the Author Contributions section.

Abbreviations

The following abbreviations are used in this manuscript:

ADN	Active distribution network
DER	Distributed energy resource
DG	Distributed generation
GAT	Graph attention network
HST-GAT	Heterogeneous spatio-temporal graph attention network
OPF	Optimal power flow
PV	Photovoltaic
VVC-OPF	Voltage-constrained Volt/Var-control optimal power flow

References

1. Luo, C.; Wu, H.; Zhou, Y.; Qiao, Y.; Cai, M. Network partition-based hierarchical decentralised voltage control for distribution networks with distributed PV systems. *Int. J. Electr. Power Energy Syst.* **2021**, *130*, 106929. [CrossRef]
2. Meng, L.; Yang, X.; Zhu, J.; Wang, X.; Meng, X. Network partition and distributed voltage coordination control strategy of active distribution network system considering photovoltaic uncertainty. *Appl. Energy* **2024**, *362*, 122846. [CrossRef]
3. Sowa, P.; Zychma, D. Dynamic equivalents in power system studies: A review. *Energies* **2022**, *15*, 1396. [CrossRef]
4. Han, P.; Lin, Z.; Wang, L.; Fan, G.; Zhang, X. A survey on equivalence modeling for large-scale photovoltaic power plants. *Energies* **2018**, *11*, 1463. [CrossRef]
5. Jiang, C.; Guo, C.; Lin, J.; Lin, J.; Zheng, S. Enhancing economic efficiency and operational stability of high-penetration renewable distribution networks: A multi-timescale coordinated optimization method leveraging multi-agent graph reinforcement learning. *Renew. Energy* **2026**, *256*, 124345.
6. Yu, N.; Zhang, S.; Qin, J.; Hidalgo-Gonzalez, P.; Dobbe, R.; Liu, Y.; Dubey, A.; Wang, Y.; Dirkman, J.; Zhong, H.; et al. Data-driven control, optimization, and decision-making in active power distribution networks. *Appl. Energy* **2025**, *397*, 126253.
7. Qiu, R.; Bai, H.; Yao, R.; Liu, C.; Xu, M.; Chen, Q.; Yang, W. A data-physics-driven modeling approach of key equipment for large-scale distribution network simulation. *Electronics* **2024**, *13*, 1663. [CrossRef]
8. Liu, Y.; Wang, M.; Zhang, X.; Duan, J.; Gao, H.; Liu, J. Kriging surrogate model enabled heuristic algorithm for coordinated Volt/Var management in active distribution networks. *Electr. Power Syst. Res.* **2022**, *210*, 108089. [CrossRef]
9. Yu, H.; Wang, Y.; Wang, Q.; Cheng, Z.; Li, S.; Xiao, X.; Chen, Y. Dynamic network partition and voltage regulation method by PVs considering reactive power compensation benefits. *Int. J. Electr. Power Energy Syst.* **2025**, *165*, 110464. [CrossRef]
10. Pineda, S.; Pérez-Ruiz, J.; Morales, J.M. Beyond the Neural Fog: Interpretable Learning for AC Optimal Power Flow. *IEEE Trans. Power Syst.* **2025**, *40*, 4912–4921. [CrossRef]
11. Ruan, H.; Gao, H.; Liu, Y.; Wang, L.; Liu, J. Distributed voltage control in active distribution network considering renewable energy: A novel network partitioning method. *IEEE Trans. Power Syst.* **2020**, *35*, 4220–4231. [CrossRef]
12. Zhang, Z.; Li, P.; Ji, H.; Zhao, J.; Xi, W.; Wu, J.; Wang, C. Combined central-local voltage control of inverter-based DG in active distribution networks. *Appl. Energy* **2024**, *372*, 123813.

13. Li, J.; Wang, X.; Guo, Q.; Yang, C.; Chen, Y.; Zhao, R.; Li, M. A cluster-optimized regulation method for distribution networks considering distributed photovoltaic heterogeneous characteristics. *Int. J. Electr. Power Energy Syst.* **2025**, *170*, 110857. [CrossRef]
14. Lu, N.; Fu, X.; Zhang, P.; Qiu, D.; Badihi, H.; Abdel-Salam, M.; Gu, H. Enhancing representative photovoltaic scenario extraction for multiple power stations with a shared-weight and adaptively fused graph clustering method. *Appl. Energy* **2026**, *406*, 127291.
15. Qiu, S.; Deng, Y.; Ding, M.; Han, W. An Optimal Scheduling Method for Distribution Network Clusters Considering Source–Load–Storage Synergy. *Sustainability* **2024**, *16*, 6399.
16. Pan, L.; Han, Z.; Zhao, S.; Wang, F. An optimal allocation method for power distribution network partitions based on improved spectral clustering algorithm. *Eng. Appl. Artif. Intell.* **2023**, *123*, 106497. [CrossRef]
17. Mahdavizadeh, S.F.; Aghamohammadi, M.R.; Ranjbar, S. Frequency stability-based controlled islanding scheme based on clustering algorithm and electrical distance using real-time dynamic criteria from WAMS data. *Sustain. Energy Grids Netw.* **2022**, *30*, 100560. [CrossRef]
18. Aghamohammadi, M.R.; Mahdavizadeh, S.F.; Rafiee, Z. Controlled Islanding Based on the Coherency of Generators and Minimum Electrical Distance. *IEEE Access* **2021**, *9*, 146830–146840. [CrossRef]
19. Fallahzadeh-Abarghouei, H.; Hasanvand, S.; Nikoobakht, A.; Doostizadeh, M. Decentralized and hierarchical voltage management of renewable energy resources in distribution smart grid. *Int. J. Electr. Power Energy Syst.* **2018**, *100*, 117–128. [CrossRef]
20. Ma, C.; Sun, T.; Chen, Z.; Zhu, M.; Song, M.; Huang, D.; Yang, M.; Sun, L. Methods for distribution network partitioning and hybrid time-scale voltage/var control considering photovoltaic uncertainty. *Electr. Power Syst. Res.* **2026**, *258*, 113059. [CrossRef]
21. Huang, Y.; Deng, W.; Zhang, K.; Liu, M.; Liu, M.; Li, Q.; Zhang, Y. Charge-discharge hierarchical-multi-stage decoupling optimization for large-scale EVs interacting with volt/var control in distribution systems. *Appl. Energy* **2025**, *400*, 126557.
22. Qi, J.; Guo, Q.; Zhu, Y.; Yu, Y.; Tu, L.; Luo, C.; Sun, C.; Tang, Y.; Liu, Y. Dynamic equivalent modeling of distributed photovoltaic generation systems in microgrid considering LVRT active power response difference. *Electronics* **2025**, *14*, 4355. [CrossRef]
23. Tang, Y.; Sun, S.; Zhao, B.; Ni, C.; Che, L.; Li, J. Distributed photovoltaic generation aggregation approach considering distribution network topology. *Energies* **2024**, *17*, 2990. [CrossRef]
24. Yang, S.; Hu, X.; Tu, Z.; Yin, R.; Zhang, Y.; Li, J.; Wang, L.; Meng, L. A distributed photovoltaic aggregation method considering photovoltaic output characteristics and active voltage sensitivity. *Electronics* **2025**, *14*, 917. [CrossRef]
25. Tan, H.; Miao, S.; Wang, T.; Wang, B.; Wang, J.; Yao, F.; Zhang, D. Research on transient equivalent modeling method for distribution networks with high-penetration distributed photovoltaics. *Sol. Energy* **2025**, *299*, 113725. [CrossRef]
26. Dan, Y.; Zhong, H.; Wang, C.; Wang, J.; Fei, Y.; Yu, L. A Graph Deep Reinforcement Learning-Based Fault Restoration Method for Active Distribution Networks. *Energies* **2025**, *18*, 4420. [CrossRef]
27. Hua, H.; Zou, Y.; Chen, X.; Yu, K.; Gan, L.; Lei, S.; Gertrudes, J.B.; Zheng, J.; Sidorov, D. Dynamic network partitioning for a large-scale microgrid cluster. *IEEE Trans. Smart Grid* **2025**, *16*, 5615–5627. [CrossRef]
28. Liao, W.; Bak-Jensen, B.; Pillai, J.R.; Wang, Y.; Wang, Y. A review of graph neural networks and their applications in power systems. *J. Mod. Power Syst. Clean Energy* **2022**, *10*, 345–360. [CrossRef]
29. Chen, Y.; Liu, Y.; Zhao, J.; Qiu, G.; Yin, H.; Li, Z. Physical-assisted multi-agent graph reinforcement learning enabled fast voltage regulation for PV-rich active distribution network. *Appl. Energy* **2023**, *351*, 121743.
30. Luo, F.; Wang, S.; Lv, Y.; Mu, R.; Fo, J.; Zhang, T.; Xu, J.; Wang, C. Domain knowledge-enhanced graph reinforcement learning method for Volt/Var control in distribution networks. *Appl. Energy* **2025**, *398*, 126409.
31. Wang, W.; Suo, X.; Wei, X.; Wang, B.; Wang, H.; Dai, H.-N.; Zhang, X. HGATE: Heterogeneous graph attention auto-encoders. *IEEE Trans. Knowl. Data Eng.* **2023**, *35*, 3938–3951.
32. Hu, X.; Hu, H.; Verma, S.; Zhang, Z.-L. Physics-guided deep neural networks for power flow analysis. *IEEE Trans. Power Syst.* **2021**, *36*, 2082–2092.

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