IEEE PES Power & Energy Society®

Graph Learning-based Power System Health Assessment Model

KOJI YAMASHITA*, MEMBER, IEEE, NANPENG YU*, SENIOR MEMBER, IEEE, EVANGELOS FARANTATOS†, SENIOR MEMBER, IEEE, AND LIN ZHU †, SENIOR MEMBER, IEEE

[1]University of California, Riverside, Riverside, CA 92521 USA
[2]Electric Power Research Institute, Knoxville, TN 37932 USA

Corresponding Author: Nanpeng Yu (e-mail: nyu@ece.ucr.edu).

ABSTRACT As the power transmission system's energy sources become increasingly diversified, the grid stability is experiencing increased fluctuations, thereby necessitating more frequent and near real-time monitoring by grid operators. The power system security has been monitored through real-time contingency analysis and dynamic security assessment framework, both of which are typically based on time-domain simulations or power flow calculations. Achieving higher accuracy in grid health level prediction often requires time-consuming simulation and analysis. To improve computational efficiency, this paper develops machine learning models with phasor measurement unit (PMU) data to monitor the power system health index, focusing on rotor angle stability and frequency stability. The proposed machine learning models accurately predict frequency and angle stability indicators, essential for evaluating grid health considering various contingencies, even when dealing with limited PMU deployment in transmission grids. The proposed framework leverages a physics-informed graph convolution network and graph attention network with ordinal encoders, which are benchmarked with multi-layer perceptron models. These models are trained on dataset derived from an augmented IEEE 118-bus system with different demand levels and fuel mix, including tailored dynamic generator models, generator controller models, and grid protection models. The numerical studies explored the performance of the proposed and baseline machine learning models under both full PMU coverage and various partial PMU coverage conditions, where different data imputation methods are used for substations without PMUs. The findings from this study offer valuable insights, such as machine learning model selection and critical PMU locations regarding power equipment, into the design of data-driven grid health index prediction models for power systems.

INDEX TERMS Attention mechanism, dynamic security, graph neural network, grid health index.

I INTRODUCTION

Conventional power system security assessment applications primarily rely on time-domain simulation results of power system dynamics models, often supplemented by direct methods such as screening techniques for preliminary stability assessments [1]. Such applications are in practical use by grid operators [2], [3]. Despite the remarkable improvement in computer hardware and simulation algorithms, it remains a challenge to analyze numerous power system operating scenarios and contingencies in a short period of time [4], [5]. On the other hand, machine learning (ML)-based grid security monitoring applications are in the early stage of development. Such ML-based applications are in limited practical use. With the growing diversification of power system components, the need for near real-time grid security monitoring is increasing. ML-based grid security monitoring emerges as a promising solution to address this need. By employing ML-based algorithms, computationally intensive stability studies in bulk power systems can be accelerated, enabling rapid bulk power system security assessment without relying solely on model-based simulations.

Naturally, the accuracy of the ML-based grid security monitoring algorithms could approach but not surpass that of the time-domain simulations with power system dynamic models. Instead of replacing the existing model-based grid security monitoring systems, the ML-based approaches can be used for preliminary assessment of the operating scenarios and contingencies in real time. The operating scenarios and contingencies that pose the highest level of threat can then be fed to near real-time time-domain simulations for more thorough analysis.

Simulation-based grid security assessment systems usually automatically generate a power flow snapshot of the current grid condition based on measurement information obtained through state estimation. The monitoring system then performs power system dynamic analysis using the

power flow snapshot as the initial system state for various contingencies. Despite the advancement in accelerating dynamic simulations, this process is computationally-intensive, typically requiring a few minutes or more for a single contingency scenario. To train an ML-based grid health prediction algorithm, a sufficiently large number of power flow snapshots and grid sensor data at substations are required as input data. The corresponding grid security or health levels regarding frequency, transient, and voltage stability are needed as output data. Once the ML-based models are trained, the online security and grid health assessment will only take a very short amount of time.

In recent years, research on ML-based grid security monitoring primarily leveraged various graph neural networks (GNNs) [6]–[12]. The GNNs are a class of deep learning models tailored for data structured as graphs, where nodes represent buses in a power grid and edges represent transmission or distribution lines, making GNNs well-suited for analyzing complex and interconnected systems such as the power grid. References [6], [7] provide a comprehensive review of machine learning-based contingency analysis algorithms. The literature surveys identified that data-driven system monitoring and dynamic studies have not fully explored the time-varying grid structure and system operating conditions. Most existing data-driven approaches are designed to examine the rotor angle stability phenomenon, and research on dynamic security assessment for frequency stability is limited. In one such study, convolutional neural networks were utilized to address frequency stability [13]. Voltage stability has also been studied with GNNs [10]. However, this paper primarily focuses on predicting steady-state voltage violations, thus overlooking dynamic aspects of power system behavior.

Reference [8] develops a graph attention network (GAT) that predicts the dynamic response of synchronous generators using real-time system measurements. Specifically, the developed model predicts future dynamic behaviors of the power grid with the input of the first 10-step dynamic response following disturbances. This model may be used for real-time power system control purposes. Reference [9] proposes a data-driven transient stability assessment model using the graph convolutional network (GCN) with a multi-pooling mechanism. Both max pooling and mean pooling are leveraged to improve the power system stability prediction performance. This model may be used for dynamic security assessment purposes. These studies [8]–[10], [13] hypothesize measurement devices are installed in all nodes, i.e., substations and power stations, which is not realistic in real-world transmission systems. In other words, the partial grid sensor coverage study has been missing in the research on ML-based grid security assessment algorithms. On the other hand, grid sensor placement has been separately studied with no ML algorithms or data-driven approaches [14]–[17].

Reference [14] provides a comprehensive literature review on grid sensor placement, covering concepts, methods, and research needs. This reference highlights that most studies in transmission grids focus on static state estimation, neglecting dynamic behavior. While some research addresses power system dynamics, such as rotor angle stability, they primarily focus on small-signal stability, with minimal attention to frequency stability. Reference [15] also discusses research gaps for optimal phasor measurement unit (PMU) placement, pointing out that more research is needed on optimal PMU placement that considers specific applications such as controlled islanding, fault tolerance, small-signal analysis, and voltage stability. Reference [18] specifically addresses frequency stability application but with a limited number of PMUs at power stations. In [18], the focus is solely on tracking the center of inertia during dynamic frequency response, ignoring frequency nadirs at critical locations.

Many PMU placement studies emphasize the minimum number of PMUs required to achieve complete system observability. The optimization problem's objective function is often set as minimizing PMU cost or maximizing grid observability [19]. However, from a grid health monitoring perspective, which focuses on assessing system strength, particularly in dynamic security, it is essential to examine critical substations, considering the connected power component to the substation (e.g., tie transformers, generators, and load feeders). Conventional research has predominantly focused on observability, with cutting-edge zero injection bus integration techniques and utilization of SCADA [19]–[22], which primarily highlight topological observability, and has given less attention to PMU placements from the dynamic grid health monitoring perspective. Reference [23] introduces a PMU placement method using binary particle swarm optimization that incorporates aspects of power component-based considerations, optimizing the number of substations requiring installations to ensure complete network observability while addressing constraints such as critical measurements, restricted installation locations, and the upgrade of line relays to dual-use digital relay PMUs. Although clarifying the optimal number of PMUs is critical, this approach primarily focuses on the minimum number of PMUs for observability, rather than specifically optimizing placements based on critical power equipment for dynamic grid health monitoring.

Moreover, determining the optimal number of PMUs is crucial, not only for complete system observability but also for ensuring that the solution represents a global optimum rather than a local one. The practical implementation is further complicated by limitations in *PMU channel capacity,* as each PMU can monitor only a finite number of channels, which underscores the need for efficient placement strategies and capacity management [19], [20], [24]. Reference [25] examines critical buses in terms of both observability and rotor angle stability with pre-determined critical buses for transient and small-signal stability. However, it does not address machine learning-based stability assessments, which are vital for evaluating dynamic grid health through advanced computational methods.

Motivated by the lack of thorough studies on grid health assessment algorithms with limited grid sensor availability, this paper aims to fill this gap by proposing a measurement-driven grid health assessment algorithm designed to accommodate partial PMU coverage scenarios. While grid health assessment encompasses both adequacy (the ability to meet demand) and security (the resilience to disturbances), our focus in this work is on security—specifically, assessing the grid's dynamic stability across different contingency scenarios. We propose a physics-informed GCN and GAT network to predict the frequency and angle health index of power systems. Additionally, this paper offers guidance to industry practitioners unfamiliar with ML techniques on how to implement and utilize the proposed algorithm. The primary contributions of this paper include:

- To promote data-sharing and reproducibility, this paper creates publicly available pre-contingency power flow snapshots under various loading levels and realistic contingency analysis results using time-domain simulations on an augmented IEEE 118-bus system with generator controller and protection models.
- We propose physics-informed GNNs with an ordinal encoder, which achieves state-of-the-art performance in grid frequency and angle health index prediction tasks. The numerical results provide useful guidance to industry practitioners in terms of machine learning model design for grid health index monitoring, which assesses system strength using hardware and software to ensure efficient operation.
- The comprehensive numerical study results involving partial PMU coverage provide useful insights into the optimization of PMU placement, which should prioritize substations with critical power equipment.

The rest of the paper is organized as follows. Section II overviews the grid health index prediction framework. Section III presents grid health index prediction models. Section IV describes the preparation of the dataset and models for dynamic simulations. Section V evaluates the grid health index prediction performance through case studies. Section VI concludes the paper.

II Enhanced Dynamic Security Assessment Framework with Measurement-driven Grid Health Index Prediction

As mentioned earlier, the proposed measurement-driven grid health assessment algorithm and model functions as a supplementary tool to the existing dynamic security assessment (DSA) framework, shown in Fig. 1. As shown in this figure, the measurement-driven DSA considers all contingency cases, including those not on the contingency list, after the initial contingency scenarios screening process that quickly classifies scenarios into critical and non-critical categories using simplified assessment methods, such as energy function approaches. If the proposed grid health index prediction tool identifies contingency cases that may lead to instability, then these cases will be further examined by grid operators

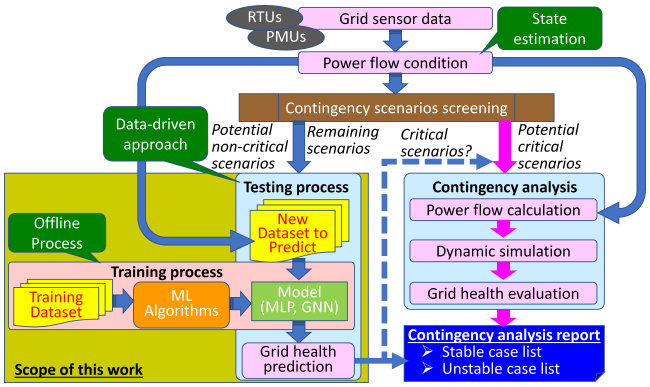


FIGURE 1. Overview of dynamic security assessment framework enhancement by integrating data-driven grid health index prediction.

using time-domain simulation, focusing primarily on the identified critical contingencies for computational efficiency.

The measurement-driven DSA will be especially useful when the power grid is undergoing a rapid transition to a new operation condition, and time-domain simulation is too slow to keep up. In such cases, the measurement-driven grid health index prediction model is the preferred tool for generating contingency analysis reports.

This paper proposes a measurement-driven grid health index prediction model that mitigates the weakness of the existing simulation-based real-time contingency analysis system for frequency and angular stability studies. The proposed model is based on GNNs, which are capable of dealing with both node and branch features. We extended the GNNs with an attention mechanism and ordinal encoder to further improve the grid health index prediction performance. The attention mechanism is a technique in machine learning that enhances model performance by prioritizing relevant features.

Comprehensive numerical studies to assess the performance of the proposed measurement-driven grid health index prediction model are critical. In particular, we performed detailed ablation studies to analyze the impacts of attention mechanism and ordinal encoder on the grid health index prediction accuracy. In addition, the performance of the proposed model is evaluated considering different partial grid sensor coverage scenarios. Specifically, we have considered the following types of substations lacking grid sensors: substations with synchronous generators, synchronous condensers, tie transformers, step-down transformers, and switching stations. Furthermore, we studied how different missing value replacement techniques impact the prediction accuracy for grid health index. In particular, we considered zero value imputation, peak demand value, and pseudo grid sensor measurement with errors. These in-depth numerical studies reveal new insights into the robustness of grid health prediction performance given partial grid sensor coverage.

TABLE 1. Inputs of Health Index Prediction Model

	Node/branch feature	Feature type	Measurable
1	Active power output, P_G	Node	Yes
2	Reactive power output, Q_G	Node	Yes
3	Active power load, P_L	Node	Yes
4	Reactive power load, Q_L	Node	Yes
5	Voltage magnitude, $\|V\|$	Node	Yes
6	Voltage angle, $\angle V$	Node	Yes
7	Active power output deviation, ΔP_G	Node	Yes
8	Active power load deviation, ΔP_L	Node	Yes
9	Reactive power output deviation, ΔQ_G	Node	Yes
10	Reactive power load deviation, ΔQ_L	Node	Yes
11	Fault location, ΔV*	Node	No**
12	Fault duration, F_{dur}	Node	No**
13	Active power transfer, P_{tie}	Branch	Yes
14	Reactive power transfer, Q_{tie}	Branch	Yes

* Three simplified levels: 0, 0.5, or 1 pu of residual voltage.
** auxiliary information (e.g., synthetic value) corresponding to contingency

III Overview of Grid Health Index Prediction Model

This section provides a high-level overview of the grid health index prediction model, covering its structure and key components, including physics-informed GNN, attention mechanism, and ordinal encoder. The model was implemented in Python using the PyTorch framework and is available in the GitHub repository: github.com/KojiGitCode/OAJPE.

A. Grid Health Index Model Structure

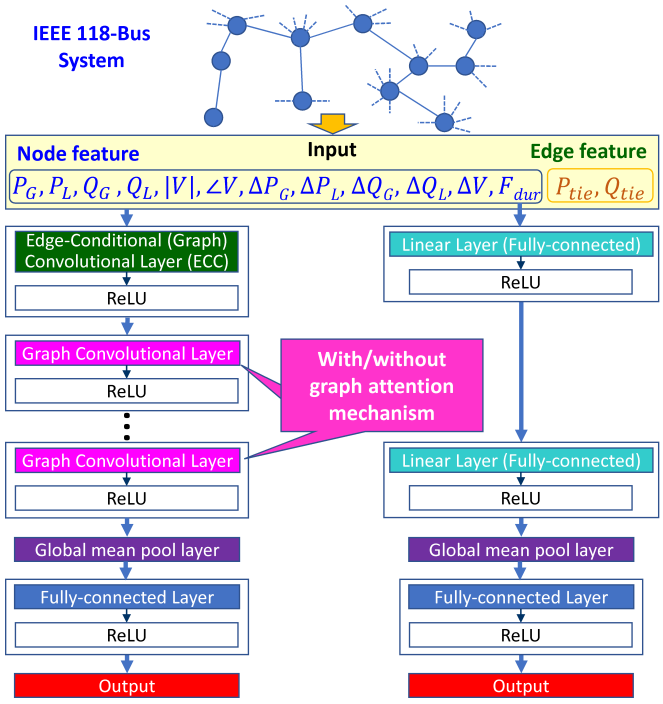
1) Neural Network Architecture

The GCN [26] is selected as a candidate deep learning model for the grid health index prediction. This is because power grids can be directly modeled as an undirected graph in GCN, where each node represents a bus and each edge represents a branch. An undirected graph is appropriate here because the direction of active and reactive power flow can vary with changing grid conditions and configurations. In addition, the GCN model can easily leverage both the nodal and the branch data from the power grid by encoding local and global network information through an edge-conditional mechanism [27]. This mechanism is implemented using an edge-conditional (ECC) layer, which allows the model to process edge features along with node features.

While the GCN is highly effective for supervised learning on graph-structured data, integrating edge weights into its training process is challenging. Thus, we propose to leverage the graph attention network (GAT) [28], [29] to predict the grid health index. By incorporating the attention mechanism, the GAT model excels at focusing on critical power lines by capturing the strength of inter-correlation between substations and buses. The overall architecture of the proposed GCN/GAT model and a *baseline*, multilayer perceptron (MLP) model are illustrated in Fig. 2. Unlike the GAT model and GCN with edge-conditional (ECC) layer, the MLP model cannot directly handle edge features.

2) Input Features

The input features of the GCN/GAT and MLP models are illustrated in Fig. 2. There are two categories of inputs to the grid health prediction model. The first category of inputs represents the pre-fault steady-state system information associated with the buses and branches. The second category of inputs is related to the contingency information, which corresponds to the 11-th and 12-th node features in Table 1. The fault location is indirectly represented by a synthetic voltage sag value , simplifying fault differentiation without expanding the feature set. We assume all bus voltages operate around 1.0 pu under normal conditions. For faults along transmission lines, modeled at the midpoint, we assign a residual voltage of 0.5 pu to reflect a typical 50% voltage drop at both line ends. For transformer faults, such as severe three-phase inter-turn faults, both primary and secondary nodes experience nearly zero voltage, represented by a residual value of 0.0 pu. Using these residual voltage values—1.0 pu for normal conditions, 0.5 pu for line faults, and 0.0 pu for transformer faults—enables effective fault type differentiation while maintaining a consistent feature set for efficient model training. The fault duration is normalized to 200 ms, meaning that, for example, 50 ms of the fault duration is represented as 0.25. Note that detailed input features for nodes and edges shown in Table 1 are assumed to be collected from PMUs.



Graph convolutional layers with attention become graph attention layers.

FIGURE 2. GCN/GAT model (left) and MLP model (right).

3) Outputs Representing Grid Health Index

The frequency health index is derived based on the frequency nadir, which is the lowest frequency level before reaching the post-steady-state frequency. The angle health index is derived based on the peak-to-peak angle difference between two particular units following the fault clearance. Detailed labels for these grid health indices are defined by separating the continuous index values into several different categories, as shown in Table 2. To create a balanced representation of sta-

ble and unstable scenarios, the range of angle differences is gradually decreased for smaller angle differences. Similarly, the frequency range gradually decreases as the frequency level approaches the rated frequency. This technique allows the dataset to capture both severe and moderate stability margins, reflecting a broad spectrum of system behavior across possible contingencies. This approach generalizes to other test systems by identifying generator pairs with the largest pre-contingency angle differences, which are often most critical for transient stability. Stability thresholds, such as 90 and 180 degrees, can serve as initial guidelines, adaptable to the specific grid characteristics. In contrast, frequency thresholds require adjustment for grid size; for example, small island grids may experience larger deviations than bulk power systems. The class labels will be used as the output of GNNs. Since both outputs are a single indicator per contingency, these grid health index prediction tasks are formulated as a graph classification problem. Given that the task can be viewed as a multi-class classification problem, the cross-entropy loss function is used to train the model's parameters and assess its performance, ensuring both prediction accuracy and handling any ordinal biases introduced by the encoder.

B. Graph Neural Network

The inputs of the GCN/GAT model are first passed through the ECC layer, which updates each node i's feature vector from x_i to x_i'. This is done by activating the sum of the product of the weight matrix and the nodal feature vector and the aggregation of nodal and edge feature vectors via a parameterized neural network, as shown in (1).

$$\boldsymbol{x}_i' = \text{ReLU}\left[\boldsymbol{W}\boldsymbol{x}_i + \sum\nolimits_{j\in\mathcal{N}(i)} \boldsymbol{x}_j \cdot h_\Theta(\boldsymbol{e}_{i,j})\right], \quad (1)$$

where $\boldsymbol{W}$ denotes the weight matrix, h_Θ denotes a neural network (e.g., a multilayer perceptron), $\boldsymbol{x}_i$ denotes the feature vector of the i-th node, and $\boldsymbol{e}_{i,j}$ denotes the edge feature vector for the edge from the source node, i, to the target node, j.

The output of the ECC layer is then used as the input of graph convolutional layers with/without the graph attention mechanism. The node representations $\boldsymbol{x}_i$ are updated at each graph convolutional layer according to (2).

$$\boldsymbol{x}_i' = \text{ReLU}\left[\widetilde{\boldsymbol{W}}^{\boldsymbol{T}} \sum_{j\in\mathcal{N}(i)} \hat{A}_{j,i} \cdot \boldsymbol{x}_j \cdot C\right], \quad (2)$$

$$C = \left(1 + \sum_{j\in\mathcal{N}(i)} \hat{A}_{j,i}\right)^{-\frac{1}{2}} \left(1 + \sum_{i\in\mathcal{N}(j)} \hat{A}_{i,j}\right)^{-\frac{1}{2}}, \quad (3)$$

where $\hat{A}_{i,j}$ denotes an ij-th element of $\hat{A} = A + I$, where A is the adjacency matrix. It is noted that $A_{i,j}$ equals 1 if node i is connected to node j and 0 otherwise.

C. Attention Mechanism

The GAT model [28] employs a refined approach to analyze graph-structured data through an attention mechanism, allowing it to focus on specific nodes and their interconnected relationships (represented by edges). The GCN model leverages pre-assigned edge weights between nodes, while the GAT model trains the edge weights using the attention mechanism. The nodal representations, x_i, are updated in each of the graph layers as shown below:

$$x_i' = \alpha_{i,i}\boldsymbol{W_s}x_i + \sum\nolimits_{j\in N(i), j\neq i} \alpha_{i,j}\boldsymbol{W_i}x_j \quad (4)$$

In (4), the weight matrix, $\mathbf{W}$, is separately expressed with $\mathbf{W}_s$ and $\mathbf{W}_i$ that correspond to the self-loop elements and other elements. The attention coefficient, $\alpha_{i,j}$ is calculated from:

$$\alpha_{i,j} = \text{softmax}(e_{i,j}) = \frac{e^{e_{i,j}}}{\sum_{k\in N_i} e^{e_{i,k}}} \quad (5)$$

$$e_{i,j} = a(\boldsymbol{W}\boldsymbol{h_i}, \boldsymbol{W}\boldsymbol{h_j}) = a^T\text{LeakyReLU}(\boldsymbol{W}\cdot[\boldsymbol{h_i}||\boldsymbol{h_j}]), \quad (6)$$

where $a \in \mathbb{R}^{2F'}$, $\boldsymbol{W} \in \mathbb{R}^{F'\times F}$, and $||$ denotes vector concatenation. F and $F^{'}$ denote the number of features at each node in the input and output of the graph attention layer, respectively. $\boldsymbol{h}$ denotes a set of node features. (6) proposed in [29] is the result of shifting a^T outside the original LeakyReLU$(\cdot)$ that was proposed in [28]. (6) enables the algorithm to handle the attention mechanism dynamically.

D. Ordinal Encoder

The range of levels (referred to as *labels* in machine learning terminology) in each health index can be described as an ordinal variable, reflecting a natural ordering in system health levels. For example, the difference between system health for labels 4 and 5 is much smaller than that of levels 4 and 1. To capture this inherent relationship, ordinal encoding, which introduces inductive bias is introduced. This inductive bias leverages the ordinal structure, guiding the model to interpret smaller differences between adjacent labels as less significant than differences between labels that are farther apart. This structured encoding allows the model to better understand and preserve the relative relationships between health states during training, adding a "hint" about the hierarchy in health index categories. To implement the ordinal encoder, we take the output values from the model's last fully-connected layer —known as logits L_i — and run it through Algorithm 1, which calculates the probability distribution over health index label s.

By encoding labels in this ordinal manner (e.g., "low," "medium," and "high" as 1, 2, and 3), we impose a structured representation that reinforces the ordinal relationships inherent in the data. This approach not only contributes to prediction accuracy but also enhances the model's understanding of label relationships, particularly in scenarios where the magnitude of difference between labels is meaningful.

Because the grid health index prediction task is formulated as a classification problem, the cross-entropy loss function is selected to compute gradients, optimizing the model's parameters to respect both accuracy and the ordinal bias introduced by the encoder. During training, the neural network learns by adjusting its weights and biases through

Algorithm 1 Classification with Ordinal Encoder

for every epoch **do**
K logits L_i are derived at the last layer
$s_i = sigmoid(L_i), \forall 1 \leq i \leq K$
$L'_i = \sum_{j \leq i} \log(s_j) + \sum_{i<j} \log(1 - s_j), \forall 1 \leq i \leq K$
$p'_i =$softmax$(L'_i), \forall 1 \leq i \leq K$
end for

TABLE 2. Labels Set for Grid Health Index

Class	Frequency	Angle bet. Buses 10-12	Angle bet. Buses 10-49
	Range (Hz) # of contingency	Range (deg) # of contingency	Range (deg) # of contingency
Label 0	(59.975, ∞] 13217	(180, ∞] 285	(180, ∞] 357
Label 1	(59.95, 59.975] 2581	(90, 180] 174	(90, 180] 527
Label 2	(59.85, 59.95] 3309	(60, 90] 481	(65, 90] 1024
Label 3	(59.70, 59.85] 3402	(45, 60] 586	(55, 65] 811
Label 4	(-∞, 59.70] 2716	(35, 45] 659	(45, 55] 1124
Label 5	N/A	(25, 35] 1162	(35, 45] 1291
Label 6	N/A	(17, 25] 1982	(25, 35] 2024
Label 7	N/A	(11, 17] 2580	(15, 25] 3062
Label 8	N/A	(8, 11] 2082	(10, 15] 2832
Label 9	N/A	(5, 8] 3121	(7, 10] 2496
Label 10	N/A	(3, 5] 3497	(4, 7] 3506
Label 11	N/A	(2, 3] 2341	(2, 4] 2992
Label 12	N/A	(1, 2] 2883	[0, 2] 3179
Label 13	N/A	[0, 1] 3392	N/A

backpropagation. The model calculates gradients of the loss function with respect to each parameter using the chain rule, allowing for the optimization of weights and biases via gradient descent. This process minimizes the loss function by iteratively updating the parameters, improving the model's ability to predict the grid health index, even in the presence of class imbalance. However, while the ordinal encoder generally improves performance, it is not always effective. There are instances, as observed in Section V, where the MLP achieves better results without the encoder, possibly due to the model converging to a local minimum.

IV Numerical Study Setup

A. Augmented IEEE 118-bus Test System

The IEEE 118-bus system [30] is adopted as the base testing environment. To make the system more realistic, step-up/down transformers and additional units at power stations are included [31]. These modifications expand the system to encompass 264 buses and 325 branches.

Reference [31] recommends modeling multiple units at selected power stations. Specifically, the number of generation units at bus 100 increases from one to two (see Table 3). Considering power plant types, the rated capacity of a single unit of the advanced combined cycle power plant powered by natural gas is limited to a maximum of 200 MW. Fuel types and the number of units at power stations are set based on [31] and [32]. Run-of-river hydropower facilities' units have smaller capacity, each consisting of a single generator. Thermal power stations, on the other hand, can host between one and four units.

Based on insights from [30], [31], the following steps are used to generate a wide range of grid operating conditions:

1) Loading Condition

Different loading scenarios are generated by adjusting the overall system demand from 40% to 100% of the peak demand, increasing in increments of 5%.

2) Power Station-level Dispatch

Advanced combined cycle power units powered by natural gas are fuel efficient. Thus, their active power outputs are fixed at over 90% for all units at the power station. When the loading level lowers, e.g., below 50%, 3 out of 5 power stations reduce the number of connected units without changing active power output. The run-of-river hydropower station is treated as a constant power output unit regardless of the loading level. The rest of the power stations, such as the coal/gas-fired and hydropower, are treated as variable active power units.

3) Unit-level Dispatch Scenario

Two unit-level dispatch scenarios are employed: even unit dispatch, where all units at a power station have the same active power output without unit disconnection, and uneven unit dispatch, where one unit's active power output is adjusted while others remain unchanged.

It is noted that on-load tap changers, shunt capacitors/reactors, and the reactive power output from synchronous generators and synchronous condensers are individually adjusted to satisfy the entire grid voltage level constraints (nominal voltage $\pm 5\%$) for a particular demand.

B. Dynamic Model of the Test System

1) Generator Model

The generator model used in this study is a round rotor generator model with a quadratic saturation, known as GENROU in the PSS/E. This is a two-axis model with a saturation curve using an exponential function. The leveraged time-domain simulation tool [33] expresses the saturation curve with the broken line approximation instead. These data points are provided in [31].

2) Generator Controller Model

a: Automatic Voltage Regulator (AVR)

The AVR model used in this paper, showcased in Fig. 3, is based on the EXST1 model used in [31]. The terminal voltage magnitude, V_T, is the input, and the field voltage magnitude, E_F, is the output of this model. However, there are notable differences between the representations in [31] and this paper. In [31], certain elements, such as the washout

TABLE 3. Generation Resource with Two Dispatch Scenarios

Bus number	Fuel type	Unit capacity	# of connected units for demand level	
			Even dispatch	Uneven dispatch
10	Coal	248 MW	2: [40%, 100%]	3: [90%, 100%] 2: [65%, 85%] 1: [40%, 60%]
12	AGCC	23.5 MW	4: [40%, 100%]	4: [40%, 100%]
25	Coal	215 MW	1: [40%, 100%]	1: [40%, 100%]
26	Coal	153 MW	2: [40%, 100%]	2: [65%, 100%] 1: [40%, 60%]
31	ROR	8.0 MW	1: [40%, 100%]	1: [40%, 100%]
46	Gas	10.9 MW	2: [40%, 100%]	2: [65%, 100%] 1: [40%, 60%]
49	AGCC	150 MW	1: [40%, 100%]	1: [40%, 100%]
54	Coal	40 MW	3: [40%, 100%]	2: [50%, 100%] 1: [40%, 45%]
59	AGCC	115 MW	2: [55%, 100%] 1: [40%, 50%]	2: [50%, 100%] 1: [40%, 45%]
61	AGCC	119 MW	2: [55%, 100%] 1: [40%, 50%]	2: [55%, 100%] 1: [40%, 50%]
65	Coal	220 MW	2: [40%, 100%]	2: [65%, 100%] 1: [40%, 60%]
66	Coal	238 MW	2: [40%, 100%]	2: [65%, 100%] 1: [40%, 60%]
69	Hydro	115 MW	1: [40%, 100%]	1: [40%, 100%]
80	Coal	162 MW	2: [40%, 100%]	2: [65%, 100%] 1: [40%, 60%]
87	ROR	12.5 MW	1: [40%, 100%]	1: [40%, 100%]
89	Coal	238 MW	3: [40%, 100%]	3: [85%, 100%] 2: [60%, 80%] 1: [40%, 55%]
100-1	AGCC	100 MW	1: [40%, 100%]	1: [40%, 100%]
100-2	AGCC	119 MW	2: [55%, 100%] 1: [40%, 50%]	2: [55%, 100%] 1: [40%, 50%]
103	Hydro	75 MW	1: [40%, 100%]	1: [40%, 100%]
111	Gas	67 MW	1: [40%, 100%]	1: [40%, 100%]
116	Gas	217 MW	1: [40%, 100%]	1: [40%, 100%]

AGCC: Advanced Gas Combined Cycle, ROR: Run-of-River.

element for high-frequency filtering, are excluded, resulting in the control diagram shown in Fig. 3. Additionally, [31] assumed zero delay for terminal voltage magnitude, V_T detection, which is refined in the model by assigning a small time constant of 0.01 seconds. Therefore, while Fig. 3 resembles the EXST1 model, it includes modifications to better reflect practical considerations compared to the representation in [31].

b: Turbine-governor Controller Model

The primary frequency control model is derived from an IEEE standard model known as IEESGO in the PSS/E, as shown in Fig. 3. This model represents a typical thermal power plant governor. Although its parameters are based on [21], the upper limit, converted from MVA base value to MW base (from 0.9 pu to 1.0588 pu to include a 5.88% overload margin reflecting turbine governor models), is calculated as 0.9/0.85, with 0.85 being the rated power factor of synchronous generators and synchronous condensers based on design data from the early 1970s. Also, the lower limit is set at 0.0, ignoring the in-house loss and mechanical loss.

3) Grid Protection Model

a: Underfrequency Protection

The grid protection for frequency stability generally detects the frequency drop using underfrequency relays with or without timer or rate-of-change-of-frequency relays. The developed dynamic model employs the underfrequency re-

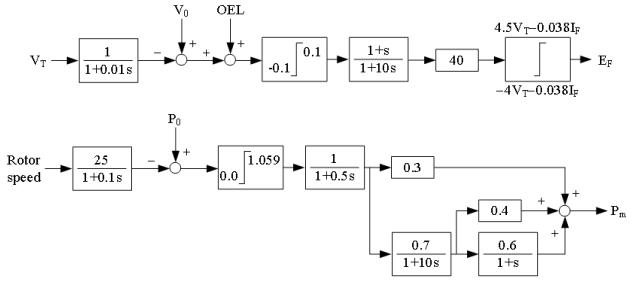


FIGURE 3. Generator controller: automatic voltage regulator model (top) and turbine-governor model (bottom).

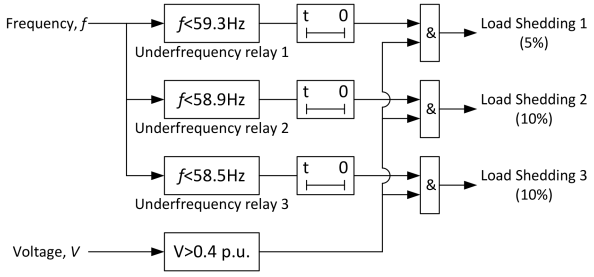


FIGURE 4. Underfrequency relay logic.

lay without the timer as the grid protection for frequency stability. This decision prioritizes simplicity in the modeling process, as determining timer settings requires a comprehensive study of various generation dispatching scenarios. Relay setting values are generally determined according to the following principles:

- The frequency nadir is not lower than the operating frequency (lower) limit of synchronous generators.
- The total load shedding amount is approximately equal to the assumed capacity of the largest single generator that could trip.
- The post-fault frequency should not exceed the steady-state frequency.

Based on the above considerations, referring to a publicly available source [34], the underfrequency protection has been modeled as shown in Fig. 4. The underfrequency relays are designed to remain inactive when the voltage level falls below 0.4 pu.

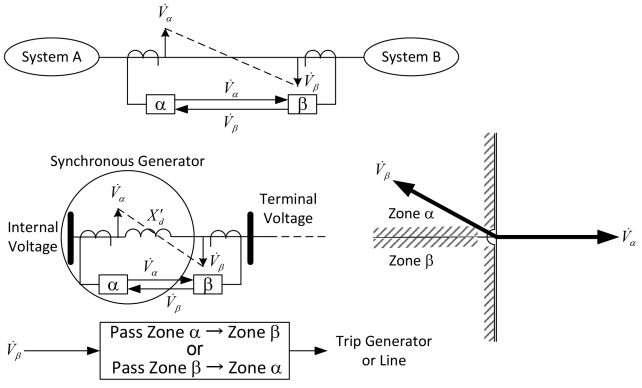


FIGURE 5. Out-of-step relay characteristics for transmission lines and generators.

b: Out-of-step Relay

The grid protection for transient stability aims to prevent cascading failures by disconnecting out-of-step (OOS) synchronous generators or tie-lines. OOS relays are designed based on impedance or voltage angle difference [35], though they are not always placed on all transmission lines in many countries. The former relay is used for unit protection, while the latter relay is used for grid/line protection. For simplicity, the voltage angle difference is used for both lines and generators (see Fig. 5). This decision emphasizes simplicity in the modeling process, given that determining impedance settings necessitates thorough consideration of diverse generation dispatch scenarios typically handled by grid operators. As time-domain simulation tools calculate the internal induced voltage, the angle difference between terminal voltage and internal induced voltage is used to detect OOS generators.

c: Load Model

Load characteristics are represented using the exponential load model with the voltage and frequency dependency as shown in (7). The coefficients of load frequency characteristics are set based on CIGRE working group report [36].

The load self-disconnection characteristics are equivalent to that of the under-voltage load-shedding system at the system level, expressed as (8) [36]. The starting and saturated voltage levels are set as 0.8 p.u. and 0.4 p.u., respectively. The amount of self-disconnected loads is assumed to increase linearly with the upper limit of 25% of the initial load. It is noted that this relay is not applied to reactive power compensators.

$$\begin{cases} P = (P_0 - P_{\text{drop}}) \left(\frac{V}{V_0}\right) \left(1 + \frac{3.33}{100}\Delta f\right) \\ Q = (Q_0 - Q_{\text{drop}}) \left(\frac{V}{V_0}\right)^2 \left(1 + \frac{0}{100}\Delta f\right) \end{cases} \tag{7}$$

$$P_{\text{drop}}, Q_{\text{drop}} = \begin{cases} 0 & (V_{\min} > 0.8) \\ -1.6V_{\min} + 0.8 & (0.8 \geq V_{\min} \geq 0.4) \\ 0.25 & (0.4 > V_{\min}) \end{cases} \tag{8}$$

where P_{drop} and Q_{drop} denote the amount of load self-disconnection, and $V_{\min}$ denotes the lowest load bus voltage.

C. N-1 Contingency List

Contingency analysis is generally performed by disconnecting one or more power components from the grid. The single power component disconnection is commonly called N-1 contingency. When two power components are disconnected, it is referred to as N-2 contingency. As this study is an initial attempt to apply GCN/GAT for grid health index prediction, the scope of the contingency analysis is limited to N-1 contingencies. N-1 contingencies in the IEEE 118-bus system are listed by the power equipment, as shown in Table 4 and Fig. 6. All power flow snapshots, outputs (i.e., grid health indices), and the employed model parameters for the generator controller models and grid protection models are available in the GitHub repository: github.com/KojiGitCode/OAJPE.

TABLE 4. Number of N-1 Contingency

Type	Line	Unit	Transformer
Transmission line	163	N/A	N/A
Synchronous generator (SG)	N/A	19	N/A
Synchronous condenser (SC)	N/A	36	N/A
Step-up transformer for SG	N/A	N/A	19
Step-up transformer for SC	N/A	N/A	36
Tie transformer	N/A	N/A	9
Step-down transformer for load	N/A	N/A	91
Number of fault duration pattern	3*	N/A	3*
Total **	489	55	465

* Three fault durations are set to 50, 200, and 367 ms.
** Calculation: Numbers of same contingency types times 3 fault durations

V Grid Health Index Prediction Performance

As shown in Table 2, the number of unstable cases for angular stability ranges from 285 to 357, and if a 90-degree threshold for steady-state angle stability is considered, the number increases to between 459 and 884. Given the total number of cases (25,225), unstable cases constitute a relatively small fraction, resulting in a class imbalance.

To ensure a fair evaluation of the model's performance, especially for the minority unstable cases, the F1-score was used instead of accuracy. The F1-score balances both precision (correctly predicted unstable cases) and recall (actual unstable cases detected), providing a more reliable assessment of the model's ability to detect instability despite the class imbalance.

For this study, 60% of the dataset is assigned for the training dataset, then around 10% of the dataset is allocated for the validation dataset, and the remaining 30% of the dataset is leveraged for the testing dataset.

The generator pairs selected for the angle health index (buses 10 and 12, and buses 10 and 49) were chosen based on their largest pre-contingency angle differences among all generator pairs in the test system. These pairs are critical indicators of transient stability, as they tend to show significant angle deviations following a disturbance.

A. Prediction Accuracy with Full Grid Sensor Coverage

1) Prediction Performance of GCN and GAT Models

The grid health indices, specifically the frequency health index and angle health index, are predicted by the proposed GCN and GAT models. MLP models are selected as the *baseline* model. The validation losses and testing accuracies obtained from the MLP, GCN, and GAT models are shown in Table 5. Note that the hyperparameters, e.g., learning rate and the number of hidden layers and neurons, were fine-tuned and chosen based on validation dataset performance. An extension of the Adam optimizer, named Adamax, was employed as the optimizer, which provides better performance in terms of minimizing model error.

Table 5 shows that GNN models, such as GCN and GAT, improve grid health index prediction accuracy by more than 10% compared to the baseline model (i.e., MLP models). When comparing the performance between GCN and GAT models, it is observed that both models demonstrate nearly the same prediction accuracy for the frequency health

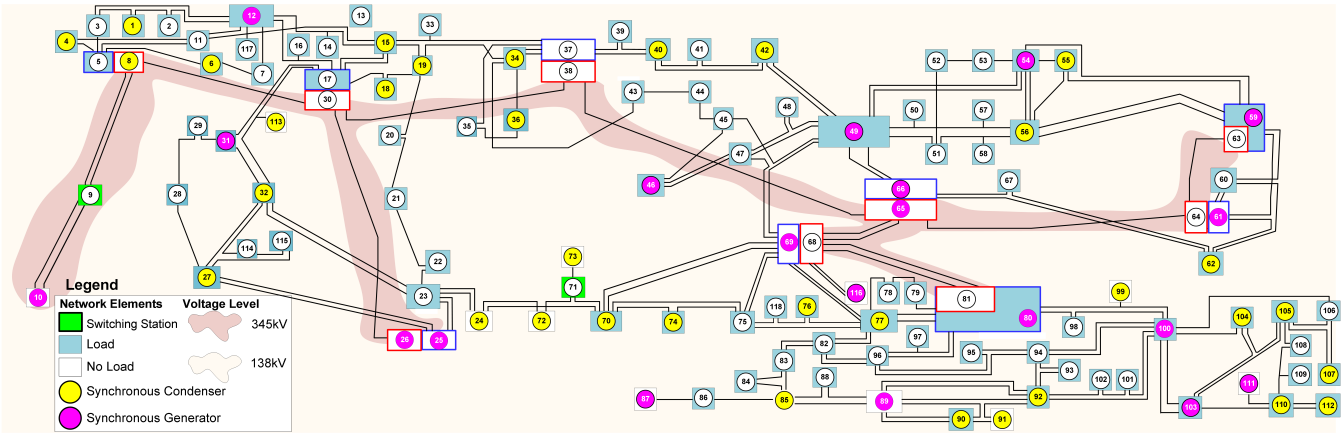


FIGURE 6. IEEE 118-Bus system single-line diagram with annotations of key power system components.

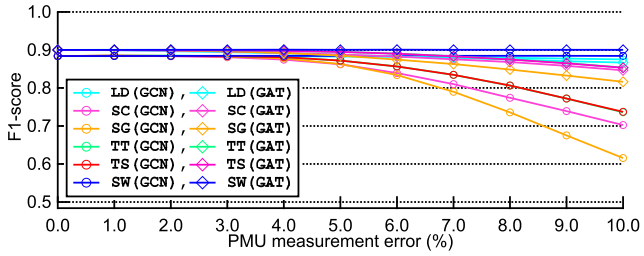


FIGURE 7. Accuracy for frequency health index prediction for 6 partial PMU coverage scenarios with MLP/GCN/GAT model, considering measurement error levels ranging from 0 to 10%.

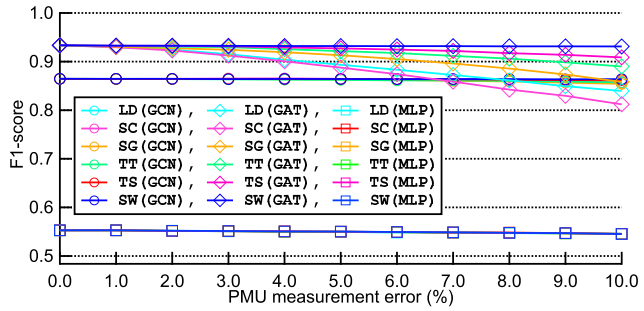


FIGURE 8. F1-score for angle health index prediction (Buses 10-49) for 6 partial PMU coverage scenarios with MLP/GCN/GAT model, considering measurement error levels ranging from 0 to 10%.

index prediction. However, GAT models exhibit significantly higher accuracy than GCN models for the angle health index. These differences may be attributed to the GAT model's superior ability to handle edge features, which are more crucial for angular stability than for frequency stability.

2) Impact of Ordinal Encoder on Prediction Accuracy

The upper three rows of Table 5 show that the ordinal encoder does not always contribute to an increase in frequency health prediction accuracy for both GCN and GAT models. On the other hand, the lower six rows of Table 5 shows that the ordinal encoder contributes to an increase in angle health prediction accuracy for both GCN and GAT models, by $2.3\sim3.2\%$ and $0.1\sim2.7\%$, respectively. While the latter range appears smaller, the rate of accuracy improvement for the angle health prediction relative to the baseline model is higher in GAT models compared to GCN models.

In summary, both the attention mechanism and the ordinal encoder enable GNNs to predict the grid health indices more accurately.

3) Assessing Real-Time Applicability of GAT Models in Grid Health Index Prediction

The GAT models are tested on a system running Ubuntu, equipped with an Intel i7-10700F CPU, 32 GB of RAM, and an NVIDIA GeForce RTX 4090 GPU with 24 GB of VRAM. With the system mentioned above, the training time is a few hours. However, the calculation times for the frequency and angle health indices are 1.22–1.27 ms and 2.89–2.98 ms, respectively. The computation times are for a single contingency and calculated based on 100 attempts across all testing datasets, including data loading, model initialization, and health index prediction, but excluding communication delays involved in acquiring measurement data.

B. Partial PMU Coverage

Six scenarios with different levels of Phasor Measurement Unit (PMU) coverage in the power system were analyzed, as outlined in Table 6. Each scenario represents varying PMU coverage for distinct power equipment types (see second column in Table 6). This study enables system operators and planners to identify critical grid sensor locations for predicting the frequency and angle health index.

1) Pseudo PMU Measurement Setup

In the absence of grid sensors at specific bus locations, GCN and GAT models utilize pseudo measurements (approximated measurements that mimic the output of actual PMUs) derived from the state estimator. Notably, state estimator outputs, updated in intervals of $0.5\sim5$ minutes, had a much lower sampling frequency compared to the standard PMU measurement with a reporting rate of 33.33 ms that has been employed in the real world, e.g., WECC in the U.S. Consequently, state estimator-derived pseudo measurements are inherently noisy and delayed measurements. To emulate this characteristic in experiments, a Gaussian noise with a

TABLE 5. Frequency/Angle Health Index Prediction Performance (Buses 10-12 and 10-49)

Grid health index	Machine learning model	Validation loss ($\times 10^{-3}$)	Testing accuracy w/wo ordinal encoder	Testing F1-score w/wo ordinal encoder	Number of node neurons	Number of edge neurons	Number of hidden layers	Learning rate of Adam	Number of heads
Frequency	MLP	1.338 / 1.384	0.8192 / 0.8414	0.7279 / 0.7139	64	N/A	4	0.0001	N/A
	GCN	0.442 / 0.477	0.9302 / 0.9370	0.8841 / 0.9010	64	32	5	0.0001	N/A
	GAT	0.476 / 0.536	0.9392 / 0.9394	0.9001 / 0.8999	64	32	4	0.0001	4
Angle Index between buses 10-12	MLP	3.811 / 3.798	0.5151 / 0.5103	0.4670 / 0.4596	128	N/A	2	0.001	N/A
	GCN	1.316 / 1.834	0.8810 / 0.8488	0.8888 / 0.8620	192	32	2	0.001	N/A
	GAT	1.228 / 1.571	0.8992 / 0.8980	0.9123 / 0.9023	64	32	3	0.0001	2
Angle Index between buses 10-49	MLP	3.712 / 3.880	0.5365 / 0.5119	0.5529 / 0.4966	64	N/A	5	0.001	N/A
	GCN	1.640 / 1.714	0.9056 / 0.8821	0.9009 / 0.8806	192	32	4	0.001	N/A
	GAT	0.836 / 1.370	0.9389 / 0.9115	0.9334 / 0.8982	64	32	4	0.0001	2

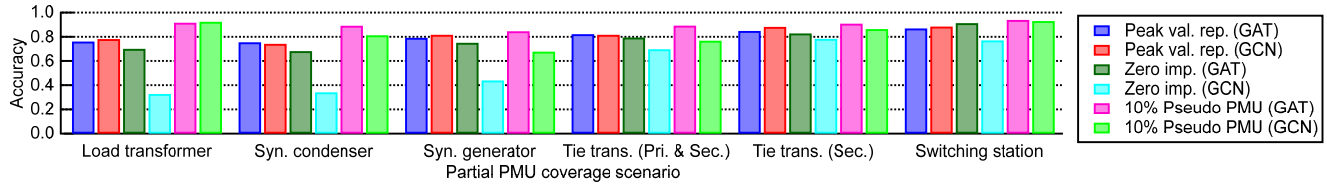


FIGURE 9. F1-score for frequency health index prediction for 6 partial PMU coverage scenarios with GCN and GAT models, using 3 imputation methods.

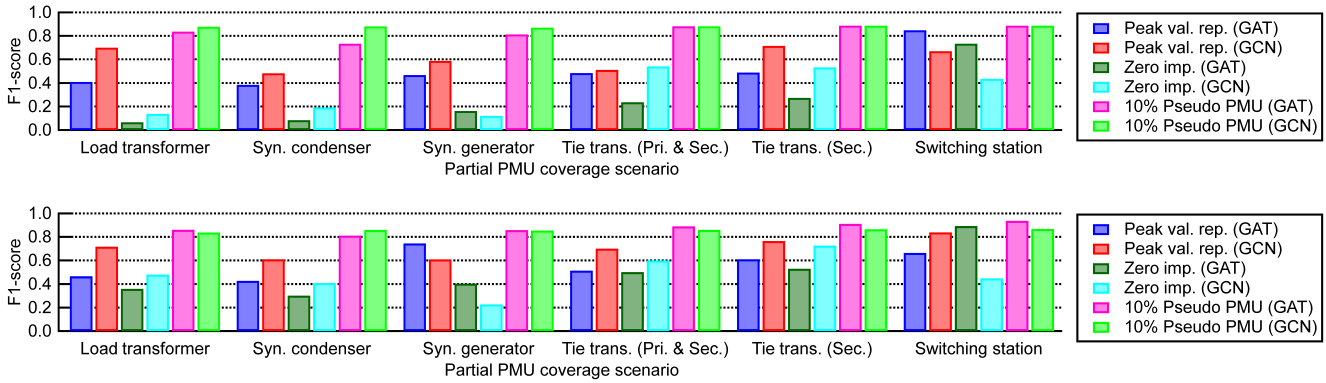


FIGURE 10. F1-score for angle health index prediction (top subgraph: angle difference between Buses 10-12, bottom subgraph: Buses 10-49) for 6 partial PMU coverage scenarios with GCN and GAT models, using 3 imputation methods.

standard deviation ranging from $1/30$ to $1/3$ is introduced into dynamic simulation results. It should be noted that a standard deviation of $1/30$ is assumed to correspond to a $\pm 1\%$ measurement error. In addition, two naive missing value replacement methods, i.e., zero imputation method and peak demand value replacement method, are also examined.

2) Analysis of Grid Health Prediction Performance Across Different Pseudo PMU Measurement Error Levels

The grid health index prediction performance with pseudo grid sensor measurement errors from 0 to 10% is shown in Figs. 7 and 8. The F1-scores were calculated as a performance indicator using the Torchmetrics library in PyTorch, which computes it as the harmonic mean of precision and recall. The F1-scores in Figs. 7 and 8, gradually decline as the measurement error increases. However, the rate of decrease is much greater when using GAT models. The prediction performance of GCN models surpasses that of GAT models for certain partial PMU coverage scenarios when the measurement error exceeds 7%. This observation in prediction performance seems to suggest a trade-off between robustness and accuracy for a prediction model.

TABLE 6. Partial PMU Coverage Scenario in the IEEE 118-Bus System

Index of buses with no PMU (abbrev.)	Power equipment	Remark
2 (SW)	Switching station	—
9 (TS)	Tie-transformer	Secondary-side only
18 (TT)	Tie-transformer	Both sides
19 (SG)	Synchronous generator	—
35 (SC)	Synchronous condenser	—
53 (LD)	Step-down transformer	Load bus with no unit

3) Enhancing Grid Health Prediction Performance under Partial PMU Coverage with Data Imputation Techniques

The performance of grid health index prediction with three different missing sensor data imputation techniques is shown in Figs. 9 and 10. The three data imputation techniques evaluated are zero value imputation, peak value replacement, and pseudo grid sensor measurement with 10% error. It can be seen from the figures that under most partial PMU coverage scenarios, both the zero imputation method and peak value replacement method perform worse than pseudo-PMU measurements with 10% error. The zero imputation method could significantly deteriorate the grid health prediction performance when a large area lacks PMUs or when critical substations miss PMUs. Finally, the results show

that GCN models demonstrate greater tolerance to limited PMU observability compared to GAT models when assessing frequency stability. In contrast, GCN models exhibit lower tolerance to limited PMU observability compared to GAT models when assessing angular stability. The last finding reveals a potential weakness in the performance of the GAT/GCN model for grid health prediction.

4) Critical Location for Grid Sensor

The numerical study results in figures 7 and 8 provide three insights for grid sensor placement for grid health index prediction. First, we observe that grid sensors should be consistently present at substations with synchronous generators when monitoring the frequency stability-related health index. Second, grid sensors must be consistently deployed at substations with synchronous condensers when monitoring the angular stability-related health index. Third, grid sensors should be reliably deployed at substations with step-down transformers connected to load feeders when monitoring angular stability-related health. Given that many step-down transformers are co-located with step-up transformers and transmission lines within the same substation, a single PMU with multiple measurement channels can aggregate data across interconnected components. Additionally, by placing the grid sensor on the secondary side of each step-down transformer, it can monitor and aggregate data from multiple load feeders connected to the same transformer. This approach reduces the need for dedicated PMUs at each step-down transformer, streamlining monitoring infrastructure while enhancing data coverage. The second finding implicitly reveals the importance of both active power and reactive power in assessing angular stability. In summary, we must consider the type of power equipment in a substation when identifying ideal locations of grid sensors for monitoring grid health indices.

VI Conclusion

This paper introduces a physics-informed graph learning model for real-time assessment of grid health index, focusing on frequency and angular stability studies. Comprehensive numerical study results reveal that the proposed graph convolutional network (GCN) and graph attention (GAT) models significantly outperform the multi-layer perception model in terms of prediction accuracy by $12\sim40\%$ when considering full PMU coverage. The study results also show that the GAT model boosted the angle health index prediction accuracy by $1 \sim 3\%$ compared to the GCN model. Both models achieve similar accuracy for frequency health index under full PMU observability. The inclusion of an ordinal encoder improves accuracy between 0.1% and 2.7% for GAT models for predicting angle and frequency health indices. However, the ordinal encoder only improves accuracy for GCN models $(2.3\sim3.2\%)$ when predicting the angle health index.

The numerical study results highlight the importance of selecting appropriate models and sensor data imputation methods for grid health index prediction under partial PMU coverage scenarios. In practice, one should avoid utilizing the zero imputation and peak value replacement methods. Furthermore, GCN models not only often yield better performance for frequency stability analysis but also exhibit greater resilience to limited PMU observability compared to GAT models. In contrast, GAT models achieve superior performance for angular stability analysis with partial PMU coverage, demonstrating higher tolerance to limited PMU observability than GCN models.

The numerical studies of this paper show that the placement of grid sensors at substations with critical power equipment is crucial for improving grid health index prediction algorithms. Specifically, strategically placing sensors at these critical locations enhances the data quality used by machine learning models, improving their accuracy in predicting grid health. It is important to note that the criticality of power equipment varies depending on the type of stability analysis being conducted, and understanding these variations enables machine learning models to identify and prioritize the most important features for accurate prediction. These observations can guide system operation engineers in developing more effective and efficient grid health prediction algorithms, tailored to real-world power grid conditions.

Future work will first focus on incorporating a wider range of grid configurations, including alternative topologies and off-peak maintenance scenarios, into the training dataset. This approach will enhance the model's ability to generalize to rare configurations that may arise in real-world power systems. Additionally, real-world demonstrations with utility and network operators will be conducted to implement the proposed graph neural network-based real-time contingency analysis tool for monitoring grid health conditions.

Acknowledgment

This work was supported by the Electric Power Research Institute under Agreement No. 10015026.

REFERENCES

[1] P. Sauer, K. Tomsovic, and V. Vittal, "Dynamic security assessment," in *Power System Stability and Control*, L. L. Grigsby, Ed. CRC Press, 2007, pp. 223–246.

[2] J. Baranowski and D. J. French, "Operational use of contingency analysis at PJM," in *Proc. 2012 IEEE PES GM*, San Diego, CA, USA, July 2012.

[3] CIGRE Working Group C2-C4.37, "A proposed framework for coordinated power system stability control," CIGRE, Tech. Rep. TB742, Sep. 2018.

[4] S. Cao, N. Lin, and V. Dinavahi, "Faster-than-real-time hardware emulation of extensive contingencies for dynamic security analysis of large-scale integrated AC/DC grid," *IEEE Trans. Power Syst.*, vol. 38, no. 1, pp. 861–871, Jan. 2023.

[5] S. Huang and V. Dinavahi, "Real-time contingency analysis on massively parallel architectures with compensation method," *IEEE Access*, vol. 6, pp. 44 519–44 530, Aug. 2018.

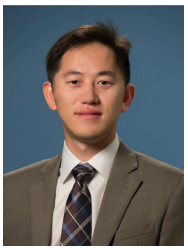
[6] S. Yang, B. Vaagensmith, and D. Patra, "Power grid contingency analysis with machine learning: A brief survey and prospects," in *Proc. of 2020 Resilience Week*, Salt Lake City, UT, USA, Oct. 2020.

[7] W. Liao, B. Bak-Jensen, J. R. Pillai, Y. Wang, and Y. Wang, "A review of graph neural networks and their applications in power system," *J. Mod. Power Syst. Clean Energy*, vol. 10, no. 2, pp. 345–360, Mar. 2022.

[8] T. Zhao, M. Yue, and J. Wang, "Structure-informed graph learning of networked dependencies for online prediction of power system

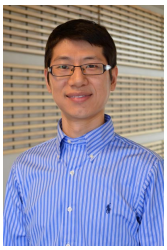
transient dynamics," *IEEE Trans. Power Syst.*, vol. 37, no. 6, pp. 4885–4895, Jan. 2022.
[9] J. Huang, L. Guan, Y. Su, H. Yao, M. Guo, and Z. Zhong, "System-scale-free transient contingency screening scheme based on steady-state information: A pooling-ensemble multi-graph learning approach," *IEEE Trans. Power Syst.*, vol. 37, no. 1, pp. 294–305, Jan. 2022.
[10] M. Dezvarei, K. Tomsovic, J. S. Sun, and S. M. Djouadi, "Graph neural network framework for security assessment informed by topological measures," 2023, arXiv preprint. [Online]. Available: https://arxiv.org/abs/2301.12988
[11] K. Yamashita, N. Yu, E. Farantatos, and L. Zhu, "Predicting power system frequency health index with PMUs and graph attention networks," in *2024 IEEE SmartGridComm*, pp. 353–358.
[12] K. Yamashita, J. Qin, N. Yu, E. Farantatos, and L. Zhu, "Predicting power system voltage health index with PMUs and graph convolutional networks," in *2023 IEEE PESGM*, pp. 1–5.
[13] H. Ma, H. Tian, D. Cheng, K. Zhao, F. Xing, H. Qi, R. Su, and C. Li, "Data-driven multiple-task dynamic security assessment and its application on optimization of emergency control strategy for frequency," in *2023 Int. Conf. on Power System Tech. POWERCON*. Jinan, China: IEEE, Sep. 2023, pp. 1–5.
[14] M. Netto, V. Krishnan, Y. Zhang, and L. Mili, "Measurement placement in electric power transmission and distribution grids: Review of concepts, methods, and research needs," *IET Generation, Transmission & Distribution*, vol. 36, no. 5, pp. 805–838, Oct. 2021.
[15] M. M. Ahmed, M. Amjad, M. A. Qureshi, K. Imran, Z. M. Haider, and M. O. Khan, "A critical review of state-of-the-art optimal PMU placement techniques," *Energies*, vol. 15, no. 6, Mar. 2022.
[16] A. Pal, A. K. S. Vullikanti, and S. S. Ravi, "A PMU placement scheme considering realistic costs and modern trends in relaying," *IEEE Trans. Power Syst.*, vol. 10, no. 2, pp. 345–360, Mar. 2022.
[17] T. Chen, H. Ren, Y. Sun, M. Kraft, and G. A. J. Amaratunga, "Optimal placement of phasor measurement unit in smart grids considering multiple constraints," *J. Mod. Power Syst. Clean Energy*, vol. 11, no. 2, pp. 479–488, Mar. 2023.
[18] J. D. Lara-Jimenez, J. M. Ramırez, and F. Mancilla-David, "Allocation of PMUs for power system-wide inertial frequency response estimation," *IET Generation, Transmission & Distribution*, vol. 11, no. 11, pp. 2902–2911, Aug. 2017.
[19] N. M. Manousakis and G. N. Korres, "Optimal allocation of phasor measurement units considering various contingencies and measurement redundancy," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, 2020.
[20] N. P. Theodorakatos, R. Babu, C. A. Theodoridis, and A. P. Moschoudis, "Mathematical models for the single-channel and multi-channel PMU allocation problem and their solution algorithms," *Algorithms*, vol. 17, no. 191, 2024.
[21] G. N. Korres, N. M. Manousakis, T. C. Xygkis, and J. Löfberg, "Optimal phasor measurement unit placement for numerical observability in the presence of conventional measurements using semidefinite programming," *IET Gener. Transm. Distrib.*, vol. 9, no. 15, 2015.
[22] B. Gou, "Generalized integer linear programming formulation for optimal PMU placement," *IEEE Trans. Power Syst.*, vol. 23, no. 3, 2008.
[23] C. Mishra, K. D. Jones, A. Pal, and V. A. Centeno, "Binary particle swarm optimisation-based optimal substation coverage algorithm for phasor measurement unit installations in practical systems," *IET Gener. Transm. Distrib.*, vol. 10, no. 2, 2016.
[24] N. Fan and J.-P. Watson, "On integer programming models for the multi-channel PMU placement problem and their solution," *Energy Syst.*, vol. 6, 2015.
[25] A. Pal, G. A. Sanchez-Ayala, V. A. Centeno, and J. S. Thorp, "A PMU placement scheme ensuring real-time monitoring of critical buses of the network," *IEEE Trans. Power Del.*, vol. 29, no. 2, 2014.
[26] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. of 5th Int. Conf. on Learning Representations (ICLR)*, Apr. 2017.
[27] M. Simonovsky and N. Komodakis, "Dynamic edge-conditioned filters in convolutional neural networks on graphs," in *Proc. 2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2017.
[28] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *6th Int. Conf. Learn. Represent. ICLR 2018, Vancouver, BC, Canada*.
[29] S. Brody, U. Alon, and E. Yahav, "How attentive are graph attention networks?" in *10th Int. Conf. Learn. Represent. ICLR 2022, Virtual, April 25 - 29, 2022*.
[30] Power Systems Test Case Archive. (Aug., 1993) 118 bus power flow test case. University of Washington. [Online]. Available: https://labs.ece.uw.edu/pstca/pf118/pg_tca118bus.htm
[31] CIGRE Working Group C4.503, "Power system test cases for EMT-type simulation studies," CIGRE, TB 736, Aug. 2018.
[32] A. Anderson, S. Kincic, B. Jefferson, B. Mcgary, C. Fallon, D. Ciesielski, J. Wenskovitch, and Y. Chen, "A real-time operation manual for the IEEE 118 bus transmission model," PNNL, Tech. Rep. 33499, Sep. 2022.
[33] Power System Stability Study Group, "Integrated analysis software for bulk power system stability," CRIEPI, Tech. Rep. ET90002, Jul. 991.
[34] Electric Reliability Council of Texas. (2018, Jun.) ERCOT nodal operating guides, section 2: System operations and control requirements. [Online]. Available: http://www.ercot.com/content/wcm/current_guides/53525/02-060118.doc
[35] CIGRE Working Group B5.19, "Protection Relay Coordination," CIGRE, Tech. Rep. TB432, Oct. 2010.
[36] CIGRE Working Group C4.605, "Modelling and aggregation of loads in flexible power networks," CIGRE, Tech. Rep. TB566, Feb. 2014.

Koji Yamashita(M'04) received the B.S. and M.S. degrees in Electrical Engineering from Waseda University, Japan, in 1993 and 1995, respectively. He also received his Ph.D. degree in Electrical and Computer Engineering from Michigan Technological University, MI, USA, in 2020. He was a research scientist with the power system division of CRIEPI, Japan in 1995–2018. He is currently a postdoctoral scholar at University of California Riverside.



Nanpeng Yu(M'11-SM'16) received his B.S. in Electrical Engineering from Tsinghua University, Beijing, China, in 2006. Dr. Yu also received his M.S. and Ph.D. degree in Electrical Engineering from Iowa State University, Ames, IA, USA in 2007 and 2010 respectively. He is a Professor in the Department of Electrical and Computer Engineering at University of California, Riverside, CA, USA. His research interests include machine learning and optimization for smart grid and electricity market design. Dr. Yu is an Editor of IEEE Power Engineering Letters and Open Access Journal of Power and Energy.

Evangelos Farantatosreceived the Diploma in Electrical and Computer Engineering from the National Technical University of Athens, Greece, in 2006 and the M.S. and Ph.D. degrees from the Georgia Institute of Technology, Atlanta, GA, USA, in 2009 and 2012, respectively. He is a Sr. Principal Team Lead with the Transmission Operations and Planning R&D Group at EPRI. He is managing and leading the technical work of various R&D projects related to synchrophasor technology, power systems monitoring and control, power systems stability and dynamics, renewable energy resources modeling, grid operation protection and control with high levels of inverter-based resources. He is a Senior Member of IEEE.



Lin Zhu(S'07–M'11-SM'20) received the B.S. and Ph.D. degrees in electrical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2005 and 2011, respectively. He was a Research Assistant Professor with the Department of Electrical Engineering and Computer Science at the University of Tennessee, Knoxville, from 2017 to 2021. He is currently a Senior Technical Leader at EPRI, USA.