

Highlights

Graph neural networks for operational risk assessment under evolving grid topology

Yadong Zhang, Pranav M Karve, Sankaran Mahadevan

- Graph neural network as a computationally efficient optimization proxy for hours-ahead operational risk assessment
- Reliability and risk assessment at system, zone and branch levels by considering evolution of grid state over the next few hours
- Load shedding and branch overloading as failure modes
- Methodology demonstrated using large, synthetic power grids (Case1354pegase, Case2848rte)

Graph neural networks for operational risk assessment under evolving grid topology

Yadong Zhang^a, Pranav M Karve^a, Sankaran Mahadevan^{a,*}

^a*Department of Civil and Environmental Engineering, Vanderbilt University, 400 24th Avenue South, Nashville, 37212, TN, USA*

Abstract

This article investigates the ability of graph neural networks (GNNs) to identify risky conditions in a power grid over the subsequent few hours, without explicit, high-resolution information regarding future generator on/off status (grid topology) or power dispatch decisions. The GNNs are trained using supervised learning, to predict the power grid's aggregated bus-level (zonal or system-level) or individual branch-level state under different power supply and demand conditions. The uncertainty in, and correlations among, the stochastic grid variables (wind/solar generation and load demand) are rigorously considered while generating the inputs for training data. The outputs in the training data, obtained by solving numerous mixed-integer linear programming (MILP) problems, correspond to system-level, zonal and transmission line-level quantities of interest (QoIs). QoIs predicted by the GNNs are used to conduct hours-ahead, sampling-based reliability and risk assessment w.r.t. zonal and system level (load shedding) as well as branch level (overloading) failure events. The proposed methodology is demonstrated for three synthetic grids with sizes ranging from 118 to 2848 buses. Our results demonstrate that GNNs are capable of providing fast and accurate prediction of QoIs and can be good proxies for computationally expensive MILP algorithms. The excellent accuracy of GNN-based reliability and risk assessment suggests that GNN models can substantially improve situational awareness by quickly providing rigorous reliability/risk estimates.

Keywords: Power grid, uncertainty, graph neural network, reliability, risk

*Corresponding author

Email address: mahadevan.sankaran@vanderbilt.edu (Sankaran Mahadevan)

Nomenclature

\mathcal{G}	Graph
\mathcal{E}	Set of edges in a graph
\mathcal{V}	Set of nodes in a graph
$ \cdot $	Cardinality of a set
u	Node in a graph
\mathbf{h}_u^k	Embedding of u in k -th layer
\mathbf{m}_u	Aggregated node embedding from neighbours of u
$\mathcal{N}(u)$	Set of 1-hop neighbours of node u
N	Number of samples
M	Number of marginal distributions
Q	Number of branches in a power grid
T	Number of time steps in the reliability unit commitment horizon
ΔT	Number of time steps in the multi-step reliability/risk assessment
$\epsilon_{\mathcal{N}}$	Random matrix, each column follows $\mathcal{N}(0, 1)$, $\mathbb{R}^{N \times M}$
\mathbf{x}_t	Random matrix, each column follows $\mathcal{N}(0, t)$, $\mathbb{R}^{N \times M}$
\mathbf{x}_t^c	Random vector with covariance between columns specified by \mathbf{C} , $\mathbb{R}^{N \times M}$
\mathbf{u}_t	CDF of \mathbf{x}_t^c , $[0, 1]^{N \times M}$
\mathcal{W}_i	i -th marginal distribution
\mathbf{C}	Covariance matrix, $\mathbb{R}^{M \times M}$
\mathbf{L}	Lower triangular matrix of Cholesky factorization of \mathbf{C} , $\mathbb{R}^{M \times M}$
Φ	CDF of standard normal probability distribution
$\Phi_{\mathcal{W}_i}$	CDF of the i -th marginal distribution
Ψ_t	Indicator variable $\{0, 1\}$ for load shedding at time t
$\psi(t)$	The amount of load shed at time t (in MW)
$C_s(\psi)$	Consequence cost function for load shedding
Γ_t^i	Indicator variable $\{0, 1\}$ for branch i overloading at t
γ	Power flow in a transmission line (network branch)
γ_{max}	Maximum allowed branch power flow
ϵ	Fraction of γ_{max} that defines the security threshold of branch power flow
$\tilde{\gamma}$	The amount of branch overloading, $\tilde{\gamma} = \gamma - \epsilon \gamma_{max}$
$C_o(\tilde{\gamma})$	Consequence cost function for branch overloading

1. Introduction

The power grid is poised for a major transformation due to the increasing participation of renewable energy sources (RES), plugin devices, and flexible loads [1, 2, 3]. This paradigm shift, while pivotal in steering towards a more sustainable and diversified energy future, introduces substantial uncertainty in grid behavior [4, 5, 6, 7]. To effectively handle these challenges, grid operators are in urgent need of (i) optimization techniques that take into account uncertainty [8, 9, 10, 11] and (ii) comprehensive risk quantification methods [12, 13, 14, 15, 16, 17]. Stochastic optimization techniques are essential in the complicated grid operation tasks like unit commitment (UC) and economic dispatch (ED) [18, 19, 20] in the presence of uncertainty. On the other hand, risk quantification methods are required to understand risks associated with the increasing uncertainty in the grid variables [21, 22, 23, 24]. By assessing hours-ahead risks, grid operators can develop strategies to enhance grid safety and ensure uninterrupted power supply at optimal costs even under volatile conditions.

Mixed-integer linear programming (MILP) is a widely used decision-making algorithm for performing UC and ED tasks [25, 26, 27, 28]. Methods such as stochastic programming, robust optimization, and chance-constrained optimization are commonly employed for making decisions under uncertainty [29, 30, 31, 32]. These methods come with their own sets of challenges. Unlike UC or ED, which provide a decision for each generation unit, reliability and risk assessment in power systems has often been done at a more aggregated level, typically at the system or zone level [33, 34, 35, 36]. A range of potential contingencies, including loss of load [37, 38, 39, 40], reserve inadequacy [41, 42, 43], reserve inflexibility [44, 45], insufficient ramp capacity [46, 47], etc., have been considered in previous work. By focusing on the system or zone level, the corresponding risk assessment provides a macroscopic view of the grid's operational state. It enables operators to identify vulnerabilities, allocate resources more effectively, and develop strategies that enhance the overall stability and safety of the power network. The zonal/system level risk assessment requires computing the probability distribution of grid states corresponding to the given (forecast) joint probability distribution of stochastic grid variables (wind/solar generation and load). This is accomplished by drawing Monte Carlo (MC) samples of stochastic grid variables from the forecast distribution and solving a deterministic MILP problem to obtain the grid state for each MC sample. This process, however, is computationally demanding and time consuming due to the computational cost of solving numerous MILP problems. The granularity afforded by MILP solution (bus-level grid states) is also not needed for zonal/system

level risk analysis. It is therefore important to develop computational tools that can directly predict system/zone-level quantities (i.e., without solving the MILP problem) and thus enable fast and accurate reliability and risk assessment [48, 49]. Such tools are expected to predict the grid state corresponding to numerous future (hours-ahead) forecast scenarios of the stochastic grid variables, without explicitly solving the stochastic optimization problems. They are also expected to be aware of and to account for the intricate interdependencies within the power grid, while providing accurate and rapid future grid state predictions [50, 51].

Fast and accurate prediction of future grid behavior under different forecast scenarios requires learning the patterns of grid behavior under different conditions. Neural networks, particularly graph neural networks (GNNs), emerge as attractive options for this task due to their superior ability to handle graph-structured information, like that encountered in power grid analysis [52, 53, 54]. The core principle of GNN is to learn nodal features within a graph by assimilating information from neighboring nodes via message passing [55, 56]. This unique capability has led to GNNs being used in various fields, including recommender system [57], social network analysis [58], knowledge graph completion [59], etc. In the power grid domain, GNN has been used for fault diagnosis [60], power outage prediction [61], line flow control [62], load or generation forecasting [63], and so on. In recent work [64, 65, 66], GNN has also been used as a surrogate for optimal power flow computation with a fixed (known) generator on/off status (UC). To the best of our knowledge, however, the utilization of GNNs in performing hours-ahead prediction of system-, zone- and branch-level quantities has not been previously reported. Such predictions are necessary for performing hours-ahead risk assessment. These predictions need to account for future operator actions corresponding to the given probabilistic forecast. For example, if the forecast shows higher than expected load demand, then the zonal/system level grid variables need to be predicted by anticipating the operator bringing more generators online to help with increasing demand. If the future operator actions are not considered, then the risk assessment will be inaccurate.

In this work, we focus on hours-ahead prediction of quantities of interest (QoIs) at system, zone and branch levels. Accurate GNN-based system-level/zonal or branch-level reliability and risk prediction will help operators tune their intra-day reliability unit commitments (RUCs) [67, 68, 69, 70, 71, 72] and ensure grid safety. At the system/zone level, we consider thermal power generation and load shedding as QoIs, while transmission line flow is considered as QoI at the branch level. Multiple GNN models are trained with each focusing on the prediction of

one QoI. Fluctuation of grid variables in one zone is often related to the changes in other zones, leading to spatial correlation. Temporal correlations also emerge due to grid inertia and sequential nature of grid operations. These correlations must be meticulously accounted for in reliability and risk assessment. For that purpose, a large number of spatio-temporally correlated samples are drawn from probability distribution forecasting of RES power supply and load demand. These samples are then used to solve MILP problems and to obtain the grid state corresponding to these forecast scenarios. Following the supervised approach, the data (samples and numerical solutions) is split into training, validation and testing, and model evaluation is performed on testing data. The GNN-based QoI predictions are then used for reliability and risk quantification, which are compared with that obtained from ground truth, to evaluate the accuracy. Our main contributions include:

1. Investigation of GNN as models' utility to predict QoIs at system/zone/branch level under changing grid topology. This is the first investigation aimed at assessing the utility of GNNs for hours-ahead grid state prediction, without explicitly considering the future commitment/dispatch decisions (grid topology).
2. Development of novel reliability and risk quantification methods for *standalone* (at a given future time instance) and *multi-step* (over multiple future time instances) hours-ahead operational risk assessment. The developed methods can be applied in both system/zone and branch level. This is the first attempt that considers these two important aspects of hours-ahead grid risk quantification.
3. Development of a methodology for separating the influence of reserve-related and security constraints on load shedding risk quantification. Such cause-aware risk quantification can help improve operators' situational awareness and help illuminate various risk trade-offs.
4. Demonstration of the proposed methodology on medium to large sized synthetic power grids.
A methodology for generating spatio-temporally correlated samples of grid variables (synthetic forecasts) is developed and exercised for risk assessment of medium/large power grids.

The rest of the paper is organized as follows: Section 2 provides a brief introduction of hours-ahead operational decision-making in a power grid, fundamentals of graph neural network, and reliability/risk assessment metrics. In Section 3, we discuss the elements of the proposed methodology, including GNN model development, risk quantification, and training

data generation. Numerical experiments used to demonstrate the proposed methodology are described in Section 4 and the results are summarized in Section 5. The conclusion is drawn in Section 6

2. Background

100 2.1. Power grid and its operation

A power grid is a complex interconnected network that facilitates the generation, transmission, and distribution of electrical power to end consumers [10, 73]. The basic requirements of grid operation are balancing demand and supply, and at the same time, minimizing the overall power generation cost. For that purpose, grid operators usually schedule the on/off status of 105 generation units for a certain period (e.g., 12 hours) leveraging the hours-ahead planning techniques such as security-constrained unit commitment (SCUC) [74, 75]. Given the forecast of demand and supply, the outcome of SCUC plays a fundamental role in subsequent grid operation. For instance, due to its strategically important position in the operational framework, SCUC directly influences the calculation of economic dispatch, which aims at refining generation 110 schedule to every five to ten minutes [27, 76]. A precise and high-resolution solution in SCUC is essential to ensure optimal grid operations.

A comprehensive hours-ahead or intra-day grid risk assessment methodology needs to consider the evolving on/off status of the generators in response to the changing grid state. When sampling-based operational risk assessment methods are used, this necessitates solving thousands of SCUC problems corresponding to the forecast scenarios. This task is computationally 115 prohibitive for iteration-based numerical solvers. Note, however, that grid operators are often interested in risk quantification of the entire power grid or specific zones [33, 35, 36]. Therefore, a machine learning model that can directly predict system/zone-level quantities, without requiring explicit SCUC problem solutions, could be very useful for the desired reliability and 120 risk assessment. The objective of this paper is to investigate the performance of graph neural networks for this task.

2.2. Graph neural network (GNN)

Graph neural network (GNN) is a type of neural network that can process graph-structured data [77, 78, 54]. The core of GNN is message passing mechanism, which systematically aggregates and updates node features by assimilating information from the neighboring nodes [55, 56]. 125

Within the topology of a graph, nodes directly linked to a given node are designated as its *one-hop* neighbors. Extending this nomenclature, the immediate neighbors of these one-hop nodes are identified as *two-hop* neighbors, and so on. Each layer of a GNN model is dedicated to orchestrating the exchange of information between a node and its one-hop neighbors, and multiple layers can be stacked together to assimilate information from remote neighbors in the graph.

Message passing is comprised of *aggregation* and *updating*, which can be represented as:

$$\mathbf{m}_u = \mathbf{AGGREGATE}(\mathbf{h}_v^k), v \in \mathcal{N}(u), \quad (1)$$

$$\mathbf{h}_u^{(k+1)} = \mathbf{UPDATE}(\mathbf{h}_u, \mathbf{m}_u). \quad (2)$$

The application of GNN has witnessed remarkable success in recommender systems [57], knowledge graph enhancements [59], social network analyses (anomaly detection and community identification) [58], etc. Various GNN architectures have been developed by changing the aggregation and update functions. One of the earliest successful applications of GNN models utilized a graph convolutional network (GCN) [79], which considers symmetric-normalized aggregation with self-loop added as:

$$\mathbf{m}_u = \sum_{v \in \mathcal{N}(u) \cup u} \frac{\mathbf{h}_v}{\sqrt{|\mathcal{N}(u)||\mathcal{N}(v)|}}. \quad (3)$$

GCN has been successfully used in many application areas [77]. In power grids, GCN has been used for fault detection [80], outage prediction [81], etc. In a recent work [64], GCN is also used to develop a surrogate model for optimal power flow (OPF) analysis of the power grid. However, GCN requires to access the entire graph in every aggregation and update operation, which is computationally expensive and hinders its application to large graphs. To alleviate the computational burden, another category of GNNs that can work on partial graphs has been developed. GraphSAGE [82] is one such GNN. It generates embeddings by sampling a fixed-size neighborhood and then aggregating the features from these neighbors. GraphSAGE can be represented as:

$$\mathbf{m}_u = \text{CONCAT} \left(\mathbf{h}_u, \frac{1}{|\mathcal{N}(u)|} \sum_{v \in \mathcal{N}(u)} \mathbf{h}_v \right). \quad (4)$$

Note that instead of using the mean value of neighbors' state, max, min, or other complicated

functions could also be used. GraphSAGE does not require the entire graph to be loaded into the computer memory. It gradually samples nodes over the graph, making it attractive for the 150 learning task on large graphs. In this work, we explore GraphSAGE’s utility for learning the grid behavior pattern over a few hours into the future, for large grids (thousands of buses).

2.3. Reliability and risk metrics for the power grid

In the context of evaluating hours-ahead operational risk, [24] provides a comprehensive review of state-of-the-art reliability and risk quantification methods, and proposes a new sampling-based framework that encompasses various failure modes. This framework develops metrics at 155 three levels to characterize the risk: the conditional expectation of the failure event (Level 1), the probability of this event occurring (Level 2), and the consequence (monetary cost) (Level 3). Within the purview of this study, we place emphasis on the Level 2 and Level 3 metrics. The Level 2 metric captures the probability of system failure with respect to a failure mode, 160 e.g., reserve inadequacy, loss of load. The Level 3 metric adopts a more holistic view, not only gauging the likelihood of failure but also factoring in the economic implications or consequence costs linked to the adverse event. This dual consideration ensures a comprehensive understanding of the operational risk, melding both the probabilistic and economic dimensions of potential system failures. We use this framework for hours-ahead system/zone/branch level reliability 165 and risk prediction under *temporally evolving* grid topology (i.e., unit commitments).

3. Methodology

Three key elements of the proposed methodology, namely, GNN model development, reliability and risk assessment, and training data generation, are discussed in this section.

3.1. GNN model development

170 In this work, we utilize GNN models consisting of three components: *Encoder*, *GNN layers*, and *Decoder*. *Encoder* generates a high-dimensional representation (*embedding*) of the input vector, i.e., node features. This embedding acts as an enriched representation of the input data, capturing intricate patterns and relationships inherent in the input. Once the encoder processes the input, it is passed on to the *GNN layers*, which aggregate and update node 175 representations by incorporating information from neighboring nodes, allowing the network to understand the broader structure and context of the graph. Such aggregation aids in refining the node embedding by assimilating both local and global information from the graph. Finally, the

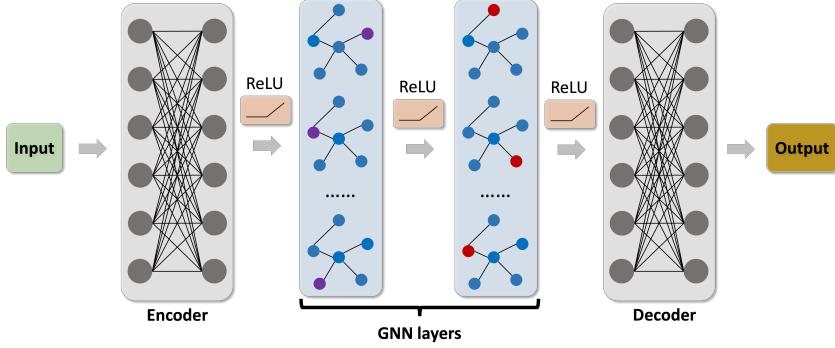


Figure 1: Schematic of the GNN surrogate model

Decoder takes these enhanced node representations and translates them into the desired output format. Depending on the specific application, the decoder might reconstruct the original data,
 180 predict certain properties of the graph, or even generate new graph structures. The combination
 of *Encoder*, *GNN layers*, and *Decoder* ensures that the GNN model can efficiently capture,
 process, and utilize graph-structured data. Note that both *Encoder* and *Decoder* consist of two
 185 fully connected neural layers (multi-layer perceptron, or MLP). Park et al. [69] have shown that
 the expressive power of GNN models can be improved by adding these fully connected linear
 layers. The architecture of the GNN surrogate model is shown in Fig. 1.

Three separate GNN models are developed for predicting different QoIs: (*i*) aggregated thermal power generation, (*ii*) aggregated load shedding, and (*iii*) branch power flow. The first two models perform graph-level prediction and output a single value at each time step; therefore, a graph pooling layer is added before the *Decoder* to generate global embedding. The
 190 third model (model *iii*) performs node-level prediction, i.e., net power generation at individual bus (+ for generation and - for demand), which can be converted to branch flow γ using the power transfer distribution factor (PTDF) matrix. A ReLU activation function is added after each neural network layer except for the last one. These GNN models take grid topology, load demand and RES power generation as inputs, and generate predictions of the QoIs. Following
 195 a supervised approach, numerical solutions of QoIs (from SCUC analysis) as used as ground truth in the training process. The loss function consists of mean squared error (MSE) and a regularization term to ensure that predictions are within reasonable upper and lower bounds.

3.2. Reliability and risk assessment

We consider two failure modes in reliability and risk assessment: *load shedding* and *branch
 200 overloading*. Load shedding refers to the deliberate reduction or shutdown of electrical power

supply to a consumer in a system to prevent the failure of the entire system, when demand strains the capacity of the system. Branch overloading implies that the electric current on branch exceeds a safety threshold and may lead to failure. We perform two types of hours-ahead reliability and risk assessments: (a) *standalone* hours-ahead assessment that considers QoIs at any given future time instant ($t + i$), and (b) *multi-step* hours-ahead assessment that considers QoIs at multiple time instances $((t+i), (t+i)+1, (t+i)+2, \dots)$ in order to provide an aggregated estimate over multiple future time instances. We use the variable ΔT to distinguish the reliability and risk aggregation horizon from the RUC planning horizon T .

3.2.1. Load shedding

Both standalone and multi-step load shedding are considered in load shedding analysis. Standalone shedding represents the immediate vulnerability within the system, highlighting areas where current demand exceeds the system's capacity or where there are breakdowns at a particular time instant. On the other hand, the multi-step shedding is a forward-looking measure that manifests potential pitfalls of the grid operation over multiple future time instants in the near future. This dual perspective establishes a comprehensive assessment of the system's reliability that can account for both current and future challenges. The probabilities of standalone and multi-step load shedding are defined as:

standalone:

$$\mathcal{P}_s(t) = p(\Psi_t = 1), \quad (5)$$

multi-step:

$$\mathcal{P}_s^+(t) = p\left(\bigcup_{t'=t+1}^{t+\Delta T} \Psi_{t'} = 1 \mid \Psi_t = 1\right), \quad (6)$$

respectively.

Load shedding usually leads to compromised service reliability, potential economic losses and safety concerns, hence there is consequence (monetary cost) associated with it. In [24], the degree of shedding is considered and the corresponding cost is defined as:

$$R_s(t) = \int_0^{\psi(t)} C_s(\psi) d\psi, \quad (7)$$

Note that $C_s(\psi)$ depends solely on ψ , while $\psi(t)$ is a function of t . The associated risk is then
225 quantified as the expectation of potential monetary cost:

$$\mathcal{R}_s(t) = \mathbb{E} [R_s(t)], \quad (8)$$

and

$$\mathcal{R}_s^+(t) = \sum_{\Delta t=1}^{\Delta T} \mathbb{E} [R_s(t + \Delta t)], \quad (9)$$

for standalone and multi-step load shedding, respectively. Here, we assume that the cost function $C_s(\psi)$ is temporally invariant. However, the consequence cost for load shedding events further away from the current time could be reduced by considering suitable, temporally diminishing cost functions, e.g., by using a weight factor $1/(1 + \Delta t)$ to modify $C_s(\psi)$. The overall
230 risk at t can be computed as:

$$\begin{aligned} \mathfrak{R}_s(t) &= \mathcal{R}_s(t) + \mathcal{R}_s^+(t) \\ &= \mathbb{E} [R_s(t)] + \sum_{\Delta t=1}^{\Delta T} \mathbb{E} [R_s(t + \Delta t)] \\ &= \sum_{\Delta t=0}^{\Delta T} \mathbb{E} [R_s(t + \Delta t)] \end{aligned} \quad (10)$$

3.2.2. Branch overloading

Branch overloading can result in large amount of heat generation that may weaken the insulating materials, accelerate the aging process of the conductors, and even lead to equipment
235 failure. Hence it is vital to monitor and manage branch loading status. Once again, we consider both standalone and multi-step risk. The respective probabilities are calculated as:

standalone:

$$\mathcal{P}_o^I(t) = p(\Gamma_t = 1), \quad (11)$$

multi-step:

$$\mathcal{P}_o^{I+}(t) = p \left(\bigcup_{t'=t+1}^{t+\Delta T} \Gamma_{t'} = 1 \mid \Gamma_t = 1 \right), \quad (12)$$

In real-world grid operations, electric current through any given branch is influenced by
240 the flows in other branches. This interdependence is due to the physical laws that govern electrical flow (Kirchhoff's laws). As a result, it is crucial to reveal the complex interactions between different branches. For that purpose, we further consider standalone (conditional) and multi-step (conditional) branch overloading. The corresponding probabilities are computed as:
standalone (conditional):

$$\mathcal{P}_o^{II}(t)|_{i,j} = \left[p\left(\Gamma_t^j = 1 | \Gamma_t^i = 1\right) \right]_{i,j}, \quad (13)$$

245 *multi-step (conditional)*:

$$\mathcal{P}_o^{II+}(t)|_{t',i,j} = \left[p\left(\bigcup_{t'=t+1}^{t+\Delta T} \Gamma_{t'}^j = 1 | \Gamma_t^i = 1\right) \right]_{t',i,j}. \quad (14)$$

The consequence of branch overloading, also with consideration of the degree of overloading, is defined as:

$$R_o(t) = \max \left\{ \int_0^{\tilde{\gamma}(t)} C_o(\tilde{\gamma}) d\tilde{\gamma}, 0 \right\}. \quad (15)$$

The risk of standalone and multi-step load shedding can be computed as:

$$\mathcal{R}_o^I(t) = \sum_{i=1}^Q \mathbb{E}[R_o(t)]_i, \quad (16)$$

$$\mathcal{R}_o^{I+}(t) = \sum_{\Delta t=1}^{\Delta T} \sum_{i=1}^Q \mathbb{E}[R_o(t + \Delta t)]_i, \quad (17)$$

and the overall risk can be obtained as:

$$\begin{aligned} \mathfrak{R}_o(t) &= \mathcal{R}_o^I(t) + \mathcal{R}_o^{I+}(t) \\ &= \sum_{i=1}^Q \mathbb{E}[R_o(t)]_i + \sum_{\Delta t=1}^{\Delta T} \sum_{i=1}^Q \mathbb{E}[R_o(t + \Delta t)]_i \\ &= \sum_{\Delta t=0}^{\Delta T} \sum_{i=1}^Q \mathbb{E}[R_o(t + \Delta t)]_i \end{aligned} \quad (18)$$

250 Note that it is not necessary to include the other two conditional probabilities in the calculation of risk. The interactions between different branches are reflected in branch flow; as a

result, their influence is already included in the risk computation above.

3.3. Training data generation

Demonstration of the proposed methodology requires probabilistic forecasts for the grid.
255 Since such forecasts are typically not available for synthetic power grids, we develop a method here for generating the same. We utilize an autoregressive model to account for temporal correlations for a given grid variable over the future time steps of interest. The spatial dependency between different grid variables is considered by specifying a covariance matrix of grid variables. By incorporating both temporal and spatial correlation, our methodology offers a robust frame-
260 work for a nuanced simulation of grid dynamics, which is valuable in demonstrating the utility of probabilistic grid behavior analysis methods. Given a planning horizon (time steps) T and probabilistic distributions \mathcal{W}_i ($i \in 1, 2, \dots, M$) of grid variables, the procedure for generating N spatio-temporally correlated training data samples for all time steps is shown in Algorithm 1.

Algorithm 1 Sampling spatio-temporally correlated stochastic grid variables

Inputs: #samples N , #time steps T , marginals $\mathcal{W}_1, \mathcal{W}_2, \dots, \mathcal{W}_M$, covariance matrix $\mathbf{C} \in \mathbb{R}^{M \times M}$

Start

do: $\mathbf{L} \leftarrow \mathbf{C} = \mathbf{LL}^T$, $\mathbf{x}_0 \leftarrow \mathbf{0} \in \mathbb{R}^{N \times M}$

for $t = 1:T$ **do:**

$\mathbf{x}_t \leftarrow \mathbf{x}_{t-1} + \epsilon_{\mathcal{N}}$, $\epsilon_{\mathcal{N}} \sim \mathcal{N}(0, 1)$

$\mathbf{s}_t \leftarrow \mathbf{x}_t / \sqrt{t}$

$\mathbf{x}_t^c \leftarrow (\mathbf{L}\mathbf{s}_t')'$

$\mathbf{u}_t \leftarrow \Phi(\mathbf{x}_t^c)$

End for

for $t = 1:T$ **do:**

for $i = 1:M$ **do:**

$\mathbf{w}_t^i \leftarrow \Phi_{\mathcal{W}_i}^{-1}(\mathbf{u}_t^i)$, $\mathbf{u}_t^i \in \mathbb{R}^N$

End for

End

Outputs: $\mathbf{W}_t = [\mathbf{w}_t^1 \ \mathbf{w}_t^2 \ \dots \ \mathbf{w}_t^M]$ consisting of stochastic variables with correlations specified by \mathbf{C} . (The N variables in the same column are from the same marginal probability distribution).

We also develop and use a methodology to generate load shedding cause-aware training
265 data to train GNN models that could identify the cause of (future) load shedding. There are two main causes of load shedding: reserve shortage and security constraint (transmission line flow limit) violations. Reserve shortage occurs when adequate backup power is not available to meet the demand in the event of a sudden drop in supply. Reserve shortage could occur due to lack of adequate power generation, inaccurate demand forecasting, or disruptions in fuel

270 supply. If the set of online generators cannot provide the minimum required reserve supply, the SCUC optimizer may shed load to ensure that sufficient reserves are available. Security constraints, on the other hand, refer to the physical limits of power transmission and distribution infrastructure. These constraints could be violated due to outdated equipment, maintenance issues, or bottlenecks in certain parts of the grid. Even if an adequate amount of power is
 275 being generated, it cannot be effectively distributed to all areas if security constraints are being violated. This leads to targeted load shedding to prevent equipment overload. When predicting future load shedding (associated risk), it is of operators' interest to know the cause of the problem (insufficient reserves or security constraints). This will allow operators to implement a suitable problem-specific mitigation strategy. We propose a methodology to identify the cause of
 280 load shedding in a MILP solution, when generating data to train the GNN models. Specifically, we propose to solve two MILP problems: (a) by enforcing all constraints, and (b) by relaxing reserve-related constraints. We record the load shedding proposed by the MILP solver in these two cases to separately quantify reserve shortage-related and security constraints-related load shedding for a given grid forecast (Fig. 2).

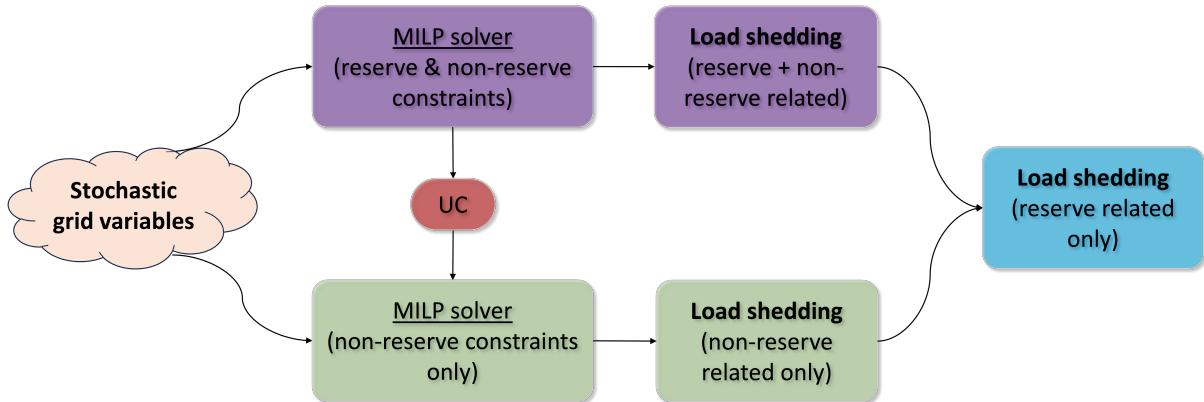


Figure 2: Separating the influence of reserve and non-reserve constraints on load shedding: training data generation

285 4. Numerical Experiments

We illustrate the utility of the GNN models using numerical experiments conducted on three synthetic grids: Case118 [83], Case1354pegase [84] and Case2848rte [85]. About 20% generators in each of the synthetic grids are taken to be RES generation units. The grids are partitioned into three, eight and sixteen zones, respectively, and the implementation details can be found
 290 in Appendix A. We highlight the following:

- Grids are partitioned based on geographic location of buses. The details of zonal partitions can be found in [Appendix A](#).
 - Wind turbines are used to represent RES generators, and wind power will be preferably used to meet the load demand before thermal generators are taken into consideration.
- 295 • Wind power is obtained by considering the probability distribution of wind speed. A given realization of wind speed is converted to wind power using an assumed power rating curve.
- Without loss of generality, the maximum generation capacity of all wind turbines is considered to be the same. The maximum generation capacity for thermal generators is chosen to simulate real-world power grids.
- 300 • Aggregated load demand and wind power generation are used for each zone. These aggregated values are utilized in modeling the spatio-temporal dependence between different zones, while a fixed scalar distribution factor is used to distribute these aggregated values to individual buses in the same zone.

We consider an intra-day planning horizon of $T = 12$ hours for SCUC calculation. Stochastic grid variables are sampled using Algorithm 1. Specifically, wind speed is assumed to follow Weibull PDF (and converted to wind power) while load is described by truncated normal PDF. Latin Hypercube Sampling (LHS) technique is utilized for sampling the grid variables in the first time step. The samples are used as inputs to an MILP solver to obtain the power grid state under the forecast distributions of the grid variables. The well-known power system optimal scheduling tool MATPOWER is used to formulate the MILP problems, which are then solved by using the state-of-the-art Gurobi solver. Overall, 1000 SCUC solutions using MILP are generated for model training and evaluation.

The GNN models are trained using a supervised learning approach in PyTorch and PyG. Inputs to the model include wind generation and load, the grid topology and branch properties. 315 Outputs are the corresponding system/zonal aggregated thermal power generation and load shedding, as well as branch flow, for the planning period (12 hours). The MILP-based SCUC solutions are split into training/validation/testing with proportions 70%/10%/20%. The GNN models are trained on a computer with 128 GB Intel Core i9-13900K processor and RTX A6000 graphical card with 48 GB RAM.

320 For the multi-step reliability/risk assessment, we consider $\Delta T = 2$ hours. This means that

we compute aggregated reliability and risk over two hour-long windows starting at multiple future time instances. For demonstration purposes, we utilize a constant cost function by setting $C_s = \$10/MW$ and $C_o = \$1/MW$. Note, however, that it is possible to implement any cost function. We conduct experiments on the system and individual zones simultaneously, while the results are evaluated separately. By doing so, we aim at investigating the relative vulnerability of different zones and their contributions to the entire power grid. We identify important branches in each grid by computing the average of maximum power flow in all branches over all training data samples. We refer to these branches as *significant* branches. We focus on these branches for branch overloading reliability and risk assessment, since the loading percentage on the remaining branches is far below the security threshold in most operating conditions. The fraction of branch power flow security threshold is set as $\epsilon = 85\%$.

5. Results

The results of the numerical experiments with the three grids are detailed in this section. Firstly, to show that the grid topology indeed changes over the 12-hour planning period, we plot the rate of change of generator ON/OFF status for all three grids. We then evaluate the performance of the GNN models in predicting the QoIs, and the results are shown in terms of mean relative error (MRE, %). Finally, we evaluate the accuracy of reliability and risk quantification based on the GNN predictions by comparing them with the reference reliability and risk estimates computed based on the MILP solution.

5.1. Evolution of thermal generator on/off status and grid topology

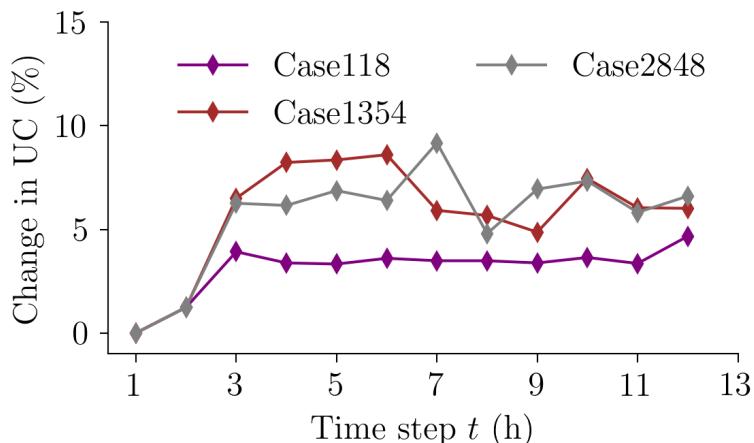


Figure 3: The percentage change in unit commitment (grid topology) over the RUC horizon of interest (the percentage of generators that change status from ‘ON to OFF’ or from ‘OFF to ON’ at each time step)

The average change in the grid topology (the % of thermal generators turned on or off) over the RUC horizon, for the testing data points, is shown in Fig. 3. All thermal generators are online (status: ON) at $t = 1$ h and some of them are turned off at $t = 2$ h, because of low anticipated load demand. Thereafter, the forecast scenarios dictate when a generator will be turned on or off. A significant change in thermal generator on-off status proportionately impacts on thermal power generation, branch power flow, etc. The results in Fig. 3 demonstrate that the testing data samples show significant change in the grid topology over the RUC horizon. The reliability/risk quantification results under such evolving grid topology are of interest in this work.

350 5.2. GNN model predictions

In this section, we evaluate the accuracy of the GNN model for system-level, zone-level, and branch-level QoI predictions.

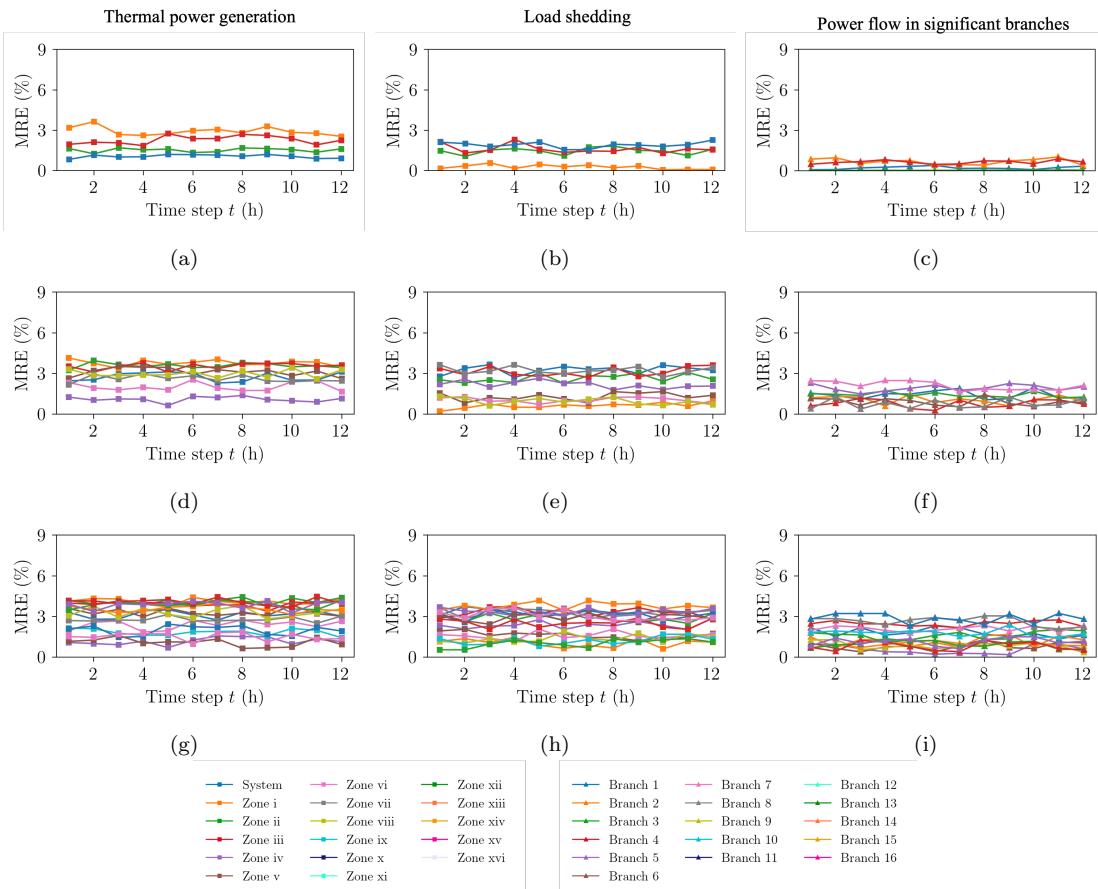


Figure 4: Mean relative error of GNN predictions for aggregated thermal power generation, aggregated load shedding, and power flow in significant branches. (a)–(c): Case118, (d)–(f): Case1354pegase and (g)–(i): Case2848rte. The number of zones for the three grids are 3, 8 and 16, and the number of (selected) significant branches are 4, 8 and 16, respectively. The same color scheme applies to all subsequent figures unless otherwise specified.

5.2.1. System-level and zone-level prediction

For Case118, as shown in Fig. 4a, the MRE of aggregated power generation at each time step within the RUC horizon remains below 4%, indicating excellent accuracy. The MRE for system and zone II is below 2%. In zone I and III, the MRE is slightly higher but still maintains a good accuracy level with MRE mostly under 3%; the MRE goes beyond 3% at two future time steps. The GNN prediction for load shedding is shown in Fig. 4b. The load shedding prediction is also extremely accurate with the MRE less than 1% for zone I and 3% for other zones as well as the system.

The results for Case1354pegase are summarized in Fig 4d and 4e, results for Case2848rte are shown in Fig. 4g and 4h, respectively. Even for these large grids GNN predictions exhibit good accuracy ($MRE < 5\%$) for all zones and the system at different time steps.

5.2.2. Branch-level prediction

In terms of the GNN prediction for branch flow, the results are shown in Fig. 4c. Note that only the results for significant branches are shown here. It is noticed that MRE is consistently below 1% for Case118, suggesting the great predicting capacity of the GNN model. The results for other two grids are shown in Fig 4f and 4i, and GNN prediction MRE remains below 3% for these large grids.

370 5.3. Reliability and risk predictions

In this section, we evaluate the accuracy of GNN-based reliability and risk assessment by comparing it with MILP-based assessment.

5.3.1. Load shedding

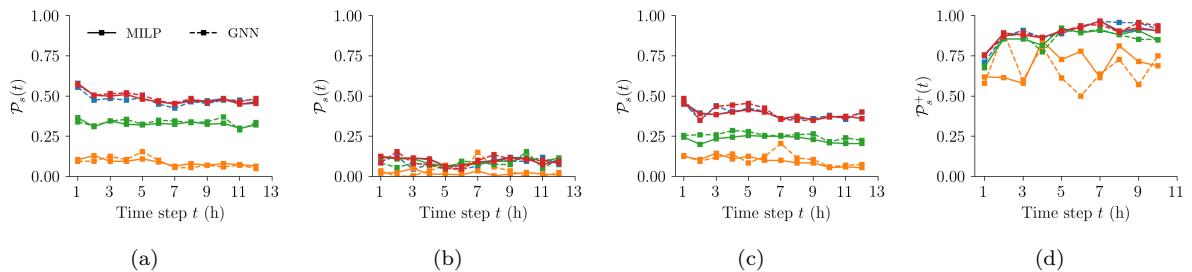


Figure 5: Reliability and risk quantification for load shedding (Case118). (a) Probability of (standalone) total (reserve- and non-reserve-related) load shedding, (b) probability of (standalone) load shedding due to reserve constraint, (c) probability of (standalone) load shedding due to non-reserve constraints and (d) probability of (multi-step) total (reserve- and non-reserve-related) load shedding.

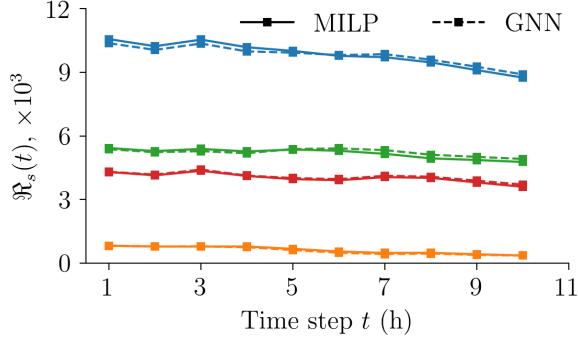


Figure 6: Risk of load shedding (Case118).

The probability of load shedding for Case118 is shown in Fig. 5a. GNN-based reliability quantification is in agreement with the ground truth (MILP solver-based quantification), for both the system and individual zones. The probabilities of standalone total (reserve- plus non-reserve-related) load shedding at zones I, II and III are around 0.1, 0.4 and 0.5, respectively, suggesting that zone II and III have higher vulnerability than zone I. At the system level, the probability of load shedding is around 0.5 over the planning horizon. Load shedding caused by reserve requirement and other constraints are also obtained and shown in Fig. 5a and 5c. The GNN model shows excellent predictive ability for cause-aware load shedding analysis. Unmet reserve requirement constraint seems to be the primary cause of load shedding, as the corresponding probability is significantly larger than that corresponding to non-reserve constraints. GNN-based predictions could be used to analyze and mitigate the reserve requirement-related load shedding in the next few hours. The results suggest that GNN is capable of decoupling different factors that lead to the adverse event, thus it can be a useful tool to improve the operator's situational awareness and enhance system reliability. The multi-step reliability quantification is shown in Fig. 5d. It is seen that GNN-based reliability calculation is very accurate (except for Zone I) and captures the trend of load shedding probability in future two-hour-long windows starting at each of the future hours of interest.

Not surprisingly, GNN-based risk quantification aligns well with that from the reference (MILP-based) solution, as shown in Fig. 6. It is observed that the risk associated with load shedding gradually decreases. We also notice that zone I is significantly less risky than other zones, and the consequence cost almost drops to zero after $T = 8$ hours. In contrast, zone II exhibits the highest risk at around \$5,000, slightly higher than that of zone III which is at \$4,000. The system-level risk caused by load shedding experiences a slight drop from around \$11,000 at $t = 1$ h to \$9,000 at $t = 12$ h, and the trend is accurately captured by the GNN-based

risk computation.

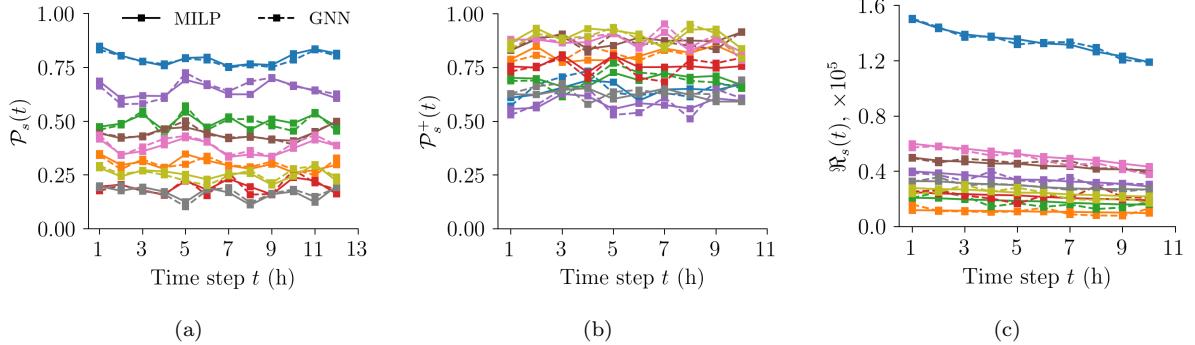


Figure 7: Reliability and risk quantification for load shedding (Case1354pegase). (a) Probability of standalone load shedding, (b) probability of multi-step load shedding, and (c) risk of load shedding.

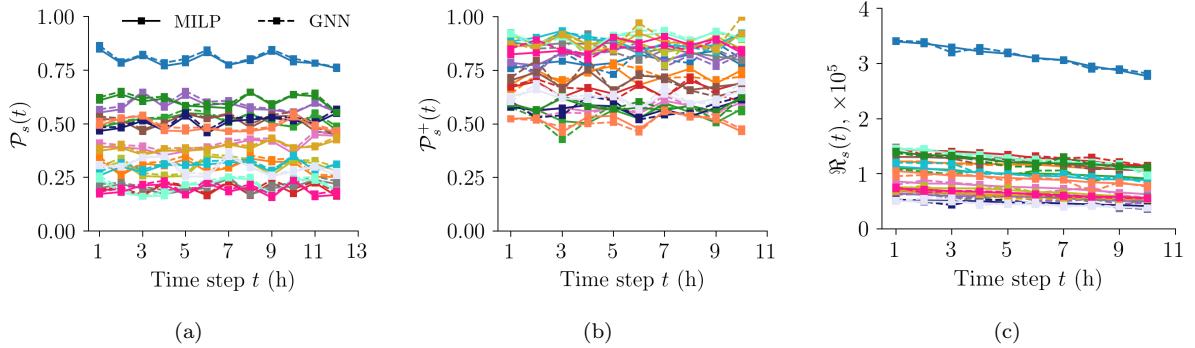


Figure 8: Reliability and risk quantification for load shedding (Case2848rte). (a) Probability of standalone load shedding, (b) probability of multi-step load shedding, and (c) risk of load shedding.

The reliability and risk quantification of load shedding for Case1354pegase and Case2848rte
400 are shown in Fig. 7 and 8, respectively. It can be seen that GNN-based risk quantification is in excellent agreement with the reference (MILP-based) solution for both individual zones and the entire system. The results suggest that GNN can be used to facilitate risk management of large, real-world grids by speeding up the computation. Since well-trained GNN models are able to give QoI predictions for thousands of operational scenarios within a second, the GNN-based approach is thus appealing for hours-ahead reliability and risk assessment.
405

5.3.2. Branch overloading

The probability of standalone branch overloading is shown in Fig. 9a. It is observed that the GNN-based analysis for the selected (significant) branches is very accurate. We notice that the first branch is almost always overloaded with probability near 1 at all time steps, and the second 410 and third branches are often overloaded with probabilities around 0.7 and 0.6, respectively. The fourth branch is overloaded with a probability of 0.5. The results also demonstrate that it is

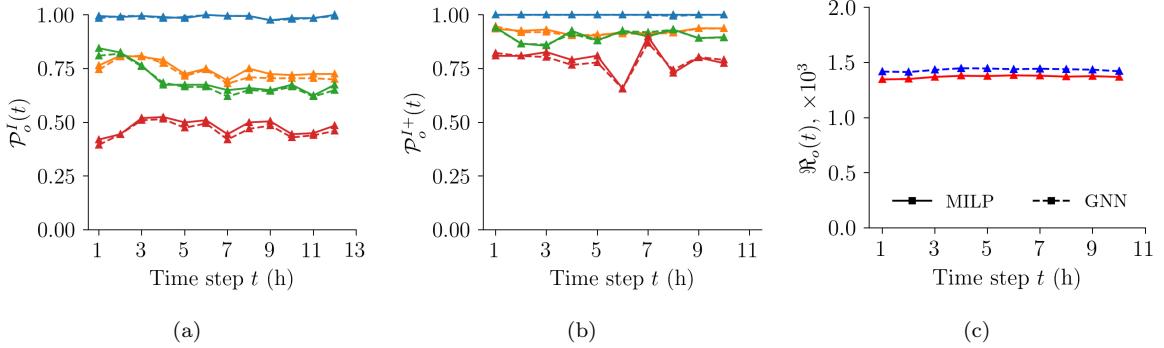


Figure 9: Reliability and risk of branch overloading (Case118). (a) Probability of standalone branch overloading, (b) Probability of multi-step branch overloading, and (c) Risk of branch overloading.

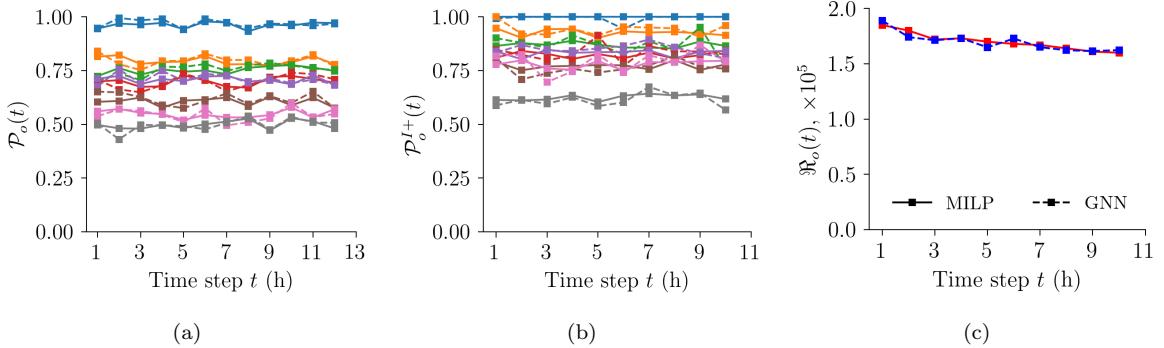


Figure 10: Reliability and risk of branch overloading (Case1354pegase). (a) Probability of standalone branch overloading, (b) Probability of multi-step branch overloading, and (c) Risk of branch overloading.

not necessary to include all the branches in reliability and risk assessment, since the overloading probability for the remaining the branches is smaller than 0.5.

The probabilities of multi-step branch overloading are shown in Fig. 9b. Compared with

standalone overloading, the probability of multi-step overloading for all branches is slightly elevated. In particular, the first branch is still the most vulnerable one with the probability of overloading increased to 1 at all time steps. For the second and third branch, multi-step overloading probability is elevated to 0.9 and 0.8, respectively. Overloading probability for the fourth branch increases to around 0.7. The risk associated with branch overloading is shown in Fig. 9c. It is observed that the GNN-based risk quantification is slightly conservative compared to the reference (MILP-based) solution (however, the error is negligibly small).

The reliability and risk quantification results of branch overloading for Case1354pegase and Case2848rte are shown in Fig. 10 and 11, respectively, and the results confirm the good accuracy GNN-based reliability/risk analysis. This demonstrates the excellent accuracy of GNN-based risk quantification; hence the GNN surrogate model can be a valuable tool to enable fast risk quantification during power grid operation, thus supporting fast, risk-informed decision-making.

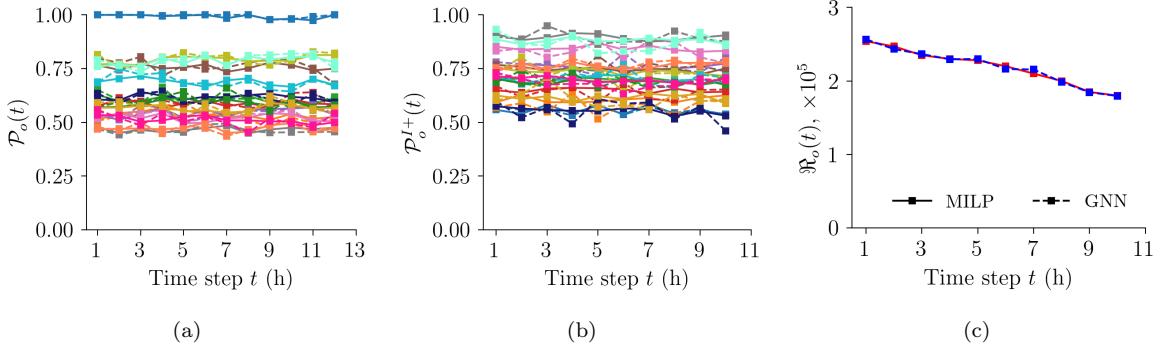


Figure 11: Reliability and risk of branch overloading (Case2848rte). (a) Probability of standalone branch overloading, (b) Probability of multi-step branch overloading, and (c) Risk of branch overloading.

6. Conclusion

We investigated the utilization of graph neural network (GNN) as a proxy to facilitate hours-ahead decision-making in power grid operations, while accounting for anticipated changes in the grid topology (i.e., changes in the generator on/off status). To this end, spatio-temporally correlated samples of stochastic variables (wind power and load demand) are drawn from their joint probability (forecast) distribution, and the corresponding unit commitment and dispatch solutions are obtained using MILP solver. The data (samples and MILP solutions) are used to train the GNN models following a supervised learning approach with stochastic grid variables (wind/solar generation and load) as input and MILP solutions as outputs. Multiple GNN models are trained to predict different quantities of interest (QoIs), i.e., load shedding at system/zonal level and power flow at branch level. The QoIs are then used for reliability and risk quantification using a recently proposed risk assessment framework. The framework employs standalone and multi-step temporal view of failure events for hours-ahead reliability/risk assessment, and thus provides a comprehensive evaluation of risk associated with adverse events. The proposed GNN-based reliability and risk estimation methodology was demonstrated on medium to large realistic power grids. The excellent prediction accuracy and lower computational cost of the GNN model indicate that GNN models can be good proxies for complicated computational tasks in power grid operation. Our results also demonstrated that GNN models can provide accurate estimate of hours-ahead grid operational risk. The proposed GNN-based reliability/risk assessment method utilizes the probabilistic forecasts over the next few hours to obtain the training data. The subsequently trained GNN proxies are only applicable to the probabilistic forecasts used for generating the training data. Future work needs to develop GNN proxies that can be generalized to a wide range of possible forecasts. The statistical distance between the

450 training data and any given probabilistic forecast could be used to gauge the accuracy of the
GNN proxies before incorporating these proxies in real-time risk assessment.

Acknowledgement

This work is partly funded by ARPA-E under PERFORM award DE-AR0001280. The support is gratefully acknowledged. The authors also acknowledge helpful discussions with
455 Dr. Oliver Stover at Charles River Associates and Dr. Ray Daniel Zimmerman at Cornell University.

References

- [1] G. Mittelman, R. Eran, L. Zhivin, O. Eisenhändler, Y. Luzon, M. Tshuva, The potential of renewable electricity in isolated grids: The case of israel in 2050, *Applied Energy* 349 (2023) 121325.
460
- [2] Y. Shang, M. Liu, Z. Shao, L. Jian, Internet of smart charging points with photovoltaic integration: A high-efficiency scheme enabling optimal dispatching between electric vehicles and power grids, *Applied Energy* 278 (2020) 115640.
- [3] Z. J. Lee, J. Z. Pang, S. H. Low, Pricing ev charging service with demand charge, *Electric Power Systems Research* 189 (2020) 106694.
465
- [4] W. Wang, B. Sun, H. Li, Q. Sun, R. Wennersten, An improved min-max power dispatching method for integration of variable renewable energy, *Applied Energy* 276 (2020) 115430.
- [5] J. Li, J. Zhou, B. Chen, Review of wind power scenario generation methods for optimal operation of renewable energy systems, *Applied Energy* 280 (2020) 115992.
470
- [6] Y. Matsuo, S. Endo, Y. Nagatomi, Y. Shibata, R. Komiya, Y. Fujii, Investigating the economics of the power sector under high penetration of variable renewable energies, *Applied Energy* 267 (2020) 113956.
- [7] Y. Ding, T. Morstyn, M. D. McCulloch, Distributionally robust joint chance-constrained optimization for networked microgrids considering contingencies and renewable uncertainty, *IEEE Transactions on Smart Grid* 13 (3) (2022) 2467–2478.
475

- [8] T. G. Hlalele, R. M. Naidoo, R. C. Bansal, J. Zhang, Multi-objective stochastic economic dispatch with maximal renewable penetration under renewable obligation, *Applied Energy* 270 (2020) 115120.
- [9] R. Ponciroli, N. E. Stauff, J. Ramsey, F. Ganda, R. B. Vilim, An improved genetic algorithm approach to the unit commitment/economic dispatch problem, *IEEE Transactions on Power Systems* 35 (5) (2020) 4005–4013.
- [10] D. T. Lagos, N. D. Hatziargyriou, Data-driven frequency dynamic unit commitment for island systems with high res penetration, *IEEE Transactions on Power Systems* 36 (5) (2021) 4699–4711.
- [11] S. Jiang, S. Gao, G. Pan, Y. Liu, C. Wu, S. Wang, Congestion-aware robust security constrained unit commitment model for ac-dc grids, *Applied Energy* 304 (2021) 117392.
- [12] S. Abedi, M. He, D. Obadina, Congestion risk-aware unit commitment with significant wind power generation, *IEEE Transactions on Power Systems* 33 (6) (2018) 6861–6869.
- [13] L. C. da Costa, F. S. Thomé, J. D. Garcia, M. V. Pereira, Reliability-constrained power system expansion planning: A stochastic risk-averse optimization approach, *IEEE Transactions on Power Systems* 36 (1) (2020) 97–106.
- [14] R. Wu, G. Sansavini, Integrating reliability and resilience to support the transition from passive distribution grids to islanding microgrids, *Applied Energy* 272 (2020) 115254.
- [15] W. Zhu, M. Han, J. V. Milanović, P. Crossley, Methodology for reliability assessment of smart grid considering risk of failure of communication architecture, *IEEE Transactions on Smart Grid* 11 (5) (2020) 4358–4365.
- [16] G. S. Seck, V. Krakowski, E. Assoumou, N. Maïzi, V. Mazauric, Embedding power system's reliability within a long-term energy system optimization model: Linking high renewable energy integration and future grid stability for france by 2050, *Applied Energy* 257 (2020) 114037.
- [17] M. Nadeem, S. A. Nugroho, A. F. Taha, Dynamic state estimation of nonlinear differential algebraic equation models of power networks, *IEEE Transactions on Power Systems* 38 (3) (2022) 2539–2552.

- [18] G. N. Psarros, P. A. Dratsas, S. A. Papathanassiou, A comparison between central-and self-dispatch storage management principles in island systems, *Applied Energy* 298 (2021) 117181.
- [19] L. Moretti, E. Martelli, G. Manzolini, An efficient robust optimization model for the unit commitment and dispatch of multi-energy systems and microgrids, *Applied Energy* 261 (2020) 113859.
- [20] M. Isuru, M. Hotz, H. Gooi, W. Utschick, Network-constrained thermal unit commitment for hybrid ac/dc transmission grids under wind power uncertainty, *Applied Energy* 258 (2020) 114031.
- [21] K. S. A. Sedzro, S. Kishore, A. J. Lamadrid, L. F. Zuluaga, Stochastic risk-sensitive market integration for renewable energy: Application to ocean wave power plants, *Applied Energy* 229 (2018) 474–481.
- [22] S. M. Mohseni-Bonab, I. Kamwa, A. Rabiee, C. Chung, Stochastic optimal transmission switching: A novel approach to enhance power grid security margins through vulnerability mitigation under renewables uncertainties, *Applied Energy* 305 (2022) 117851.
- [23] T. Adefarati, R. Bansal, Reliability and economic assessment of a microgrid power system with the integration of renewable energy resources, *Applied Energy* 206 (2017) 911–933.
- [24] O. Stover, P. Karve, S. Mahadevan, Reliability and risk metrics to assess operational adequacy and flexibility of power grids, *Reliability Engineering & System Safety* 231 (2023) 109018.
- [25] A. F. Castelli, L. Pilotti, A. Monchieri, E. Martelli, Optimal design of aggregated energy systems with (n-1) reliability: Milp models and decomposition algorithms, *Applied Energy* 356 (2024) 122002.
- [26] M. Juanpera, L. Ferrer-Martí, R. Pastor, Multi-stage optimization of rural electrification planning at regional level considering multiple criteria. case study in nigeria, *Applied Energy* 314 (2022) 118926.
- [27] C. Lyu, Y. Jia, Z. Xu, A novel communication-less approach to economic dispatch for microgrids, *IEEE Transactions on Smart Grid* 12 (1) (2020) 901–904.

- [28] S. Samsatli, N. J. Samsatli, A multi-objective milp model for the design and operation of future integrated multi-vector energy networks capturing detailed spatio-temporal dependencies, *Applied Energy* 220 (2018) 893–920.
- 535 [29] Y. Cao, W. Wei, S. Mei, M. Shafie-Khah, J. P. Catalao, Analyzing and quantifying the intrinsic distributional robustness of cvar reformulation for chance-constrained stochastic programs, *IEEE transactions on power systems* 35 (6) (2020) 4908–4911.
- 540 [30] M. Rayati, M. Bozorg, R. Cherkaoui, M. Carpita, Distributionally robust chance constrained optimization for providing flexibility in an active distribution network, *IEEE Transactions on Smart Grid* 13 (4) (2022) 2920–2934.
- [31] X. Cao, J. Wang, B. Zeng, Networked microgrids planning through chance constrained stochastic conic programming, *IEEE Transactions on Smart Grid* 10 (6) (2019) 6619–6628.
- 545 [32] N. Gu, H. Wang, J. Zhang, C. Wu, Bridging chance-constrained and robust optimization in an emission-aware economic dispatch with energy storage, *IEEE Transactions on Power Systems* 37 (2) (2021) 1078–1090.
- [33] A. Kirilenko, Y. Gong, C. Chung, A framework for power system operational planning under uncertainty using coherent risk measures, *IEEE Transactions on Power Systems* 36 (5) (2021) 4376–4386.
- 550 [34] R. Ghorani, F. Pourahmadi, M. Moeini-Aghaie, M. Fotuhi-Firuzabad, M. Shahidehpour, Risk-based networked-constrained unit commitment considering correlated power system uncertainties, *IEEE Transactions on Smart Grid* 11 (2) (2019) 1781–1791.
- [35] Y. Xu, L. Mili, A. Sandu, M. R. von Spakovsky, J. Zhao, Propagating uncertainty in power system dynamic simulations using polynomial chaos, *IEEE Transactions on Power Systems* 34 (1) (2018) 338–348.
- 555 [36] H. Gao, X. Lyu, S. He, L. Wang, C. Wang, J. Liu, Integrated planning of cyber-physical active distribution system considering multidimensional uncertainties, *IEEE Transactions on Smart Grid* 13 (4) (2022) 3145–3159.
- [37] Y. Zhang, P. M. Karve, S. Mahadevan, Power grid operational risk assessment using graph neural network surrogates, *arXiv preprint arXiv:2311.12309* (2023).

- 560 [38] J. Xu, Y. Chen, S. Liao, Y. Sun, L. Yao, H. Fu, X. Jiang, D. Ke, X. Li, J. Yang, et al.,
Demand side industrial load control for local utilization of wind power in isolated grids,
Applied Energy 243 (2019) 47–56.
- 565 [39] J. Liu, Y. Zhang, K. Meng, Z. Y. Dong, Y. Xu, S. Han, Real-time emergency load shedding
for power system transient stability control: A risk-averse deep learning method, Applied
Energy 307 (2022) 118221.
- [40] N. Luo, J. Langevin, H. Chandra-Putra, S. H. Lee, Quantifying the effect of multiple load
flexibility strategies on commercial building electricity demand and services via surrogate
modeling, Applied Energy 309 (2022) 118372.
- 570 [41] Q. Wang, C. Miao, Y. Tang, Power shortage support strategies considering unified gas-
thermal inertia in an integrated energy system, Applied Energy 328 (2022) 120229.
- [42] B. Li, C. Feng, C. Siebenschuh, R. Zhang, E. Spyrou, V. Krishnan, B. F. Hobbs, J. Zhang,
Sizing ramping reserve using probabilistic solar forecasts: A data-driven method, Applied
Energy 313 (2022) 118812.
- 575 [43] V. Mohan, J. G. Singh, W. Ongsakul, An efficient two stage stochastic optimal energy and
reserve management in a microgrid, Applied energy 160 (2015) 28–38.
- [44] R. Khatami, M. Parvania, A. Narayan, Flexibility reserve in power systems: Definition
and stochastic multi-fidelity optimization, IEEE Transactions on Smart Grid 11 (1) (2019)
644–654.
- 580 [45] M. Heleno, M. A. Matos, J. P. Lopes, Availability and flexibility of loads for the provision
of reserve, IEEE Transactions on Smart Grid 6 (2) (2014) 667–674.
- [46] G. Morales-España, R. Baldick, J. García-González, A. Ramos, Power-capacity and ramp-
capability reserves for wind integration in power-based uc, IEEE Transactions on Sustainable
Energy 7 (2) (2015) 614–624.
- 585 [47] Y. Gong, Q. Jiang, R. Baldick, Ramp event forecast based wind power ramp control with
energy storage system, IEEE Transactions on Power Systems 31 (3) (2015) 1831–1844.
- [48] M. Dolanyi, K. Bruninx, J.-F. Toubeau, E. Delarue, Risk-based constraints for the optimal
operation of an energy community, IEEE Transactions on Smart Grid 13 (6) (2022) 4551–
4561.

- [49] Y. Ryu, H.-W. Lee, A real-time framework for matching prosumers with minimum risk in
590 the cluster of microgrids, *IEEE Transactions on Smart Grid* 11 (4) (2020) 2832–2844.
- [50] S. Nikkhah, I. Sarantakos, N.-M. Zografou-Barredo, A. Rabiee, A. Allahham, D. Giaouris,
A joint risk-and security-constrained control framework for real-time energy scheduling of
islanded microgrids, *IEEE Transactions on Smart Grid* 13 (5) (2022) 3354–3368.
- [595] N. Neyestani, M. Yazdani-Damavandi, M. Shafie-Khah, G. Chicco, J. P. Catalão, Stochastic
modeling of multienergy carriers dependencies in smart local networks with distributed
energy resources, *IEEE Transactions on Smart Grid* 6 (4) (2015) 1748–1762.
- [52] X. Y. Lee, S. Sarkar, Y. Wang, A graph policy network approach for volt-var control in
power distribution systems, *Applied Energy* 323 (2022) 119530.
- [600] J. Lu, C. Zhang, J. Li, Y. Zhao, W. Qiu, T. Li, K. Zhou, J. He, Graph convolutional
networks-based method for estimating design loads of complex buildings in the preliminary
design stage, *Applied Energy* 322 (2022) 119478.
- [54] M. Liu, H. Gao, S. Ji, Towards deeper graph neural networks, in: *Proceedings of the 26th
ACM SIGKDD international conference on knowledge discovery & data mining*, 2020, pp.
338–348.
- [605] [55] J. Feng, Y. Chen, F. Li, A. Sarkar, M. Zhang, How powerful are k-hop message passing
graph neural networks, *Advances in Neural Information Processing Systems* 35 (2022)
4776–4790.
- [56] Z. Liu, J. Zhou, *Introduction to graph neural networks*, Springer Nature, 2022.
- [610] [57] W. Fan, Y. Ma, Q. Li, Y. He, E. Zhao, J. Tang, D. Yin, Graph neural networks for social
recommendation, in: *The world wide web conference*, 2019, pp. 417–426.
- [58] S. Min, Z. Gao, J. Peng, L. Wang, K. Qin, B. Fang, Stgsn—a spatial-temporal graph
neural network framework for time-evolving social networks, *Knowledge-Based Systems*
214 (2021) 106746.
- [615] [59] Z. Zhang, F. Zhuang, H. Zhu, Z. Shi, H. Xiong, Q. He, Relational graph neural network
with hierarchical attention for knowledge graph completion, in: *Proceedings of the AAAI
conference on artificial intelligence*, Vol. 34, 2020, pp. 9612–9619.

- [60] R. A. Jacob, S. Senemmar, J. Zhang, Fault diagnostics in shipboard power systems using graph neural networks, in: 2021 IEEE 13th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED), Vol. 1, IEEE, 2021, pp. 316–321.
- [620]
- [61] D. Owerko, F. Gama, A. Ribeiro, Predicting power outages using graph neural networks, in: 2018 ieee global conference on signal and information processing (globalsip), IEEE, 2018, pp. 743–747.
- [62] B. Donon, R. Clément, B. Donnot, A. Marot, I. Guyon, M. Schoenauer, Neural networks for power flow: Graph neural solver, *Electric Power Systems Research* 189 (2020) 106547.
- [625]
- [63] D. Wu, W. Lin, Efficient residential electric load forecasting via transfer learning and graph neural networks, *IEEE Transactions on Smart Grid* 14 (3) (2022) 2423–2431.
- [64] Y. Zhang, P. M. Karve, S. Mahadevan, Graph neural networks for power grid operational risk assessment, *arXiv preprint arXiv:2311.03661* (2023).
- [630] [65] T. Falconer, L. Mones, Leveraging power grid topology in machine learning assisted optimal power flow, *IEEE Transactions on Power Systems* 38 (3) (2022) 2234–2246.
- [66] F. Diehl, Warm-starting ac optimal power flow with graph neural networks, in: 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), 2019, pp. 1–6.
- [635] [67] Y. Chen, Q. Wang, X. Wang, Y. Guan, Applying robust optimization to miso look-ahead commitment, in: 2014 IEEE PES General Meeting— Conference & Exposition, IEEE, 2014, pp. 1–5.
- [68] Y. Zhou, M. Shahidehpour, Z. Wei, G. Sun, S. Chen, Multistage robust look-ahead unit commitment with probabilistic forecasting in multi-carrier energy systems, *IEEE Transactions on Sustainable Energy* 12 (1) (2020) 70–82.
- [640] [69] S. Park, W. Chen, D. Han, M. Tanneau, P. Van Hentenryck, Confidence-aware graph neural networks for learning reliability assessment commitments, *IEEE Transactions on Power Systems* (2023).
- [70] Y. Chen, D. Savageau, F. Wang, A. Casto, Voltage and local reliability commitment under electricity market operations, in: 2016 IEEE Power and Energy Society General Meeting (PESGM), IEEE, 2016, pp. 1–5.
- [645]

- [71] S. Goleijani, T. Ghanbarzadeh, F. S. Nikoo, M. P. Moghaddam, Reliability constrained unit commitment in smart grid environment, *Electric power systems research* 97 (2013) 100–108.
- [72] R. Billinton, M. Fotuhi-Firuzabad, A reliability framework for generating unit commitment, *Electric Power Systems Research* 56 (1) (2000) 81–88.
- [73] B. Zhou, J. Fang, X. Ai, Y. Zhang, W. Yao, Z. Chen, J. Wen, Partial-dimensional correlation-aided convex-hull uncertainty set for robust unit commitment, *IEEE Transactions on Power Systems* (2022).
- [74] A. S. Xavier, F. Qiu, F. Wang, P. R. Thimmapuram, Transmission constraint filtering in large-scale security-constrained unit commitment, *IEEE Transactions on Power Systems* 34 (3) (2019) 2457–2460.
- [75] K. Šepetanc, H. Pandžić, Convex polar second-order taylor approximation of ac power flows: A unit commitment study, *IEEE Transactions on Power Systems* 36 (4) (2020) 3585–3594.
- [76] Z. Tang, D. J. Hill, T. Liu, A novel consensus-based economic dispatch for microgrids, *IEEE Transactions on Smart Grid* 9 (4) (2018) 3920–3922.
- [77] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, M. Sun, Graph neural networks: A review of methods and applications, *AI open* 1 (2020) 57–81.
- [78] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, S. Y. Philip, A comprehensive survey on graph neural networks, *IEEE transactions on neural networks and learning systems* 32 (1) (2020) 4–24.
- [79] T. N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, *arXiv preprint arXiv:1609.02907* (2016).
- [80] N. Sapountzoglou, J. Lago, B. De Schutter, B. Raison, A generalizable and sensor-independent deep learning method for fault detection and location in low-voltage distribution grids, *Applied Energy* 276 (2020) 115299.
- [81] Y. Benmahamed, M. Teguar, A. Boubakeur, Application of svm and knn to duval pentagon 1 for transformer oil diagnosis, *IEEE Transactions on Dielectrics and Electrical Insulation* 24 (6) (2017) 3443–3451.

- 675 [82] W. Hamilton, Z. Ying, J. Leskovec, Inductive representation learning on large graphs,
 Advances in neural information processing systems 30 (2017).
- [83] R. D. Zimmerman, C. E. Murillo-Sánchez, D. Gan, Matpower: A matlab power system
 simulation package, Manual, Power Systems Engineering Research Center, Ithaca NY 1
 (1997) 10–7.
- 680 [84] S. Fliscounakis, P. Panciatici, F. Capitanescu, L. Wehenkel, Contingency ranking with
 respect to overloads in very large power systems taking into account uncertainty, preventive,
 and corrective actions, IEEE Transactions on Power Systems 28 (4) (2013) 4909–4917.
- [85] C. Josz, S. Fliscounakis, J. Maeght, P. Panciatici, Ac power flow data in matpower and
 qcqp format: itesla, rte snapshots, and pegaso, arXiv preprint arXiv:1603.01533 (2016).

685 **Appendix A. Synthetic power grids**

Appendix A.1. Case118

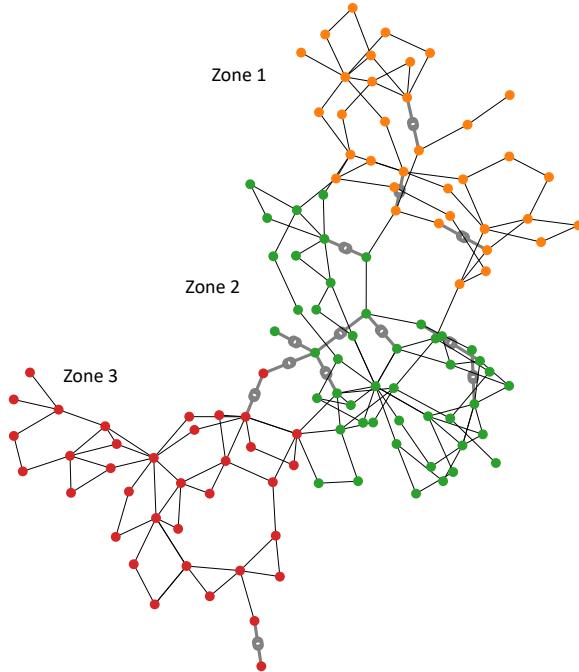


Figure A.1: Case118 power grid.

Case118 power grid contains 118 buses (54 generators, 99 loads and 186 branches). It is partitioned into three zones as shown in Fig. A.1 and Table A.1. For each zone, an aggregated value of load and wind power generation are drawn from their respective probability distributions. The relative contribution of each load bus/wind turbine is assumed to be a constant in

each zone. Consider zone i containing N_L load buses and N_W wind turbines; then the bus-level load and wind power for zone i are calculated as:

$$\begin{aligned} L_{i,j} &= r_{i,j} L_i, \quad j = 1, 2, \dots, N_L \\ W_{i,k} &= q_{i,k} W_i, \quad k = 1, 2, \dots, N_W \end{aligned} \quad (\text{A.1})$$

where $r_{i,j}$ and $q_{i,k}$ represent the relative contribution of each load bus and wind generator bus to the respective zonal (aggregated) values, and L_i and W_i denote the aggregated values for zone i . It is assumed that load follows truncated normal PDF, whereas wind power generation is converted from wind speed, which is assumed to follow Weibull PDF. The probability distribution parameters are listed in Table A.2.

Wind power generation is computed from wind speed as:

$$P = \begin{cases} \max(0, P_r (v^3 - v_{min}^3) / (v_{max}^3 - v_{min}^3)), & v < v_{min} \\ P_r (v^3 - v_{min}^3) / (v_{max}^3 - v_{min}^3), & v_{min} \leq v \leq v_{max} \\ \min(P_r, P_r (v^3 - v_{min}^3) / (v_{max}^3 - v_{min}^3)), & v > v_{max} \end{cases} \quad (\text{A.2})$$

where v_{min} represents the minimum wind speed that can rotate wind turbine, and v_{max} denotes wind speed at which maximum power generation limits is reached. P_r is the wind power generation at v_{max} . In this work, $v_{min} = 1 \text{ ms}^{-1}$, $v_{max} = 15 \text{ ms}^{-1}$ and $P_r = 100 \text{ MW}$.

Table A.1: Case118 power grid

Zone	Thermal generators	Wind generators	Loads
I	12	3	29
II	15	7	38
III	11	6	32
System	38	16	99

Table A.2: Distribution parameters used for load (MW) and wind speed ($m s^{-1}$) for grid Case118. **TN**: truncated normal PDF, **WB**: Weibull PDF

Zone	Location	Shape	Scale	Left truncation	Right truncation
I	TN	50	—	15	10 90
	WB	0	2	8	—
II	TN	75	—	20 25	125
	WB	0	1.8	8.2	—
III	TN	100	—	15 60	140
	WB	0	2.2	7.8	—

Appendix A.2. Case1354pegase

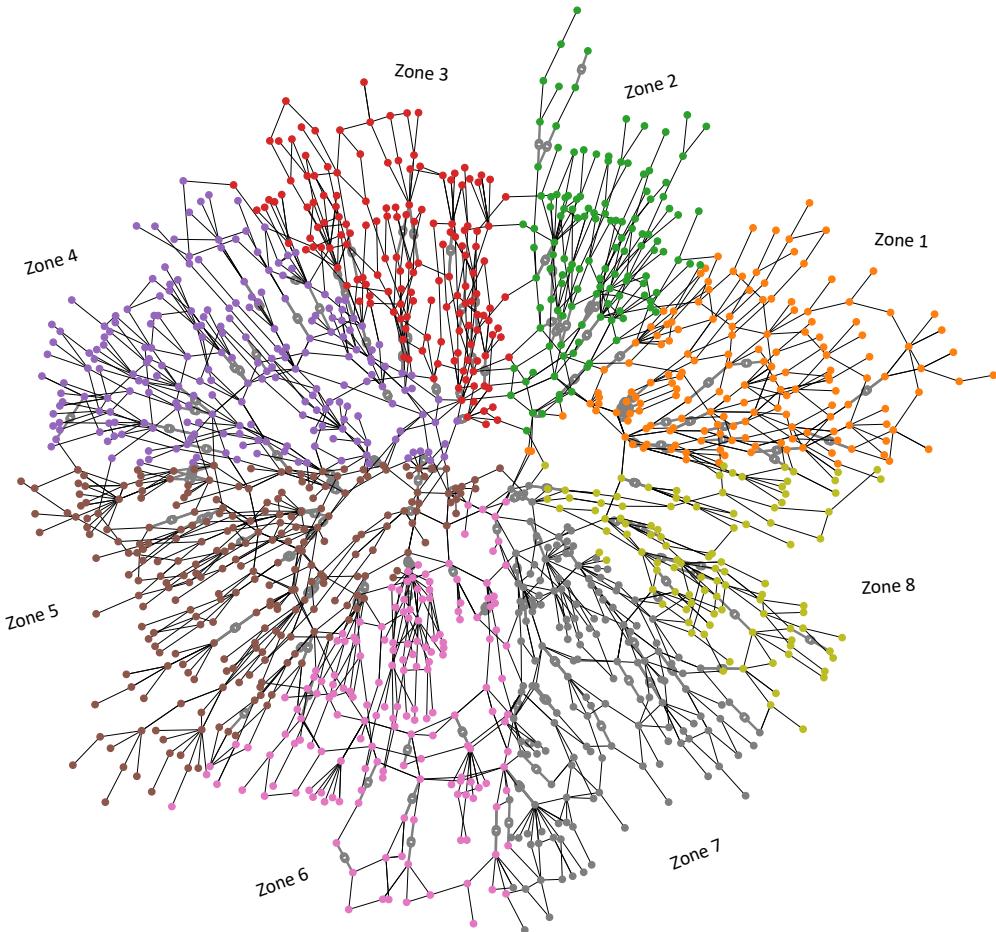


Figure A.2: Case1354pesage power grid.

The Case1354pegase power grid accurately represents the size and complexity of part of the European high voltage transmission network. The network contains 1354 buses (260 generators, 621 loads and 1991 branches). The data stems from the Pan European Grid Advanced Simulation and State Estimation (PEGASE) project, part of the 7th Framework Program of the European Union. Refer to [84] for more information. The grid is partitioned into 8 zones

as shown in Fig. A.2 and Table A.3. The probability distribution parameters are listed in Table A.4.

Table A.3: Case1354pegase power grid

Zone	Thermal generators	Wind generators	Loads
I	39	9	60
II	32	8	48
III	28	7	76
IV	20	5	115
V	24	5	114
VI	11	2	94
VII	34	8	62
VIII	23	5	52
System	211	49	621

Table A.4: Distribution parameters used for load (MW) and wind speed (m/s) for grid Case1354pegase. **TN**: truncated normal PDF, **WB**: Weibull PDF

Zone	Location	Shape	Scale	Left truncation	Right truncation
I	TN 50	–	15	10	90
	WB 0	2	8	–	–
II	TN 75	–	20	25	125
	WB 0	1.8	8.2	–	–
III	TN 100	–	15	60	140
	WB 0	2.2	7.8	–	–
IV	TN 48	–	12	9	87
	WB 0	1.9	8.1	–	–
V	TN 81	–	21	23	139
	WB 0	2.1	7.9	–	–
VI	TN 98	–	15	60	136
	WB 0	1.9	7.7	–	–
VII	TN 52	–	14	7	97
	WB 0	2.1	8	–	–
VIII	TN 83	–	20	21	145
	WB 0	2.0	8.3	–	–

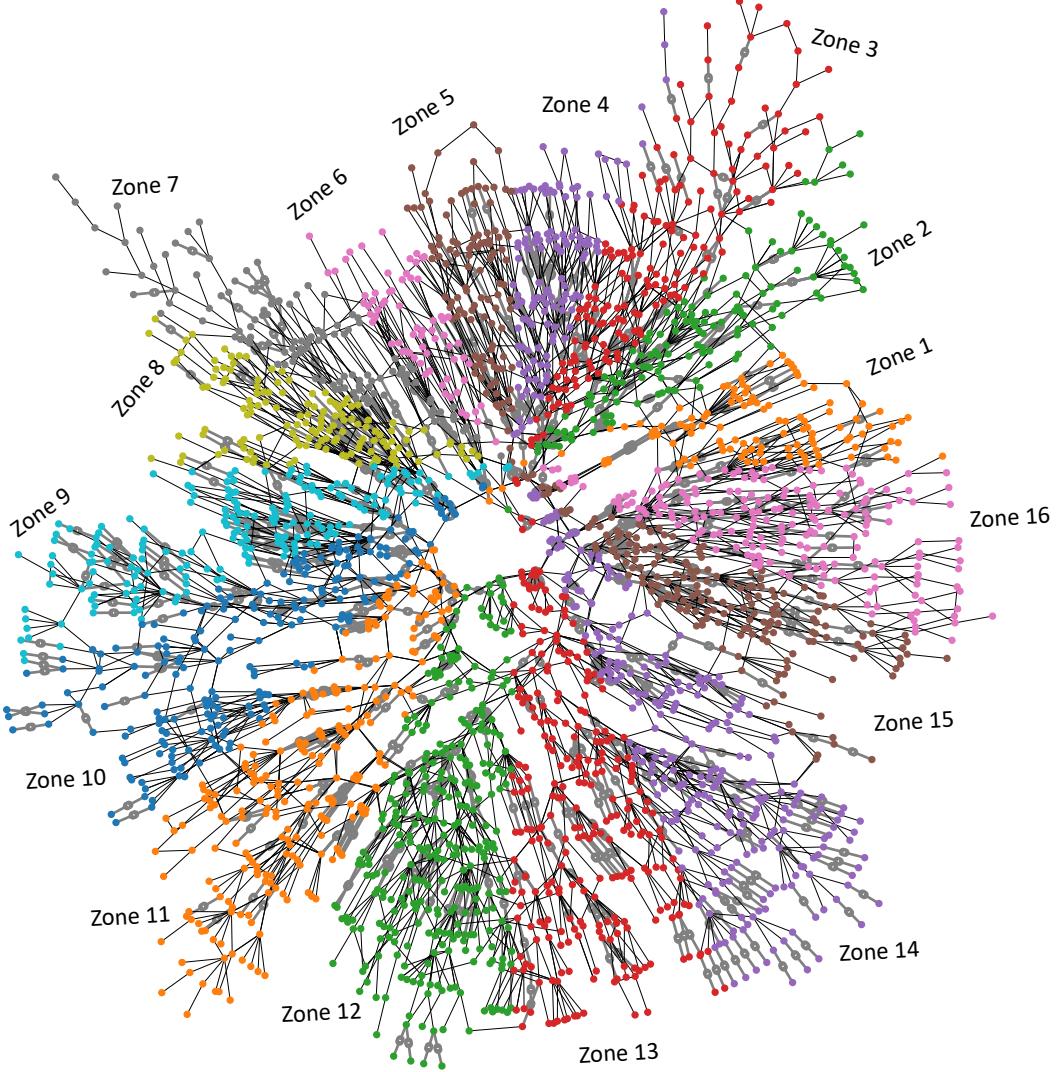


Figure A.3: Case2848rte power grid.

Case2848rte power grid accurately represents the size and complexity of French (very high voltage) transmission network. The network contains 2848 buses (270 generators, 1388 loads and 3776 branches). Refer to [85] for more details. The grid is partitioned into 16 zones (see Fig A.3 and Table A.5), and distribution parameters are shown in Table A.6.

Table A.5: Case28248rte power grid

Zone	Thermal generators	Wind generators	Loads
I	23	5	43
II	24	5	62
III	22	5	78
IV	8	2	71
V	8	2	59
VI	8	2	39
VII	22	5	57
VIII	16	3	69
IX	21	5	104
X	13	3	113
XI	15	3	113
XII	16	4	166
XIII	20	5	149
XIV	44	10	95
XV	17	4	82
XVI	24	6	88
System	301	69	1388

Table A.6: Distribution parameters used for load (MW) wind speed (m/s) for grid Case2848rte. **TN**: truncated normal PDF, **WB**: Weibull PDF.

Zone		Location	Shape	Scale	Left truncation	Right truncation
I	TN	50	–	15	10	90
	WB	0	2	8	–	–
II	TN	75	–	20	25	125
	WB	0	1.8	8.2	–	–
III	TN	100	–	15	60	140
	WB	0	2.2	7.8	–	–
IV	TN	48	–	12	9	87
	WB	0	1.9	8.1	–	–
V	TN	81	–	21	23	139
	WB	0	2.1	7.9	–	–
VI	TN	98	–	15	60	136
	WB	0	1.9	7.7	–	–
VII	TN	52	–	14	7	97
	WB	0	2.1	8	–	–
VIII	TN	83	–	20	21	145
	WB	0	2.0	8.3	–	–
IX	TN	45	–	12	14	76
	WB	0	1.8	7.8	–	–
X	TN	85	–	18	25	145
	WB	0	1.8	8.1	–	–
XI	TN	98	–	16	55	141
	WB	0	8.1	12.5	–	–
XII	TN	53	–	13	9	97
	WB	0	2.2	8.5	–	–
XIII	TN	84	–	20	23	139
	WB	0	1.8	7.5	–	–
XIV	TN	98	–	15	55	141
	WB	0	2	7.7	–	–
XV	TN	47	–	12	8	86
	WB	0	2	8	–	–
XVI	TN	85	–	20	21	149
	WB	0	2.5	8.3	–	–