Risk-Based Active Distribution System Planning for Resilience Against Extreme Weather Events

Abodh Poudyal, *Graduate Student Member, IEEE*, Shiva Poudel, *Member, IEEE*, and Anamika Dubey, *Senior Member, IEEE*

***Abstract*—Enhancing the resilience of power distribution systems to extreme weather events is of critical concern. Upgrading the distribution system infrastructure by system hardening and investing in smart grid technologies effectively enhances grid resilience. Existing distribution system planning methods primarily consider the persistent cost of the expected events (such as faults and outages likely to occur) and aim at improving system reliability. The resilience to extreme weather events requires reducing the impacts of the high impact low probability (HILP) events that are characterized by the tail probability of the event impact distribution. Thus, the resilience-oriented system upgrades solutions need to be driven by the risks imposed by extreme weather events on the power grid infrastructure rather than persistent costs. This paper aims to develop a risk-based approach for the long-term resilience planning of active power distribution systems against extreme weather events. The proposed approach explicitly models (1) the impacts of HILP events using a two-stage risk-averse stochastic optimization framework, thus, explicitly incorporating the risks of HILP events in long-term planning decisions, and (2) the advanced distribution grid operations (in the aftermath of the event) such as intentional islanding into infrastructure planning problem. The inclusion of risk in the planning objective provides additional flexibility to the grid planners to analyze the trade-off between risk-neutral and risk-averse planning strategies.**

***Index Terms*—Conditional value-at-risk, long-term planning, power distribution resilience, stochastic optimization.**

NOMENCLATURE

Parameters

α	Confidence level for VAR and CVaR.
$\mathcal{C}^{DG}_{\max}$	Maximum budget for total DG installation.
$\boldsymbol{P}_{Li} + \mathrm{j}\boldsymbol{Q}_{Li}$	Three-phase complex load demand at bus $i \in \mathcal{B}$.
$\mathbf{r}_e/\mathbf{x}_e$	Resistance/Reactance matrix of a line e : $(i,j) \in \mathcal{L}$.
$\lvert\mathcal{L}_c\rvert$	Number of switches in a cycle c.
Ξ	Total number of scenarios.
c^{DG}_i	DG siting and sizing cost at bus $i \in \mathcal{B}_{DG}$.
P_D	Total power demand.
p_ξ	Probability of realizing scenario $\xi \in \mathcal{E}$.
$P^{\max}_{DG}$	Maximum active power capacity of a DG.
$P^{\phi}_{Li} + \mathrm{j}Q^{\phi}_{Li}$	Complex power demand at $i \in \mathcal{B}$ for phase $\phi \in \Phi$.

Sets

$\mathcal{L}_c$	Switches in a cycle c.
$\mathcal{L}^v_S$	Virtual edges for DGs connection.
$\mathcal{B}$	Physical system buses.
$\mathcal{B}_S$	Remotely switchable buses.
$\mathcal{B}_{DG}$	Potential DG locations.
$\mathcal{E}$	Potential scenarios.
$\mathcal{L}$	Physical lines.
$\mathcal{L}_F$	Faulted or tripped switches.
$\mathcal{L}_S$	Switchable lines. i.e. $\mathcal{L}^s_S \cup \mathcal{L}^t_S$.
$\mathcal{L}^s_S$	Normally-closed sectionalizing switches.
$\mathcal{L}^t_S$	Normally-open tie switches.
$\Phi = \{a,b,c\}$	Phases of a bus.

Variables

β^{DG}_i	Size of DG at location $i \in \mathcal{B}_{DG}$.
δ^{DG}_i	DG location variable $i \in \mathcal{B}_{DG}$.
δ_e	Line or switch decision variable.
η	Value-at-risk.
ν	Excess variable for CVaR.
$\boldsymbol{P}_e + \mathrm{j}\boldsymbol{Q}_e$	Three-phase complex power flow from i to j.
$\boldsymbol{U}_i$	Three-phase voltage magnitude square vector.
$\boldsymbol{V}_i$	Three-phase voltage vector for bus i.
$P^{\phi}_e + \mathrm{j}Q^{\phi}_e$	Complex power flow from i to j for $\phi \in \Phi$.
s_i	Load pick-up variable.
v_i	Bus energization variable.

Manuscript received 1 March 2022; revised 9 July 2022 and 9 October 2022; accepted 3 November 2022. Date of publication 8 November 2022; date of current version 22 March 2023. This work was supported by the National Science Foundation (NSF) CAREER under Grant 1944142. Paper no. TSTE-00215-2022. *(Corresponding author: Abodh Poudyal.)*

Abodh Poudyal and Anamika Dubey are with the EECS, Washington State University, Pullman, WA 99164-1009 USA (e-mail: abodh.poudyal@wsu.edu; anamika.dubey@wsu.edu).

Shiva Poudel is with the Pacific Northwest National Laboratory, Richland, WA 99354 USA (e-mail: shiva.poudel@pnnl.gov).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSTE.2022.3220561.

Digital Object Identifier 10.1109/TSTE.2022.3220561

I. INTRODUCTION

A. Background

Extreme weather events result in an extended disruption of the electric power supply and severely affect personal safety and national security, thus posing serious concerns for the nation's electric power grid infrastructures. Between 2003–2012, in the US, extreme weather events accounted for almost 700 power outages, each affecting more than 50,000 customers, with 80–90% of outages resulting from the failure in power distribution systems [1]. In 2021 alone, in the US, there were 20 natural disaster events with losses exceeding $1 billion each [2]. The

TABLE I
SUMMARY OF EXISTING LITERATURE FOR RESILIENCE-ORIENTED PLANNING OF POWER DISTRIBUTION SYSTEMS

Planning decision	References	Objective	Formulation	Approach
Resource location	[10]	*Max.* (Expected benefit - cost)	Stochastic NLP	Heuristic search
	[11]	*Min.* Cost (operation, action)	Two-stage stochastic MILP	Progressive hedging
Line hardening & DG siting	[7]	*Min.* Load shedding cost and planning cost	Tri-level robust optimization	CCG decomposition
	[12]	*Min.* (planning and expected operating cost)	Two-stage stochastic	Progressive hedging
	[13]	*Min.* Cost (Planning + load shedding)	Tri-level robust optimization	Greedy search
	[14]	*Min.* Load shedding cost and planning cost	Two-stage stochastic	Dual decomposition
Remote controlled switch siting	[15]	*Min.* Number of RCS	Weighted set cover (WSC)	Greedy algorithm
	[16]	*Min.* (Expected loss)	Two-stage stochastic	Scenario decomposition
DG siting & sizing	[17]	*Min.* (planning and expected operating cost)	Two-stage stochastic	Progressive hedging

resulting socio-economic losses and the power grid's vulnerability to extreme weather events necessitate the incorporation of resilience in system planning to account for not only the expected events but also the extreme events that are less likely to occur. Towards this goal, different utilities have spent millions of dollars in deploying smart grid technologies such as distribution automation with automated feeder switching, intentional islanding (microgrid), as well as upgrading vulnerable feeders and substations [3]. With the increasing frequency and severity of weather-related events, a more systematic approach to smart grid expenditures is required to identify appropriate system upgrade solutions for strengthening system resilience.

B. Literature Review

Strategies for enhancing the resilience of power distribution systems can be classified into short-term (operational) planning and long-term (infrastructure) planning phases. The operational planning aims at making the best use of the existing distribution grid resources (e.g., switches, DGs) to minimize the impacts of an anticipated event in the near term [4]. On the contrary, the infrastructure planning phase targets to optimally upgrade the power systems infrastructure by strategically deploying new resources (e.g., DGs, switches, distribution lines) to improve the system response against possible HILP events [5].

The existing literature on power distribution resilience includes numerous articles on operational planning to mitigate or reduce the impact of an imminent threat such as an upcoming storm [6]. Such solutions build resilience via operational response rather than infrastructural upgrades. In operational planning, the decisions are to be made for an upcoming event that is known with a high level of certainty and thus requires considering only a limited number of scenarios for decision-making. On the contrary, long-term planning requires a probabilistic analysis over a wider range of scenarios with a higher level of uncertainty for decisions related to system hardening, infrastructure upgrades, resource allocation, and sizing [7]. These decisions must also connect to the operational problem if and when the events are realized in practice. Thus, the problem is further complicated by additional stages of operational decision-making leading to an explosion of state space to be considered for decision-making.

In general, hardening the distribution grid and investing in smart grid technologies are effective resilience-oriented designs that need to be adopted in the utilities' portfolio for long-term infrastructure planning to improve grid's response to extreme weather events. However, infrastructure planning for resilience is challenging mainly as it requires a) to include several different uncertainties (e.g., fault location, load profiles, nature and severity of extreme events, and so forth) in the decision-making process, b) ensure the validity of the planning decisions for the entire profile of weather events, and c) mitigate the critical challenge of achieving a balance between computational cost and accuracy. Thus, long-term infrastructure planning for resilience is conceptually a different problem than the prevalent methods for operational planning solutions [8]. This calls for new methods and contributions to systematically build resilience in active power distribution grid infrastructures against the HILP events [9].

The related literature on resilience-oriented design and pre-disaster resource allocation usually employ a stochastic programming model to minimize the expected cost of the future operational scenarios [10], [11]. For example, in [10], a heuristic search is employed to identify the optimal restoration path and obtain a resource allocation plan. These solutions, however, consider short-term operational requirements for a known HILP event and are not suitable for infrastructure planning. The related work on resilience-oriented distribution system long-term planning also employs stochastic optimization formulations, including a tri-level robust optimization model [7], [13], and a two-stage stochastic optimization model [12], [14], [16]. The tri-level optimization model formulates the resulting problem in a defender-attacker-defender model that is then converted into an equivalent bi-level model and solved using iterative approaches such as CCG or greedy search algorithms [7]. The tri-level approach optimizes for the worst possible outcomes and hence is not suitable for a probabilistic analysis for infrastructure planning that needs to be cost-effective and optimal for a large range of future scenarios. Alternatively, the two-stage stochastic programming method considers the overall impact of stochastic fault scenarios in planning decisions rather than just the worst-case scenarios [12], [17]. The existing two-stage stochastic optimization formulations used in resilience-oriented distribution grid design either assume that all scenarios are observed with equal probability or perform the planning based on only a targeted set of scenarios [12], [18], [19]. While such methods are generally applicable, other approaches such as importance sampling and stratified sampling techniques can be more effective in representing HILP events and their impact probabilities in

the optimization process [20], [21]. These techniques are widely adopted in power systems reliability studies [22], [23]. Another approach includes network interdiction models that plan for the worst-case scenarios [7]. However, given the emphasis on the worst-case scenario, these methods may lead to extremely conservative and expensive planning solutions as the worst-case scenario usually occurs with a very small probability. Additionally, related literature obtains planning solutions for persistent costs; they do not explicitly include the risks of extreme events. For a resilience-oriented long-term planning problem, the goal is to determine optimal investments to reduce the consequences (here, customer outages) of the HILP events. Mathematically, this amounts to minimizing the mean of consequences and, more importantly, reducing the tail of the consequence of the HILP events [3]. Table I summarizes the existing work in this domain and highlights the problem formulation and solution approaches.

C. Contributions

To appropriately incorporate the effects of HILP events, the infrastructure planning for resilience needs to be driven by risk rather than pertinent or expected cost [24]. This paper proposes a risk-based approach for infrastructure planning in active power distribution systems for resilience against extreme weather events. We employ conditional value-at-risk, CVaR_α, to quantify the risks of system outages imposed by the HILP events [25], [26]. Related works include using CVaR-metric for robust power grid operations [27], [28], [29]. Here, we employ CVaR-metric for long-term distribution planning, where the goal is to reduce the risks of outages on power distribution systems due to HILP events. A two-stage stochastic optimization framework is proposed to optimize smart grid investments, specifically DG siting and sizing, to enable advanced systems such as DG-assisted restoration and intentional islanding. The proposed model also helps evaluate the trade-off between risk-averse and risk-neutral planning decisions. To the authors' knowledge, the proposed framework for resilience is first to introduce risk-based metrics for infrastructure planning in power distribution systems with probabilistic extreme event scenarios. Note that although in [30], CVaR optimization is used for resilience enhancement, their focus is on reducing the variance of the optimization solution and not on reducing the outage risks of HILP events. The major contributions of this paper are as follows:

- *Risk-averse Two-stage Stochastic Programming for distribution grid resilience:* Existing literature on resilience-driven planning includes risk-neutral formulations assuming equal probability events. We introduce a risk-based planning approach that appropriately samples the event probabilities and their impacts on the power grid components and system. The two-stage stochastic programming problem is formulated as a MILP problem where Stage-1 decisions are the infrastructure planning decisions optimized to reduce the risks of outages due to HILP events assuming optional operational phase decisions. The Stage-2 problem models the operational phase and solves the optimal system response (restoration strategy) for the specified resource allocation (from stage-1) and a given event realization.
- *Probabilistic scenario generation and smart scenario reduction strategy:* We propose a scenario generation approach using Monte-Carlo simulations that appropriately models a given regional wind profile and its impacts on the distribution system. Next, for scenario reduction, we propose a smart scenario selection strategy based on the average loss of load representing several Monte-Carlo trials. We also extensively validate the robustness of the proposed scenario generation and reduction approach using multiple out-of-sample scenario sets for different simulation case studies.
- *Trade-off analysis on risk minimization vs. expected loss minimization:* Different case studies are presented to identify the trade-off of adopting risk-neutral vs. risk-averse policies in the planning decisions. The analysis can provide insights into adopting risk-driven solutions when the utmost priority is to maintain an uninterruptible power supply to critical customers during extreme weather events. The results suggest that risk-averse planning tends to incur higher costs to meet the resilience objectives during HILP events; however, they are more likely to restore the critical loads during the HILP events. On the other hand, risk-neutral planning decisions, while more cost-efficient, end up restoring fewer critical loads during HILP events.

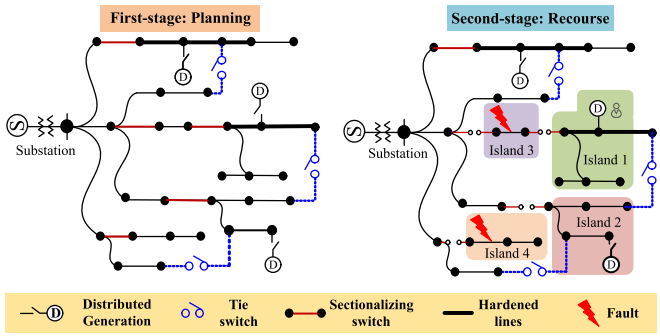


Fig. 1. Two-stage planning framework example for a specific scenario.

D. Manuscript Organization

The remainder of the paper is as follows. Section II highlights the mathematical details for a general two-stage risk-averse stochastic problem along with the general representation of the long-term planning model. Section III describes the overall risk-averse stochastic planning framework with modeling details. Results and analysis are presented in Section IV followed by conclusions in Section V.

II. Mathematical Background

A. Long-Term Planning Model Representation

A power distribution network can be graphically represented as $G(V, E)$, where the vertices V represent the buses or nodes while the distribution lines are represented by the edges E. Fig. 1

shows a general representation of the two-stage planning framework. The overall objective of the two-stage framework is to identify the first-stage optimal planning decisions that minimize the expected operational cost in the second stage. In this work, DG siting and sizing are the planning decisions whereas the second stage objective is to minimize the prioritized load loss once a scenario is realized. DGs with grid-forming inverters are assumed in this work. Such grid-forming DGs can be used for intentional islanding when some area of the distribution grid gets disconnected from the system due to an extreme event.

It is important to understand that the two stages are not decoupled but rather solved as a single optimization problem. The second stage problem minimizes the operational cost for each scenario and hence its objective function. Note that the two-stage objective function is a random variable. Thus, determining the optimal planning decision is the problem of comparing random cost variables as a function of the planning cost and the operational cost. Furthermore, it is assumed that the uncertain scenario realizations in the second stage have some form of a probability distribution. Here, we use regional wind profiles to demonstrate uncertain fault scenarios. The overall framework should provide planning decisions that remain optimal for every realization of scenarios in the second stage. In Fig. 1 it can be seen that once faults occur in the system, the tie switches, and sectionalizing switches are toggled to isolate the faulted sections i.e., Island 3 and Island 4. Furthermore, DGs form two islands i.e., Island 1 and Island 2, and continuously supply the loads inside the island. The connection of DGs is represented by a virtual switch as discussed in [31].

The two-stage problem is formulated as a risk-averse stochastic optimization problem in which the first stage problem minimizes the cost of planning and the weighted combination of the expected value and the $CVaR$ of the second stage problem. The second stage problem is the operational stage that minimizes the total prioritized loss of load for every scenario realization.

B. Two-Stage Stochastic Optimization

A general two-stage stochastic integer programming model can be formulated as [32]:

$$f(x) := \min_x c^T x + \mathbb{E}_P[Q(x, \mathcal{E})] \tag{1}$$

subject to,

$$Ax = b, \quad x \in \mathbb{R}^+_{m1} \times \mathbb{Z}^+_{n1}$$

where

$$Q(x, \mathcal{E}) := \min_y q^T y \tag{2}$$

subject to,

$$Wy = h - Tx \quad y \in \mathbb{R}^+_{m2} \times \mathbb{Z}^+_{n2}$$

where x is the first stage decision variable and y is the second stage decision variable, $\mathcal{E}$ refers to the set of uncertain data (or scenarios) with a known probability distribution P, and (q, h, T, W) are scenario-dependent variables which vary for each $\xi \in \mathcal{E}$. The objective in a general two-stage stochastic optimization is to solve (1) which seeks a first-stage decision x that minimizes the first stage cost and the expected cost of the second stage, $Q(x, \mathcal{E})$. The second stage decisions are also known as the recourse decisions and are scenario-dependent. The algorithms and model formulation in two-stage stochastic optimization depends on the stage variable types.

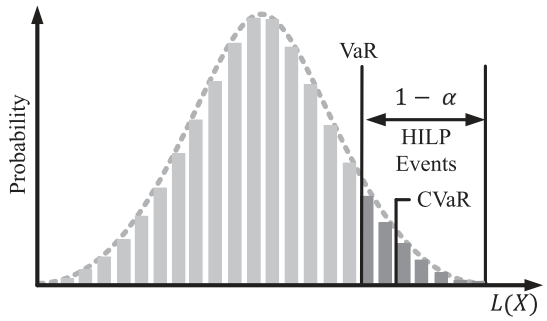


Fig. 2. VaR and CVaR representation for a discrete loss function. HILP events are the ones with $(1 - \alpha)$ probability of occurrence.

C. Optimization of Conditional Value-at-Risk

The general stochastic optimization model considers only the expected cost of the second stage as shown in (1) and does not directly incorporate the tail of the probability distribution. Unlike routine outages caused by known and credible threats, adequately anticipating and predicting system performance during HILP events is inherently difficult as they are rare [33]. While resilience metrics similar to reliability measures such as expected energy not served (EENS), loss of load expectation (LOLE), and service availability [34], [35], [36] have been investigated, these measures mostly provide a reliability-oriented view and do not explicitly quantify expected system performance under unseen HILP events. Thus, it is desired to include tail probability events when planning for resilience to reduce the impacts of HILP events on system outages. There are several metrics used to quantify the tail probabilities or risks; VaR and CVaR are commonly used risk metrics in the domain of financial risk management. Fig. 2 shows a discrete distribution of a loss function $L(X)$ along with its VaR and CVaR. VaR is the α-quantile of the distribution function whereas CVaR is the expected value of the remaining $1 - \alpha$ region that represents the HILP events. Interestingly, CVaR can be formulated as an optimization problem if the random variable under consideration is discrete [37]. The CVaR optimization problem is shown in (3).

$$CVaR_\alpha(X) = \min \left\{ \eta + \frac{1}{1-\alpha} \mathbb{E}([(X - \eta)]_+ : \eta \in \mathbb{R} \right\} \tag{3}$$

where, X is a random variable with N discrete scenarios $x_1, x_2, \ldots, x_n$ each having a probability of $p_1, p_2, \ldots, p_n$ respectively, $\alpha \in [0, 1)$ is the confidence level which gives the VaR of X (VaR_α), and η be the VaR_α of X. Here, η is independent of probability and is the same for each realization of X. Here, $[(X - \eta)]_+ = \max\{X - \eta, 0\}$ which represents the

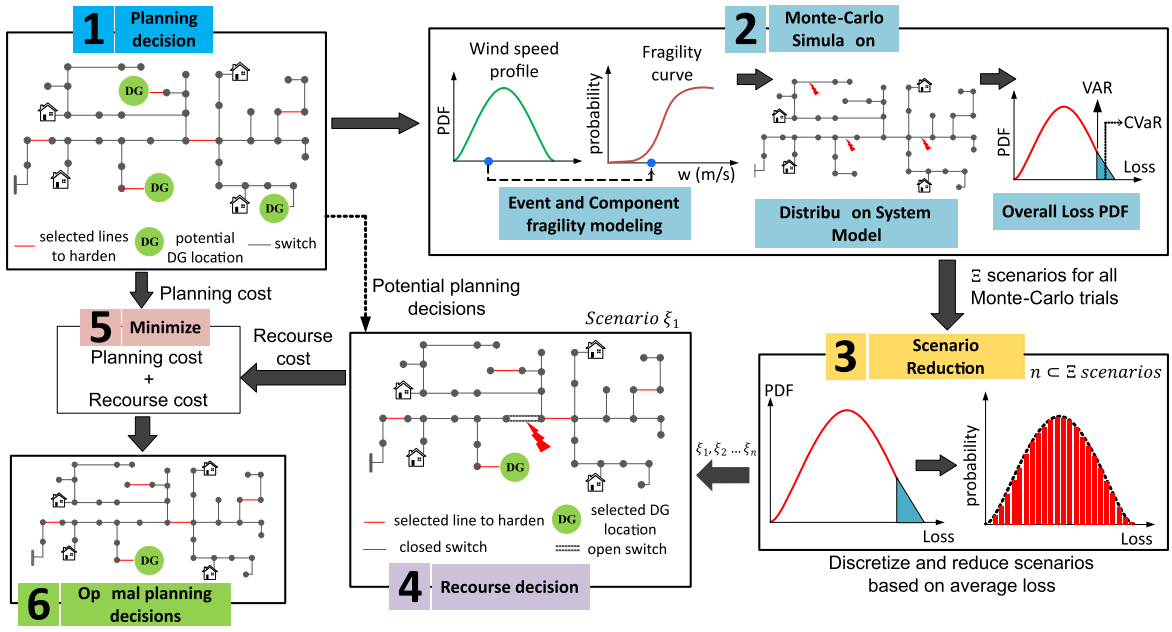


Fig. 3. Overall architecture of risk-averse two-stage planning problem. The first stage seeks the optimal planning decisions that minimize the expectation and risk of the recourse cost in the second stage for several scenario realizations. The scenarios are generated using Monte-Carlo simulation and reduced based on average loss for each scenario.

point-wise maximum of convex functions and hence, (3) can be reformulated using its epigraph form as:

$$CVaR_\alpha(X) = \min\left\{\eta + \frac{1}{1-\alpha}\sum_{\xi\in\mathcal{E}} p_\xi \nu_\xi\right\} \tag{4}$$

subject to,

$$\nu_\xi \geq x_\xi - \eta \quad \eta \in \mathbb{R}, \nu \in \mathbb{R}^n_+ \tag{5}$$

where, p_ξ is the discrete probability of each scenario ξ, and ν_ξ is an excess variable which ensures that $CVaR_\alpha$ is calculated only for realizations beyond VaR_α for each scenario ξ. Here, (4) is linear and can be solved using existing linear programming techniques.

D. Risk-Based Stochastic Optimization

As discussed earlier, the general two-stage stochastic optimization framework neglects the tail events of the probability distribution that tend to have higher risks in terms of loss in the system. $CVaR_\alpha$ is the coherent risk metric that can quantify the associated risks given the probability distribution of the losses associated with an event. In (4), $CVaR_\alpha$ is formulated as an optimization problem and hence can be introduced in a general two-stage stochastic optimization framework given in (6)–(7) [38].

$$\min_x c^T x + (1-\lambda)\mathbb{E}[Q(x,\mathcal{E})] + \lambda CVaR_\alpha(Q(x,\mathcal{E})) \tag{6}$$

where

$$\mathbb{E}(\mathcal{Q}(x,\mathcal{E})) = \sum_{\xi\in\mathcal{E}} p_\xi \mathcal{Q}(x,\xi)$$

$$CVaR(\mathcal{Q}(x,\mathcal{E})) = \eta + \frac{1}{1-\alpha}\sum_{\xi\in\mathcal{E}} p^\xi \nu^\xi \tag{7}$$

where $\lambda \in [0,1]$ is the risk multiplier or factor that defines the trade-off between $\mathbb{E}[\cdot]$ and $CVaR_\alpha(\cdot)$. By selecting different values of λ, the first stage decisions are termed as either risk-averse ($\lambda = 1$), risk-seeking/mean-risk ($\lambda = 0.5$), or risk-neutral ($\lambda = 0$). The formulation in (6) not only minimizes the expected loss but also the $CVaR_\alpha$ of those loss distributions depending on the value of λ. Hence, with this formulation, one can identify how planning decisions vary with the risk avoidance potential.

III. Risk-Based Resilience-Oriented Distribution System Planning

The overall architecture of the proposed method is shown in Fig. 3. Only wind-related events are used in this work and the probability distribution of extreme wind events is considered to generate the scenarios. Monte Carlo simulations (MCS) are conducted to identify the impact of probabilistic events, and an appropriate scenario reduction method is implemented to identify representative scenarios. Finally, the long-term planning problem is solved in a two-stage stochastic optimization setting based on a selected number of scenarios. This section details the overall optimization process along with detailed problem formulations.

A. Probabilistic Scenario Generation and Reduction

An event is characterized by its intensity and probability of occurrence. Fig. 4 shows the event probability distribution for a windfall in three different regions observing extreme, high, and normal wind profiles. The extreme regional wind profile

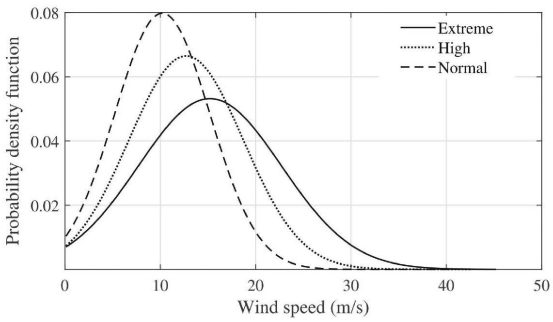


Fig. 4. Regional wind profile.

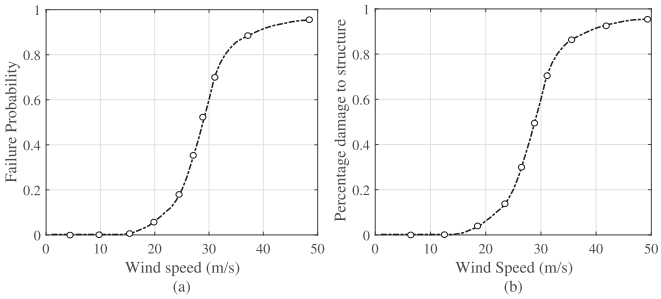


Fig. 5. Component level impact assessment for an extreme event: (a) fragility curve for a wind profile [43], (b) prototype curve fit models for a percentage of equipment damaged as a function of wind speed [41].

is used to model extreme events in this work. For simplicity, only distribution lines are assumed to be affected by wind in this work. Although wind-related events have spatiotemporal dynamics [39], we assume that for a distribution system, that covers a small region, the wind speed for the entire region is the same. MCS is performed for each wind speed case so as to also include the extreme tail probability events.This process is represented by block 2 in Fig. 3. For each wind speed scenario u, the component level failure probability $p_f(u)$ determines the operational state of a particular component in the distribution grid. Component level fragility curves [40] or prototype curve fit models [41] can be used to model the impacts of extreme events such as hurricanes or other high-speed wind events on power systems. In this work, we have used the component fragility curve that maps the probability of failure of distribution system components conditioned on the intensity of the hazard (e.g., a wind speed). An example of the fragility curve is shown in Figs. 5 a and 5 b. The fragility curve values are randomly selected for simulation purposes; however, if available, empirical data can be used to adjust the parameters [42].

$$p_f(u) = \begin{cases} P_f^n, & \text{if } u < u_{critical} \\ P_f(u), & \text{if } u_{critical} < u < u_{collapse} \\ 1, & \text{if } u > u_{collapse} \end{cases}$$

where, $P_f(u)$ is the failure probability of a component as a function of wind speed, u; P_f^n is the failure rate at normal weather conditions; $u_{critical}$ is the wind speed at which the failure probability rapidly increases. The equipment has a negligible probability of survival at $u_{collapse}$. In the problem formulation, the failures due to an extreme event are modeled as open or faulted line/switch variables as discussed later in this section. The location of these faults is determined based on the fragility curve for wind speed greater than $u_{critical}$.

Several MCS are conducted to obtain the system loss associated with the failure probability of a component in the distribution grid. In this work, the amount of active prioritized load (kW) disconnected from the system following an event is considered the system loss for a particular wind speed scenario. The critical loads (CL) should always be prioritized in any restoration methods. Hence, higher weights are assigned to the CLs that reflect a higher value of prioritized load loss if any CLs are disconnected and not restored. The average prioritized load loss is then mapped onto the regional wind profile PDF to get a probabilistic representation of the loss in the system when subjected to a given weather event.

MCS provides an extremely large number of scenarios. One major challenge in any stochastic optimization setting is handling many scenarios within the optimization framework. Furthermore, the solution should be optimal for all scenarios that make the stochastic problem computationally intractable. Existing works use special sampling techniques such as stratified sampling [44] or importance sampling [45] to include the tail probability scenarios in the optimization model appropriately. Distance-based scenario reduction methods have also been used where a probabilistic distance measure is minimized to obtain a reduced scenario distribution that closely represents the overall scenario distribution [46]. We introduce a new approach to scenario reduction inspired by stratified sampling and distance reduction methods. The proposed approach uses stratification to sample representative scenarios for each wind speed and generates a reduced scenario distribution that closely matches the original scenario distribution.

In this work, the overall number of scenarios is reduced by selecting a representative scenario for each wind speed based on the average Monte-Carlo loss. This process is represented by block 3 in Fig. 3. Let N_u be the total discrete wind speeds under consideration, $N_{\xi,u}$ be the number of scenarios obtained from MCS for each wind speed u, and $L^u_{avg} = \mathbb{E}(L_{\xi,u})$ be the average prioritized load loss in kW corresponding to $N_{\xi,u}$ scenarios. Let $\Xi = N_{\xi,u} \times N_u$ be the total number of scenarios for the entire MCS. Note that we cannot randomly select a subset of these scenarios as it significantly degrades the accuracy of the optimization solutions. Here, we use a unique sampling technique to drastically reduce the number of scenarios while maintaining the representation of the overall scenarios described next. If ξ_u is a representative scenario for all $N_{\xi,u}$ scenarios corresponding to u, then ξ_u is selected such that the prioritized load loss in the system due to ξ_u ($L^u_{\xi_u}$) is the one nearest to L^u_{avg}. In the case of multiple scenarios with losses nearing L^u_{avg}, one of the scenarios is randomly selected as ξ_u from the identical scenario representations. The proposed scenario reduction technique reduces the total number of scenarios to N_u from Ξ such

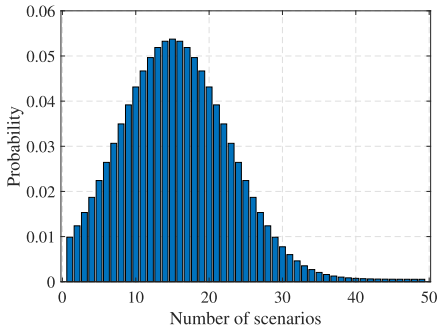


Fig. 6. A set of 49 representative scenarios with respective probabilities.

that p_ξ corresponds to the wind speed profile. Fig. 6 shows a set of $N_u = 49$ representative scenarios obtained from the overall Monte-Carlo scenarios based on the abovementioned methods. This smart scenario selection strategy ensures the practical realization of the second stage problem while incorporating HILP events within the scenarios. The overall scenario generation and reduction process is detailed in Algorithm 1. This work does not apply the restoration schemes in the scenario generation and reduction phase. Hence, the obtained scenarios are base case scenarios that only give information on the amount of prioritized load loss in the network based on each of wind scenario.

Algorithm 1: Probabilistic Approach for Scenario Generation and Reduction.

Available information and initialization:
regional wind-speed profile, distribution system model, component fragility curve with $p_f(u)$, line failure status L_F, representative scenario ξ_u
MCS for Scenario Generation:
for $u = 1, 2, ..., N_u$ **do**
  $N_{\xi,u} = 0$
  while *enough convergence trials* **do**
    $N_{\xi,u} = N_{\xi,u} + 1$
    if $p_f(u) \geq$ *rand* $[\mathcal{U}(0,1)]$ **then**
      L_F = 1 (True)
    else
      L_F = 0 (False)
    end
    update distribution system topology using L_F
    calculate $L_{\xi,u}(N_{\xi,u})$
  end
  $L^u_{avg} = \mathbb{E}[L_{\xi,u}]$
end
map L^u_{avg} with the respective u to create a loss PDF
Scenario Reduction:
for $w = 1, 2, ..., N_u$ **do**
  for $\xi = 1, 2, ..., N_{\xi,u}$ **do**
    // temporary variable
    ξ'_u = ξ
    if $|L^u_{avg} - L^u_{\xi'_u}| \approx \min(|L^u_{avg} - L^w_{\xi'_u}|)$ **then**
      $\xi_\mathbf{u} \leftarrow \xi'_u$
    end
  end
  // randomly sample one representative scenario
  $\xi_u \leftarrow$ rand ($\xi_\mathbf{u}$, 1)
end

B. Two-Stage Stochastic Optimization Problem Formulation

We detail the two-stage stochastic optimization problem for resilient distribution system planning.

1) Objective Function: In this paper, the resilience-driven distribution system planning problem is formulated as a two-stage stochastic optimization problem where the overall objective function can be defined as:

$$\min(1-\lambda)\mathbb{E}(Q(\delta,\mathcal{E})) + \lambda CVaR_\alpha(Q(\delta,\mathcal{E})) \tag{8}$$

where,

$$\mathbb{E}(Q(\delta,\mathcal{E})) := \left(\sum_{\xi\in\mathcal{E}}\sum_{i\in\mathcal{B}_S}\sum_{\phi\in\{a,b,c\}} (1-s_i^\xi)\, w_i\, P_{Li}^{\phi,\xi}\right)$$

$$CVaR_\alpha(Q(\delta,\mathcal{E})) := \left(\eta + \frac{1}{1-\alpha}\sum_{\xi\in\mathcal{E}} p^\xi \nu^\xi\right)$$

The problem objective in the first stage is to minimize the weighted sum of expected value and $CVaR_\alpha$ of the second stage cost, represented by $Q(\delta,\mathcal{E})$. To analyze the trade-offs, this formulation has not used minimization of planning cost. Instead, we use a budget constraint and observe the associated trade-offs for risk-averse and risk-neutral decisions when system planners have a limited investment budget. The objective of the second stage of the problem, $Q(\delta,\mathcal{E})$, is to minimize the prioritized load loss or maximize the restoration of prioritized loads for each $\xi \in \mathcal{E}$. The second stage costs correspond to the optimal restoration decisions once a scenario has been realized. Hence, each variable corresponding to the second stage of the problem is scenario-dependent. Here, DG location (δ_i^{DG}) and size of the DG (β_i^{DG}) are the first stage decision variables. $P_{Li}^{\phi,\xi}$ represents the active power demand at node i for phase ϕ and scenario ξ and $s_i^\xi \in 0, 1$ is the load pick-up status variable that determines whether the load at node i is picked up or not. The CLs are prioritized by a weight variable w_i. Since the CLs are critical for any scenario, w_i remains the same for all scenarios. Furthermore, the scenarios have a specific probability, p_ξ, associated with them, which comes from the scenario reduction method discussed before. The parameters for $CVaR_\alpha(Q(\delta,\mathcal{E}))$ are defined similarly as discussed in Section. II.

2) First Stage Constraints: The first stage constraints correspond to the planning decisions made in the first stage. In this work, the per unit cost for DG installation and sizing is assumed to be the same for each location; these assumptions can be easily relaxed. Constraint (9a) ensures that the total cost of DGs should be between \$[0, $\mathcal{C}^{DG}_{\max}$] regardless of the cost of installation in an individual location. This gives the freedom of utilizing the overall budget for a single big-sized DG or distributing the budget to multiple smaller-sized DGs. Constraint (9a) contains a non-linear term $\delta_i^{DG} \times \beta_i^{DG}$ which is linearized using big-M method as discussed in [47]. Constraint (9b) restricts the DG

location variable to binary. The DG location variable δ_i^{DG} is 1 if a DG is located in node i, else 0. Furthermore, constraint (9c) ensures that VaR_α for the distribution of load loss in the second stage is a real number. Furthermore, VaR_α is independent of scenarios and is obtained with the solution of the first stage.

$$\sum_{i\in\mathcal{B}_{DG}} c_i^{DG}\delta_i^{DG}\beta_i^{DG} \le \mathcal{C}_{\max}^{DG} \tag{9a}$$

$$\delta_i^{DG} \in \{0,1\} \tag{9b}$$

$$\eta \in \mathbb{R} \tag{9c}$$

3) Second Stage Constraints: The second stage of the stochastic optimization problem is the operational stage in which DG-assisted restoration is performed for each ξ. The inner-loop operational stage consists of several constraints corresponding to the restoration problem [31]. Since the second-stage variables change with respect to each scenario, each of these variables has ξ to differentiate them from the first-stage variables.

Connectivity constraints:

- Constraint (10a) ensures that a load is picked up if and only if it is connected to an energized bus, v_i. Similarly, based on the constraint (10b) loads connected to non-switchable buses will also be picked up if the corresponding bus is energized.
- The line energization status can be observed through the constraints set (11). According to constraint (11a), a switchable line without fault is energized if any of the buses connecting the line is energized. On the other hand, (11b) ensures that a non-switchable line connected to any energized bus is also energized. Finally, constraint (11c) ensures that a line experiencing a fault is disconnected from the grid.

$$s_i^\xi \le v_i^\xi,\ \forall i \in \mathcal{B}_S \tag{10a}$$

$$s_i^\xi = v_i^\xi,\ \forall i \in \mathcal{B}\backslash\mathcal{B}_S \tag{10b}$$

$$\delta_e^\xi \le v_i^\xi,\ \delta_e^\xi \le v_j^\xi,\ \forall e \in \mathcal{L}_S\backslash\mathcal{L}_F^\xi \tag{11a}$$

$$\delta_e^\xi = v_i^\xi = v_j^\xi,\ \ \forall e \in \mathcal{L}\backslash(\mathcal{L}_S \cup \mathcal{L}_F^\xi) \tag{11b}$$

$$\delta_e^\xi = 0,\ \ \forall e \in \mathcal{L}_F^\xi \tag{11c}$$

Power Flow Constraints: In this work, a three-phase unbalanced linearized power flow model is used in the optimization framework [48]. Since we are solving a long-term planning model with CL restoration in the second stage, the linearized model is sufficiently accurate and applicable for our problem [31]. Furthermore, the power flow will only be valid for the energized section of the system. Hence, the power flow equations are coupled with line and bus energization variables to appropriately represent them in the second stage problem.

- Constraints (12a)–(12d) represent the three-phase unbalanced linearized power flow equations. The equations are coupled with the line decision variable δ_e and load-pick variable s_i. Constraints (12a) and (12b) represent the active and reactive power flow for each of the energized lines. Constraint (12c) is the voltage equation for non-switchable lines whereas (12d) represents the voltage equation for a set of lines that are switchable. Constraint (12d) is coupled with δ_e to ensure that the voltage drop applies only if the switch is closed. The non-linear terms associated with the power flow equations are linearized using the big-M method [47].

$$\sum_{e:(i,j)\in\mathcal{L}} \boldsymbol{P}_e^\xi = s_j^\xi\, \boldsymbol{P}_{Lj}^\xi + \sum_{e:(j,i)\in\mathcal{L}} \boldsymbol{P}_e^\xi \tag{12a}$$

$$\sum_{e:(i,j)\in\mathcal{L}} \boldsymbol{Q}_e^\xi = s_j\, \boldsymbol{Q}_{Lj}^\xi + \sum_{e:(j,i)\in\mathcal{L}} \boldsymbol{Q}_e^\xi \tag{12b}$$

$$\boldsymbol{U}_i^\xi - \boldsymbol{U}_j^\xi = 2\left(\tilde{\mathbf{r}}_e\boldsymbol{P}_e^\xi + \tilde{\mathbf{x}}_e\boldsymbol{Q}_e^\xi\right),\ \ \forall e \in \mathcal{L}^\xi\backslash\mathcal{L}_S^\xi \tag{12c}$$

$$\delta_e^\xi\left(\boldsymbol{U}_i^\xi - \boldsymbol{U}_j^\xi\right) = 2\left(\tilde{\mathbf{r}}_e\boldsymbol{P}_e^\xi + \tilde{\mathbf{x}}_e\boldsymbol{Q}_e^\xi\right),\ \ \forall e \in \mathcal{L}_S^\xi. \tag{12d}$$

where $\tilde{\mathbf{r}}_e = \text{Real}\{\alpha\alpha^H\} \otimes \mathbf{r}_e + \text{Im}\{\alpha\alpha^H\} \otimes \mathbf{x}_e$, $\tilde{\mathbf{x}}_e = \text{Real}\{\alpha\alpha^H\} \otimes \mathbf{x}_e + \text{Im}\{\alpha\alpha^H\} \otimes \mathbf{r}_e, \alpha = [1\ e^{-j2\pi/3}\ e^{j2\pi/3}]^T$

Operational Constraints: The operational constraints of the second stage problem are related to the distribution system topology and voltage limits. The distribution system operates in a radial fashion. Hence, the topology of the distribution grid should be radial at all times. Furthermore, the nodal voltages should be within the specified limits at all times.

- A radial configuration in any distribution system consists of several sectionalizing and tie-line switches. In this work, virtual edges are assumed to supply the power from DGs in case of an islanded mode of operation. In any faulted network, a radial configuration is maintained by toggling any of the switches to avoid the formation of loops or cycles. Constraint (13a) ensures that at least one of the switches is open in a cycle. In this work, a brute-force approach is applied to count and store the number of cycles in the distribution system and Constraint (13a) is enforced on each of the cycles so that the system operates in a radial fashion. The process of counting and storing cycles is completely offline and does not affect the computational complexity of the stochastic optimization procedure.
- The voltage limit on each of the buses should be within the ANSI C84.1 standard is ensured by (13b). In this work, $\boldsymbol{U}^{\min}$ and $\boldsymbol{U}^{\max}$ are set as $(0.95)^2$ and $(1.05)^2$ respectively for each of the phases. Since the limits make sense only for the buses that are energized, the limits are coupled with v_i^ξ.
- For the purpose of reconfiguration, it is required that the power flow through an open switch is zero. If $\delta_e^\xi = 0$ for any $e : i \to j$ then constraint (14) ensures that the power flow through that line is zero. If not, box constraints on the power flow are enforced where $\underline{M}_p = -\overline{M}_p$ and $\underline{M}_q = -\overline{M}_q$.

$$\sum_{e\in\mathcal{L}_c} \delta_e^\xi \le |\mathcal{L}_c| - 1, \qquad \forall e \in \mathcal{L}_c \tag{13a}$$

$$v_i^\xi\boldsymbol{U}^{\min} \le \boldsymbol{U}_i^\xi \le v_i^\xi\boldsymbol{U}^{\max},\ \ \forall i \in \mathcal{B} \tag{13b}$$

$$\delta_e^\xi\left[\underline{M}_p\ \ \underline{M}_q\right] \le \left[\boldsymbol{P}_e^\xi\ \ \boldsymbol{Q}_e^\xi\right] \le \delta_e^\xi\left[\overline{M}_p\ \ \overline{M}_q\right],\ \ \forall e \in \mathcal{L}_S^\xi \tag{14}$$

DG Constraints:

- Virtual switches represent the connection of DGs in islanded mode. The virtual edge should only be connected if a DG is located at the specific node which is ensured by (15a).
- The two stages in any stochastic optimization framework are bound by the non-anticipativity constraints [32]. This means that the location and size of the DGs should be the same for every scenario realized in the second stage of the problem. This non-anticipativity nature of the first stage decision variables is presented in (15b).
- The in-flow power of each of the DGs should be less than or equal to the size of the DG. Since the DGs are connected using virtual edges, δ_e^{ξ} is coupled with this constraint as given in (15c).

$$\delta_e^{\xi} \leq \delta_i^{DG},\ \forall e \in \mathcal{L}_S^v,\ \forall i \in \mathcal{B}_{DG} \tag{15a}$$

$$\delta_i^{DG} = \delta_i^{DG,\xi},\ \delta_i^{DG}\beta_i^{DG} = \delta_i^{DG,\xi}\beta_i^{DG,\xi} \quad \forall i \in \mathcal{B}_{DG}, \forall \xi \in \mathcal{E} \tag{15b}$$

$$\sum_{\phi \in \{a,b,c\}} P_e^{\phi,\xi} \leq \delta_e^{\xi}\delta_i^{DG}\beta_i^{DG},\ \forall e \in \mathcal{L}_S^v, \forall i \in \mathcal{B}_{DG} \tag{15c}$$

$CVaR_\alpha$ Constraints: The VaR_α is obtained as a solution in the first stage and is represented by η. VaR_α and $CVaR_\alpha$ correspond to the distribution of optimal solutions obtained in the second stage of the problem. $CVaR_\alpha$ represents $1-\alpha$ part of distribution beyond VaR_α. Hence, as discussed in (4) an excess variable is obtained for each scenario ν_ξ such that it corresponds $1-\alpha$ part of the distribution beyond VaR_α. Furthermore, this excess variable must be a positive real number. These constraints are represented by (16).

$$\nu_\xi \geq x_\xi - \eta, \quad \nu \in \mathbb{R}_+^n \tag{16}$$

IV. Results and Analysis

The effectiveness of the proposed risk-based long-term planning model is verified on a modified IEEE 123-bus case, see Fig. 7. Several case studies with multiple DG locations, variable numbers of DGs, and varying risk preferences are presented with detailed analyses of the results. The two-stage problem without DG-based restoration is referred to as the base case which is then compared with other case studies. Furthermore, to analyze the planning decisions better, we create a new test case upon hardening 15 randomly selected lines, as shown in Fig. 7. The fragility curves of hardened lines are adjusted so that their outage probability for any extreme event is less than the case when they are not hardened. For CLs, $w_i = 10$ whereas, for non-critical loads, $w_i = 1$. Thus the second stage cost reflects the total amount of prioritized loss of load (in kW). The total non-prioritized demand of the system is $P_D = 4485$ kW and the prioritized demand is $\sum_{i \in \mathcal{V}} w_i P_{Li} = 20775$ kW. In this paper, we use prioritized demand to analyze the results for different cases.

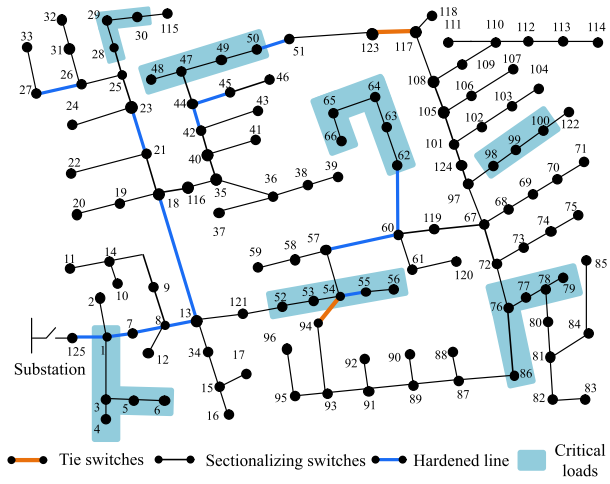


Fig. 7. Modified IEEE 123-bus test case.

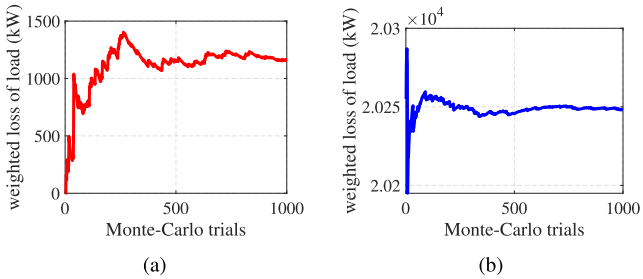


Fig. 8. Moving average of loss of load obtained for 1000 Monte-Carlo trials (a) without hardening when $u = 15$ m/s and $p_f(u) = 0.002$ and (b) with line hardening when $u = 40$ m/s and $p_f(u) = 0.915$. For each wind speed scenario, it can be guaranteed that the prioritized load loss converges after 1000 Monte-Carlo trials.

The two-stage stochastic integer programming model is formulated using PySP package in Pyomo [49]. Scenario generation and reduction using the Monte-Carlo method are implemented in MATLAB2020a. The entire simulations are carried out on a PC with a 3.4 GHz Intel i7-6700 CPU and 16 GB RAM. The proposed two-stage stochastic problem is solved as a single large mixed-integer linear programming problem for each presented case study.

A. Scenario Generation and Reduction

The wind event scenarios are generated and reduced using methods discussed in Section III-A. Using the wind speed profile for extreme wind events and failure probability of distribution lines, several trials of MCS simulation are conducted for sampled wind speeds [42]. For this experiment, $N_u = 49$ wind speeds are sampled from the wind speed profile and it was experimentally verified that 1000 Monte-Carlo trials are enough to obtain a converged value of prioritized loss of load in the distribution grid corresponding to each u. Fig. 8 shows the moving average of prioritized loss of load for 1000 Monte-Carlo trials for the base case without hardening and with hardening. It can be seen that the value of the loss is fairly converged in 1000 trials for both

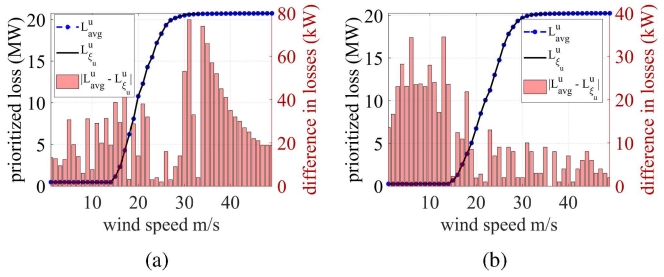


Fig. 9. Comparison of prioritized load loss obtained for two sets of reduced scenarios a) without hardening b) with hardening.

TABLE II
BASE CASE EXPECTED VALUE AND $CVaR_\alpha$ OF PRIORITIZED LOAD LOSS

line hardening	$\mathbb{E}(Q(\delta,\mathcal{E}))$ in kW	$CVaR_\alpha(Q(\delta,\mathcal{E}))$ in kW
No	5982.57	20601.58
Yes	4541.76	19839.39

cases. Since 1000 trials are conducted for each u, $N_{\xi,u} = 1000$. Hence, the total number of scenarios generated through MCS, $\Xi = 49 \times 1000 = 49000$.

Fig. 9 represents the comparison between L^u_{avg} and L^u_ξ for the test case without hardening and with line hardening. The loss due to reduced scenarios is very close to that of the actual representative scenarios for each u. The y-axis on the right represents the value of $|L^u_{avg} - L^u_\xi|$. It can be seen that the maximum difference occurs at $u = 31$ m/s in Fig. 9(a) and has a value of about 78 kW which is $<0.5\%$ of total prioritized demand. The difference in their values comes from the fact the L^u_{avg} is obtained by averaging 1000 different realizations of ξ for a specific u whereas L^u_ξ is the prioritized load loss for a specific failure scenario ξ corresponding the same u. Furthermore, it should be noted that HILP events (tail events) are also sampled in this reduction method which makes this approach highly suitable for resilience planning problems.

B. Risk-Averse Long-Term Planning

The reduced scenarios from the method mentioned above represent several scenarios that can occur on the distribution grid. Each scenario represents the line failure status due to a particular wind speed (u). In this long-term planning problem, 6 DG locations are pre-selected as potential locations for the placement of DG units. It is to be noted that the candidate locations are not the final DG locations. They are potential locations that can be used to install a DG as per the solution of the proposed optimization framework. The selected potential DG locations are nodes 95, 122, 39, 85, 56, and 66. However, the DG locations are decided by the optimization model and $\delta^{DG}_i = 1$ if and only if $\beta^{DG}_i > 0$. From the operator's perspective, it is often practical to have a limited budget while planning the siting and sizing strategies for DGs. The total budget is constrained so that the sum of the DG units is less than or equal to 900 kW. For risk-driven problems, α is set at 0.95, meaning that 5% tail scenarios (HILP) are considered to have greater risks.

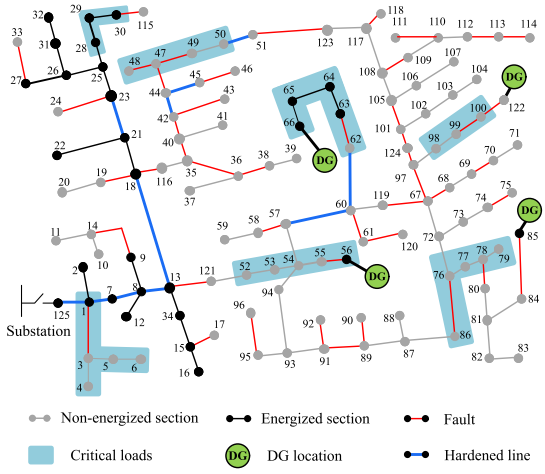
Table II shows $\mathbb{E}(Q(\delta,\mathcal{E}))$ and $CVaR_\alpha(Q(\delta,\mathcal{E}))$ when no DG-based planning measures are used (base case). Since $\alpha = 0.95$, the $CVaR_\alpha$ represents the 5% of the tail probability cases. This means that when those 5% scenarios are realized, on average, the prioritized loss of load is 20601.58 kW (with no line hardening measures in place). The losses improve to 19839.39 kW when a few lines are hardened, as shown in Fig. 7. Furthermore, the expected values of prioritized load loss, calculated over the entire scenarios, are 5982.57 kW and 4541.76 kW for the respective cases mentioned above. However, when an operator is planning to enhance the grid's resilience, the 5% of those scenarios are extremely important because the system needs to withstand or adapt to those events to maintain a constant supply of power to the CLs. Hardening a few lines is already proving to be a potential solution to minimize the expected and $CVaR_\alpha$ of the prioritized load loss. However, in the case of an islanding situation when CLs are disconnected from the system when a fault occurs, DG-based planning strategies have proven to be successful in maintaining an uninterrupted power supply to the CLs [5].

To identify the trade-off among different DG-based planning strategies, 6 locations — 39, 56, 66, 85, 95, and 122 — are selected as potential DG locations. The planning problem is then solved as a two-stage stochastic problem as discussed in Section III. First, we discuss the results for the risk-neutral case ($\lambda = 0$). The existing resilience-based planning methods, [12], [18], [19], are focused on the risk-neutral case and used as a comparison for this work. The overall capacity of each of the DGs is shown in Table III. For risk-neutral planning without line hardening measures, no DGs are required to be placed on nodes 36 and 95. However, for mean-risk and risk-averse situations, the planning strategies change significantly. For risk-involved strategies, it is required to place DGs on nodes 39 and 95 while reducing the DG sizes for the rest of the nodes as shown in Table III. Hence, the trade-off of including risk minimization in the objective is to increase the number of DG units in the system. This can be fruitful for extreme event scenarios when picking up some of the CLs is required, even though it increases the expected value of prioritized load loss. Table III also shows the expected value and $CVaR_\alpha$ of prioritized CLs picked up by different planning strategies. It can be seen that the expected value of prioritized CLs picked up does not change much regardless of the risk preference. However, for risk-based strategies (both mean risk and risk-averse), $CVaR_\alpha$ of prioritized CLs picked up increases by 200 kW compared to the risk-neutral case.

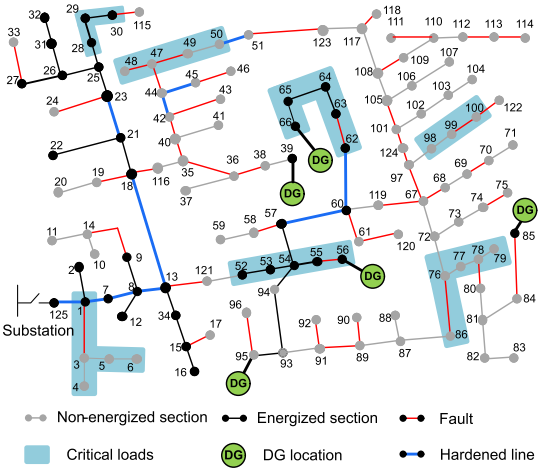
The effect of risk aversion is even more pronounced in the case with the line-hardening strategy. Fig. 10 shows a restoration and planning solution for a specific scenario of HILP nature, $u = 28$ m/s (see Fig. 6). The lines and nodes with black color are the energized section, whereas non-energized sections are represented by gray. Similarly, red lines represent out-of-service lines due to the particular outage scenario. Similar to the restoration for cases without line hardening measures, the risk-neutral solution does not include DGs in nodes 39 and 95. When the objective is risk-neutral ($\lambda = 0$) some of the prioritized critical loads are not picked up in this specific scenario as picking up critical loads in this scenario would not affect the expected value

TABLE III
EXPECTED VALUE AND $CVaR_\alpha$ OF PRIORITIZED LOAD LOSS AND PRIORITIZED CRITICAL LOAD (PCL) PICKED UP FOR DIFFERENT VALUES OF λ. THE DG PLANNING STRATEGY DIFFERS ALONG WITH THE RISK PREFERENCE DEFINED BY λ. ALL OF THE VALUES MENTIONED HERE ARE IN KW

	WITHOUT LINE HARDENING									WITH LINE HARDENING								
	Existing methods [13], [18], [19]			Proposed method						Existing methods [13], [18], [19]			Proposed method					
	$\lambda = 0$			$\lambda = 0.5$			$\lambda = 1$			$\lambda = 0$			$\lambda = 0.5$			$\lambda = 1$		
$\mathbb{E}(Q(\delta,\mathcal{E}))$	3567.12			3586.28			3595.51			2467.46			2463.95			2490.09		
$CVaR_\alpha(Q(\delta,\mathcal{E}))$	19093.89			18885.92			18885.92			18415.04			18160.18			18119.1		
Expectation of PCL picked up	15043.93			15016.64			15006.62			16065.76			16048.17			16026.12		
$CVaR_\alpha$ of PCL picked up	3406.59			3603.06			3603.06			4953.65			5580.72			6295.33		
DG planning strategies	β_{39}^{DG}	β_{56}^{DG}	β_{66}^{DG}	β_{39}^{DG}	β_{56}^{DG}	β_{66}^{DG}	β_{39}^{DG}	β_{56}^{DG}	β_{66}^{DG}	β_{39}^{DG}	β_{56}^{DG}	β_{66}^{DG}	β_{39}^{DG}	β_{56}^{DG}	β_{66}^{DG}	β_{39}^{DG}	β_{56}^{DG}	β_{66}^{DG}
	0	20	370	20	20	340	20	20	370	0	350	330	20	220	255	20	170	330
	β_{85}^{DG}	β_{95}^{DG}	β_{122}^{DG}	β_{85}^{DG}	β_{95}^{DG}	β_{122}^{DG}	β_{85}^{DG}	β_{95}^{DG}	β_{122}^{DG}	β_{85}^{DG}	β_{95}^{DG}	β_{122}^{DG}	β_{85}^{DG}	β_{95}^{DG}	β_{122}^{DG}	β_{85}^{DG}	β_{95}^{DG}	β_{122}^{DG}
	390	0	120	100	300	120	100	270	120	100	0	120	100	305	0	100	280	0



(a)



(b)

Fig. 10. DG sizing and siting solution for a specific scenario with additional hardening measures for (a) risk-neutral and (b) risk-averse planning strategy.

of load served for the overall scenarios. Since the objective is to minimize the expected value of prioritized load loss for entire scenarios, DG at location 95 is not selected. Note that the probability of HILP scenarios is low. Since the expected value contains the product of this probability with the objective function in the restoration phase, the net value is significantly low to affect the overall expected value. However, when the objective is risk-averse, any prioritized load that the nearest possible DG can pick up is given the top priority for any HILP event. For instance, it can be seen that load at node 62 is picked up by DG at node 95 through path 95-93-94-54-57-60-62. Hence, this draws an important conclusion that risk-averse decisions enhance long-term resilience planning by focusing the extreme HILP events. Contrary to the existing methods in [12], [18], [19], the prioritized CLs have a high chance of being picked up when an HILP event is realized by including risk minimization in the objective. However, when attempting to minimize the risk-averse objective (i.e., the $CVaR_\alpha$), we incur an additional DG cost in the overall planning budget to meet the requirements for risk-averse planning. Thus, through the proposed approach and by including $CVaR_\alpha$ minimization in the objective function, prioritized critical loads can be properly restored in case of HILP events. Furthermore, with the changing trade-off between the expectation and the $CVaR_\alpha$ of the prioritized load loss, the expected value *generally* decreases with the increase in λ.

For the case without line hardening, the $CVaR_\alpha$ does not change when moving from mean-risk to risk-averse setting as shown in Table III. It is important to note that tail probability events are also a part of risk-neutral planning strategies. However, the main focus is to minimize the expected loss over the entire scenario, and hence the effect of those tail events is less prominent. With risk-driven strategies, $CVaR_\alpha$ of those tail events are also minimized, and hence the value of $CVaR_\alpha$ decreases with an increase in λ. At some point, $CVaR_\alpha$ saturates as it is impossible to restore some prioritized loads regardless of the planning strategies. On the other hand, for the case with additional line hardening measures, $CVaR_\alpha$ of prioritized load loss further decreases with increasing risk-aversion. This is due to the fact that with line hardening measures, the DGs can pick up more CLs during HILP events, which ultimately reduces the prioritized load loss in the system. Hence, it is clear and obvious that with more number of resources, the $CVaR_\alpha$ can be improved further. However, the trade-off comes with the budget and feasibility. Although it is tempting to harden each and every line and install DGs in each and every location, it is almost impossible for any system operator to allocate the budget accordingly.

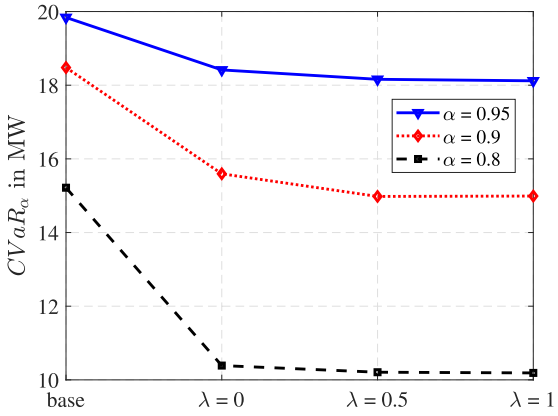


Fig. 11. Comparison of $CVaR_\alpha$ of prioritized loss of load for different values of α and risk preference.

C. Sensitivity Analysis

The value of $CVaR$ depends on several factors such as investment decisions, budget, risk preference, and scenarios under consideration. Here, we present a few of the sensitivity analyses and discuss their impacts on $CVaR$. For simplicity, the analyses are performed only on the system with additional hardening measures already in place.

1) Change in Confidence Level: The risk parameters α and λ can affect the planning decisions. The value of $CVaR_\alpha$ highly depends on α as it defines the number of scenarios to be considered in defining the risk. In other words, α can also be defined as risk percentage. For a higher value of α, the value of VAR_α increases, and hence, $CVaR_\alpha$ represents the scenarios that create greater losses in the system. Similarly, for a smaller α, $CVaR_\alpha$ incorporates a larger number of scenarios with lower losses in risk quantification. Furthermore, as discussed above, an increasing value of λ denotes an increase in risk aversion towards planning decisions. Fig. 11 shows the relation of $CVaR_\alpha$ for the prioritized loss of load for different values of λ and α. As discussed, $CVaR_\alpha$ decreases when more scenarios are considered as risky (characterized by α). Furthermore, for a fixed α, $CVaR_\alpha$ decreases with the increase in the value of λ as more importance are given to risk minimization. Appropriate values for α and λ need to be selected based on planners' risk aversion criteria.

2) Change in Investment Strategies: Changing investment strategies and allocating the budget properly can also affect the overall planning cost. First, the overall budget for DG sizing and installation is increased so that $\mathcal{C}^{DG}_{\max}$ corresponds to $P^{\max}_{DG} = 1500$ kW for the same set of DGs and their potential locations. Secondly, 3 additional DG locations (47, 27, and 114) are identified as potential DG placement locations. Fig. 12(a) and Fig. 12(b) show the distribution of prioritized loss of load when different DG planning measures are taken for risk-neutral and risk-averse cases, respectively. It is interesting to notice that increasing the budget to increase the capacity of DGs has a limited effect on the $CVaR_\alpha$ minimization. However, the expected value of prioritized load loss decreases to 2222.43 kW from 2467.46 kW. The conclusion is consistent for the risk-averse

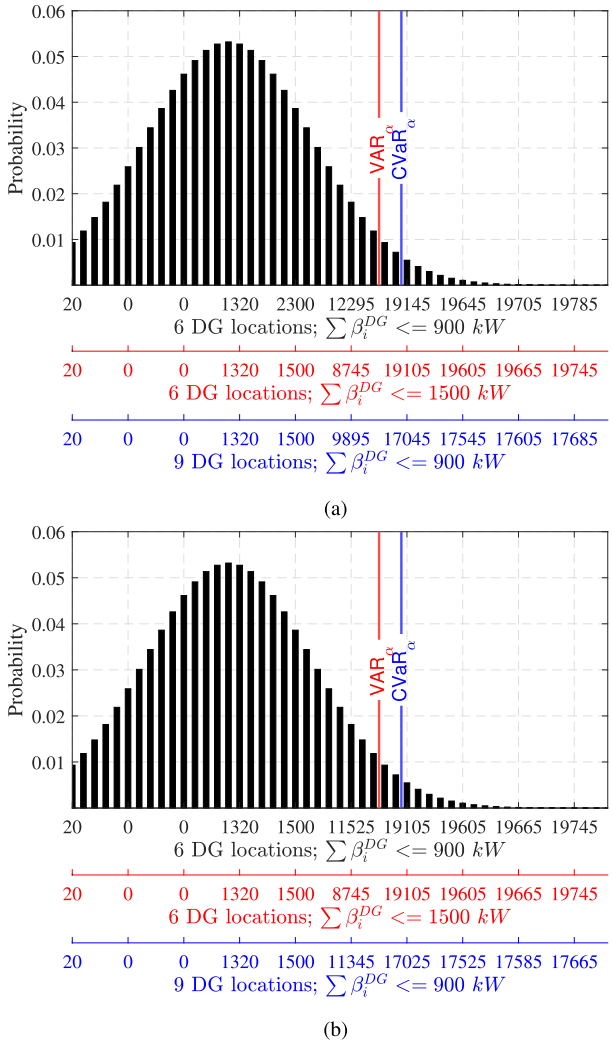


Fig. 12. $CVaR_\alpha$ for different DG investment strategies for (a) risk-neutral and (b) risk-averse planning. The value on the x-axis represents the prioritized loss of load for each ξ with corresponding p_ξ represented on the y-axis.

case. However, increasing the number of potential DG locations led to significant improvement in $CVaR_\alpha$ minimization. The change in expected loss is, however, insignificant. For the case with 9 potential DG locations, the value of $CVaR_\alpha$ decreases from 18415.04 kW to 16385.51 kW, for the risk-neutral case, and from 18119.1 kW to 15811.59 kW, for the risk-averse case. Thus, with a limited budget, multiple DG sites with smaller DGs are more effective in improving resilience.

3) Change in Number of Scenarios and Set of Scenarios: Fig. 13 shows five case studies simulated to evaluate the impacts of the number of scenarios (used in optimization) on solution quality and solve time; (a) 7 scenarios, (b) 21 scenarios, (c) 49 scenarios, (d) 98 scenarios, (e) 147 scenarios. Hence, for each case, different scenario sets are obtained using the method discussed in Algorithm 1. Fig. 14 shows the objective function value for the different number of scenarios (used in the optimization problem) along with the corresponding solve times. The result for each case is obtained by taking an average of 10

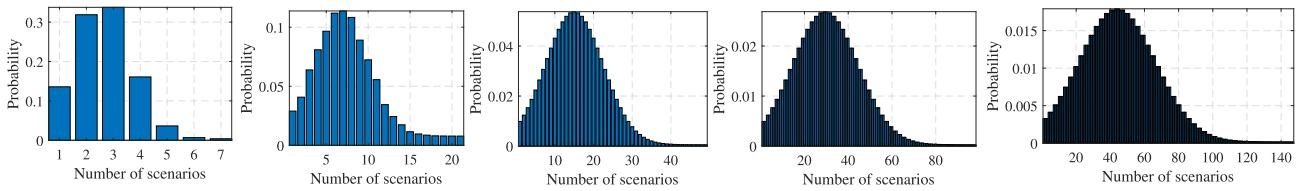


Fig. 13. Different number of scenarios with respective probabilities of occurrence: (a) 7 scenarios, (b) 21 scenarios; (c) 49 scenarios, (d) 98 scenarios e) 147 scenarios.

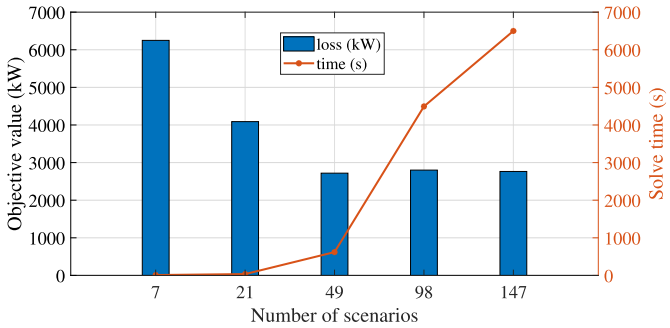


Fig. 14. Comparison of objective value and solve time for a different number of scenarios.

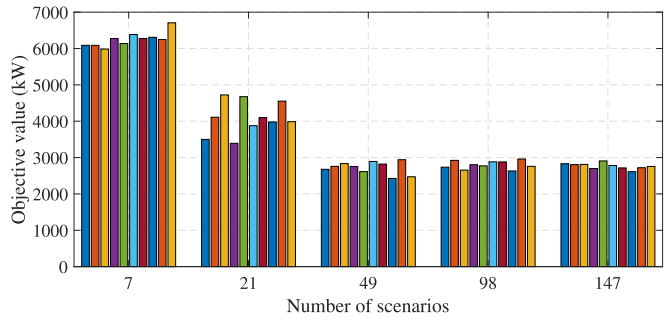


Fig. 15. Comparison of objective value on a different set of scenarios for each number of scenarios sampled.

representative scenario sets closest to the average representative scenario. We can clearly observe the trade-off between the number of scenarios, solution quality, and solve time. When a higher number of scenarios are used in optimization, the solution quality improves; however, it also leads to a significant increase in the solve time. It is also interesting to note that the solution obtained for 49 scenarios (2719.17 kW) is very close to the one obtained for 147 scenarios (2764.16 kW). However, the solve time for the problem with 147 scenarios is 11 times greater than that with 49 scenarios. Hence, 49 scenarios work well from the point of view of solution as well as solve time as the additional number of scenarios increases the computational complexity with no significant improvement in the objective value.

Fig. 15 shows our simulation results for 10 unique sets of scenarios sampled from the wind profile PDF. The simulations are done for five cases by including a different number of scenarios in the optimization problem, i.e., 7, 21, 48. 98, and

TABLE IV
95% CONFIDENCE INTERVAL OF THE SOLUTION OBTAINED USING DIFFERENT SCENARIO SETS FOR EACH NUMBER OF SCENARIOS

Number of scenarios	95% CI (kW)	Average objective value (kW)
7	[6105.95, 6393.99]	6249.57
21	[3764.35, 4414.53]	4089.44
49	[2596.13, 2842.21]	2719.17
98	[2721.20, 2879.54]	2800.37
147	[2705.10, 2823.21]	2764.16

147 (scenarios). For a specific case, it can be observed that the objective function values are very close for all 10 unique scenario sets. Each scenario set is selected based on the closest average Monte-Carlo loss as detailed in Algorithm 1. Hence, the differences in solutions for different scenario sets are not significant. The process is similar to the multiple replication method (MRP) as discussed in [18]. In addition, for each case, Table IV shows the 95% confidence interval for the objective function value along with the average objective value. As expected, including a larger number of scenarios in the optimization problem increases the granularity of information regarding the event and its impacts. As can be seen, the objective function values are consistent with the number of scenarios being considered in the optimization problem. It is interesting to note that the solution quality improves drastically when additional scenarios are considered in the optimization problem. Furthermore, the variation in the optimal function values also reduces as we consider a larger number of scenarios, see Table IV. For example, for the simulation case with 147 scenarios, the lower and upper limits of the optimal function values for all 10 unique scenario sets are very close to each other than any other scenarios. However, as shown by the results in Fig. 14, the solve time increases drastically with a higher number of scenarios. Thus, as expected, there is a tradeoff between computational complexity and solution quality.

V. CONCLUSION

This paper presents a risk-based planning framework for active power distribution systems to improve their resilience to extreme weather events. Resilience is characterized in a probabilistic sense to quantify the impacts of the HILP events on the system outages. A two-stage stochastic optimization problem is formulated to minimize the risk of system outages as a weighted function of the expected value and CVaR of the probabilistic system outages. The planning decisions include system upgrades

by siting and sizing DGs capable of intentional islanding to support CLs. A scenario reduction method is proposed based on realized loss functions that generate representative scenarios for HILP events for computational tractability. The proposed formulation makes it conducive to evaluate the trade-offs between risk-neutral and risk-averse planning decisions. The proposed risk-based planning framework is analyzed for different scenarios. It was observed that the DG-based restoration method with additional hardening measures could effectively minimize both the expectation and $CVaR_\alpha$ of the prioritized load loss. Furthermore, risk-averse planning measures were highly effective in restoring CLs during HILP events which is generally not the case for risk-neutral policy. It was also observed that it is preferable to site variable-sized DGs at multiple locations rather than a few large DGs under a limited budget. Based on the observations, the proposed risk-based framework provides distribution planners with the much-needed ability to evaluate alternate planning measures for resilience.

REFERENCES

[1] J. Furman, "Economic benefits of increasing grid resilience to weather outages," U.S. Dept. Energy, Tech. Rep., Aug. 2013. [Online]. Available: https://www.energy.gov/sites/default/files/2013/08/f2/Grid%20Resiliency%20Report_FINAL.pdf

[2] *NOAA National Centers for Environmental Information (NCEI) U. S. Billion-Dollar Weather and Climate Disasters*, 2022, doi: 10.25921/stkw-7w73. [Online]. Available: https://www.ncdc.noaa.gov/billions/

[3] R. Moreno et al., "From reliability to resilience: Planning the grid against the extremes," *IEEE Power Energy Mag.*, vol. 18, no. 4, pp. 41–53, Jul.–Aug. 2020.

[4] Q. Shi et al., "Enhancing distribution system resilience against extreme weather events: Concept review, algorithm summary, and future vision," *Int. J. Elect. Power Energy Syst.*, vol. 138, 2022, Art. no. 107860.

[5] G. Liu et al., "Resilient distribution system leveraging distributed generation and microgrids: A review," *IET Energy Syst. Integration*, vol. 2, no. 4, pp. 289–304, 2020.

[6] A. Arif et al., "Optimizing service restoration in distribution systems with uncertain repair time and demand," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6828–6838, Nov. 2018.

[7] W. Yuan et al., "Robust optimization-based resilient distribution network planning against natural disasters," *IEEE Trans. Smart Grid*, vol. 7, no. 6, pp. 2817–2826, Nov. 2016.

[8] B. A. J. Conejo et al., "Operations and long-term expansion planning of natural-gas and power systems: A market perspective," *Proc. IEEE Proc. IRE*, vol. 108, no. 9, pp. 1541–1557, Sep. 2020.

[9] A. Stankovic and K. Tomsovic, "The definition and quantification of resilience," *IEEE Power Energy Soc.*, Tech. Rep. PES-TR65, Apr. 2018.

[10] H. Gao et al., "Resilience-oriented pre-hurricane resource allocation in distribution systems considering electric buses," *Proc. IEEE Proc. IRE*, vol. 105, no. 7, pp. 1214–1233, Jul. 2017.

[11] Q. Zhang et al., "Stochastic pre-event preparation for enhancing resilience of distribution systems," *Renewable Sustain. Energy Rev.*, vol. 152, 2021, Art. no. 111636.

[12] S. Ma et al., "Resilience enhancement of distribution grids against extreme weather events," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 4842–4853, Sep. 2018.

[13] X. Wang et al., "Resilience enhancement strategies for power distribution network coupled with urban transportation system," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4068–4079, Jul. 2019.

[14] S. Ma et al., "Resilience-oriented design of distribution systems," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 2880–2891, Jul. 2019.

[15] Y. Xu et al., "Placement of remote-controlled switches to enhance distribution system restoration capability," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1139–1150, Mar. 2016.

[16] J. Liu, Y. Yu, and C. Qin, "Unified two-stage reconfiguration method for resilience enhancement of distribution systems," *IET Gener., Transmiss. Distrib.*, vol. 13, no. 9, pp. 1734–1745, 2019.

[17] Q. Shi et al., "Resilience-oriented DG siting and sizing considering stochastic scenario reduction," *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 3715–3727, Jul. 2021.

[18] Q. Zhang et al., "Stochastic pre-event preparation for enhancing resilience of distribution systems," *Renewable Sustain. Energy Rev.*, vol. 152, Dec. 2021, Art. no. 111636.

[19] B. Taheri et al., "Distribution system resilience enhancement via mobile emergency generators," *IEEE Trans. Power Del.*, vol. 36, no. 4, pp. 2308–2319, Aug. 2021.

[20] P. W. Glynn and D. L. Iglehart, "Importance sampling for stochastic simulations," *Manage. Sci.*, vol. 35, no. 11, pp. 1367–1392, 1989.

[21] R. Rush et al., "Stratified filtered sampling in stochastic optimization," *J. Appl. Math. Decis. Sci.*, vol. 4, no. 1, pp. 17–38, 2000.

[22] D. Urgun, C. Singh, and V. Vittal, "Importance sampling using multilabel radial basis classification for composite power system reliability evaluation," *IEEE Syst. J.*, vol. 14, no. 2, pp. 2791–2800, Jun. 2020.

[23] R. Preece and J. V. Milanović, "Efficient estimation of the probability of small-disturbance instability of large uncertain power systems," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1063–1072, Mar. 2016.

[24] F. D. Munoz, A. H. van der Weijde, B. F. Hobbs, and J.-P. Watson, "Does risk aversion affect transmission and generation planning? A Western North America case study," *Energy Econ.*, vol. 64, pp. 213–225, May 2017.

[25] S. Poudel, A. Dubey, and A. Bose, "Risk-based probabilistic quantification of power distribution system operational resilience," *IEEE Syst. J.*, vol. 14, no. 3, pp. 3506–3517, Sep. 2020.

[26] A. Poudyal, A. Dubey, and S. Poudel, "A risk-driven probabilistic approach to quantify resilience in power distribution systems," in *Proc. 17th Int. Conf. Probabilistic Methods Appl. Power Syst.*, 2022, pp. 1–6.

[27] R. Khodabakhsh and S. Sirouspour, "Optimal control of energy storage in a microgrid by minimizing conditional value-at-risk," *IEEE Trans. Sustain. Energy*, vol. 7, no. 3, pp. 1264–1273, Jul. 2016.

[28] T. Jiang et al., "Research on risk-based control strategy for power grid operation," in *Proc. IEEE Sustain. Power Energy Conf.*, 2021, pp. 2815–2820.

[29] X. Yang et al., "Risk-averse coordinated economic dispatching and voltage regulation in ADNS with on-site renewables and soft open points," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2021, pp. 1–5.

[30] J. Wu and P. Wang, "Risk-averse optimization for resilience enhancement of complex engineering systems under uncertainties," *Rel. Eng. Syst. Saf.*, vol. 215, Nov. 2021, Art. no. 107836.

[31] S. Poudel, A. Dubey, and K. P. Schneider, "A generalized framework for service restoration in a resilient power distribution system," *IEEE Syst. J.*, vol. 16, no. 1, pp. 252–263, Mar. 2022.

[32] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lectures on Stochastic Programming: Modeling and Theory*. Philadelphia, PA, USA: SIAM, 2021.

[33] I. B. Sperstad, G. Kjølle, and E. Ø. Norum, "Accounting for uncertainties due to high-impact low-probability events in power system development," *Electric Power Syst. Res.*, vol. 193, 2021, Art. no. 107015.

[34] D. P. Nedic et al., "Criticality in a cascading failure blackout model," *Int. J. Elect. Power Energy Syst.*, vol. 28, no. 9, pp. 627–633, 2006.

[35] R. N. Allan et al., *Reliability Evaluation of Power Systems*. Berlin, Germany: Springer, 2013.

[36] A. Kwasinski, "Quantitative model and metrics of electrical grids' resilience evaluated at a power distribution level," *Energies*, vol. 9, no. 2, 2016, Art. no. 93.

[37] R. T. Rockafellar et al., "Optimization of conditional value-at-risk," *J. Risk*, vol. 2, pp. 21–42, 2000.

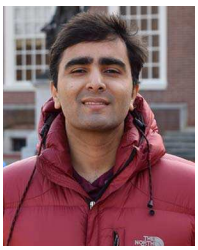
[38] N. Noyan, "Risk-averse two-stage stochastic programming with an application to disaster management," *Comput. Operations Res.*, vol. 39, no. 3, pp. 541–559, 2012.

[39] A. Poudyal et al., "Spatiotemporal impact assessment of hurricanes on electric power systems," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2022, pp. 1–5.

[40] M. Panteli et al., "Power system resilience to extreme weather: Fragility modeling, probabilistic impact assessment, and adaptation measures," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3747–3757, Sep. 2017.

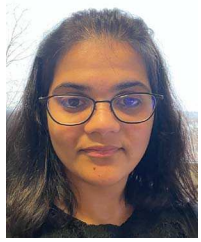
[41] M. D. Powell, S. H. Houston, and I. Ares, "Real-Time Damage Assessment in Hurricanes," in *Proc. 21st Conf. Hurricanes Trop. Meteorol.*, 1995, pp. 500–502.

[42] M. Panteli and P. Mancarella, "Modeling and evaluating the resilience of critical electrical power infrastructure to extreme weather events," *IEEE Syst. J.*, vol. 11, no. 3, pp. 1733–1742, Sep. 2017.
[43] M. Panteli et al., "Metrics and quantification of operational and infrastructure resilience in power systems," *IEEE Trans. Power Syst.*, vol. 32, no. 6, pp. 4732–4742, Nov. 2017.
[44] V. L. Parsons, "Stratified sampling," *Wiley StatsRef: Statist. Reference Online*, pp. 1–11, 2017.
[45] J. Ekblom and J. Blomvall, "Importance sampling in stochastic optimization: An application to intertemporal portfolio choice," *Eur. J. Oper. Res.*, vol. 285, no. 1, pp. 106–119, 2020.
[46] H. Heitsch and W. Römisch, "A note on scenario reduction for two-stage stochastic programs," *Operations Res. Lett.*, vol. 35, no. 6, pp. 731–738, 2007.
[47] L. C. Coelho, "Linearization of the product of two variables," *Canada Res. Chair Integr. Logistics*, 2013.
[48] L. Gan and S. H. Low, "Convex relaxations and linear approximation for optimal power flow in multiphase radial networks," in *Proc. IEEE Power Syst. Comput. Conf.*, 2014, pp. 1–9.
[49] J.-P. Watson, D. L. Woodruff, and W. E. Hart, "PYSP: Modeling and solving stochastic programs in python," *Math. Program. Comput.*, vol. 4, no. 2, pp. 109–149, 2012.



Abodh Poudyal (Graduate Student Member, IEEE) received the B.E. degree from the Department of Electrical Engineering, Pulchowk Campus, Kathmandu, Nepal, in 2016, and the M.S. degree from the Electrical Engineering and Computer Science Department, South Dakota State University, Brookings, SD, USA, in 2018. He is currently working toward the Ph.D. degree in electrical engineering with Washington State University, Pullman, WA, USA. His research interests include power distribution modeling and optimization, resilience assessment, modeling extreme events and their impact on the power grid, and AI-based applications on the power grid.

Shiva Poudel (Member, IEEE) received the B.E. degree from the Department of Electrical Engineering, Pulchowk Campus, Kathmandu, Nepal, in 2013, the M.S. degree from the Electrical Engineering and Computer Science Department, South Dakota State University, Brookings, SD, USA, in 2016, and the Ph.D. degree in electrical engineering from Washington State University, Pullman, WA, USA, in 2020. He is currently a power systems Research Engineer with Pacific Northwest National Laboratory. His research interests include distribution system modeling and analysis, resilience assessment, and transactive energy systems.



Anamika Dubey (Senior Member, IEEE) received the Ph.D. degree in electrical and computer engineering from the University of Texas at Austin, Austin, TX, USA, in 2015. She is currently Huie-Rogers Endowed Chair Associate Professor of electrical engineering with the School of EECS, Washington State University, Pullman, WA, USA. She also holds a joint appointment as a Research Scientist with the Pacific Northwest National Laboratory (PNNL). Her research interests include optimization and control of large-scale electric power distribution systems for improved efficiency, flexibility, and resilience. Her expertise is in modeling, analyzing, and operating active power distribution systems with massive penetrations of controllable grid-edge resources including DERs, EVs, and GEBs. She was the recipient of the National Science Foundation (NSF) CAREER Award. She is an Associate Editor for the IEEE TRANSACTIONS ON POWER SYSTEMS, IEEE POWER ENGINEERING LETTERS, and IEEE ACCESS. She is the current secretary of IEEE PES Distribution Systems Analysis Subcommittee and IEEE PES University Education Subcommittee and is the PES Chapter Chair of the IEEE Palouse Section.