Investment Planning Framework for Mitigating Cascading Failures

Balaji V. Venkatasubramanian[1,2], Sina Hashemi[1], Rodrigo Moreno[3], Pierluigi Mancarella[4,5], Mathaios Panteli[1]
[1]Department of Electrical and Computer Engineering, University of Cyprus, Cyprus.
[2]School of Technology, Woxsen University, Telangana, India.
[3]Departamento de Ingeniería Eléctrica, Universidad de Chile, Santiago, Chile.
[4]School of Electrical and Electronics Engineering, University of Melbourne, Parkville, VIC, Australia.
[5]School of Electrical and Electronic Engineering, The University of Manchester, Sackville Street, Manchester M13 9PL, UK
balajivenkateswaran.v@gmail.com; {hashemi.seyedsina, panteli.mathaios}@ucy.ac.cy; rmorenovieyra@uchile.cl; pierluigi.mancarella@unimelb.edu.au;

***Abstract*—Critical component outages can lead to widespread cascading propagation, which is however typically ignored in existing investment planning approaches. To address this gap, this paper seamlessly integrates advanced cascading failure analysis into resilient investment planning. It first deploys a stochastic simulator to generate spatiotemporal high-impact low-probability (HILP) events, which are then assessed using a cascading failure model, generating various cascading quantification metrics (CQMs). The framework explicitly quantifies tail risks (i.e., HILP events) using Conditional Value-at-Risk (CVaR) with a confidence level determined by unsupervised clustering, instead of using a predetermined confidence level. This enables the more tailored identification of a set of worst-case scenarios for the system under investigation, improving its practicality. An optimization model then utilizes the outputs of the cascading analysis and the defined CVaR confidence level to identify investment portfolios that provide a hedge against cascading failures. The proposed work is demonstrated on the IEEE 39-bus system, revealing reduced cascading propagation.**

***Index Terms*—Cascading failures, Power system resilience, Power system planning, Risk assessment, Unsupervised clustering.**

NOMENCLATURE

Sets and Indices:

ς, l	Indices for hazard scenario and lines
t, t_i	Indices for time and time at which the first wind-induced line outage occurs
γ	Indices representing network enhancement
N_{RS}	Number of reduced scenarios

Parameters:

$w_{critical}$	Critical wind gust speed
w_t	Wind gust speed at time t
$w_{t_i}^{\varsigma,l}$	Wind speed at line l at time t_i for scenario ς
C_{OHL}	Cost of hardening OHL in €/km
C_{ESS}	Cost of installing ESS in €/kW
ll_{OHL}	Length of OHL line l in km
P_{ESS}^{max}	Maximum power obtained from an ESS
α	CVaR confidence level
$Budget_{max}$	Maximum budget allocated in Euros.

Variables:

b_{OHL}^{l}	Binary decision variable representing hardening of OHL at line l
b_{ESS}^{n}	Binary decision variable representing the installation of ESS at node n
$ls_{\varsigma,l}$	Line status of line l for scenario ς
ν	Value-at-risk (VaR) at $\alpha \in [0,1)$ of $f_{CFM}(\gamma, \varsigma)$
ϱ_{ς}	Excess variable that ensures that CVaR_{α} is calculated only for values beyond VaR_{α} in ς scenarios

I. INTRODUCTION

Climate change has increased the frequency, intensity, and duration of severe weather events, impacting critical infrastructures, including power systems. For instance, Hurricane Ian in 2022 in the USA resulted in widespread power outages affecting over 2.7 million customers and costing an estimated $113 billion in damages [1]. The outage of critical components due to high-impact low-probability (HILP) events, such as natural hazards and extreme weather, has the potential to trigger protection mechanisms, which in turn could lead to a series of cascading failures throughout the power system [2]. Several researchers have proposed various approaches to enhance power system resilience through different planning strategies. In [3], a resilience-centered approach is proposed for assessing the benefits of different investment options with an application to earthquakes, distinguishing the fundamental differences between reliability- and resilience-driven enhancements. A few researchers have proposed resilient planning frameworks that integrate the fragility-based event models with stochastic optimization, identifying the location or size of optimal resources such as energy storage systems (ESS), distributed generators (DGs), hardening of overhead lines (OHLs), etc., to deal with earthquakes and windstorms [4], [5]. Similarly, several methodologies are proposed in the literature that follow either a two-stage or multi-stage approach to enhance the system resilience by optimally allocating resources such as energy storage, DG,

Submitted to the 23rd Power Systems Computation Conference (PSCC 2024).

and line hardening, [6]-[9]. In [10], the authors have introduced a framework for enhancing resilience through DG allocation, where resilience is assessed using expected energy not served (EENS) and conditional value-at-risk (CVaR). Many of these approaches assess system resilience by employing metrics such as ENS, EENS, or CVaR at a predefined confidence level.

Although several researchers have proposed various approaches to enhance power system resilience, very few have considered the impact of cascading outages in their decision-making processes. For example, in [11], a transmission expansion planning approach is proposed to improve power system resilience by considering cascading outages generated using the 'N-1' security criterion. However, the outage scenarios generated using traditional security criterion 'N-1' may not reflect the stochastic characteristics of widespread cascading outages initiated by natural hazards or extreme weather, leading to unrealistic solutions. In [12], a framework is introduced to identify the vulnerable lines and assess a resilience index against cascading failures triggered by windstorm-induced outages. However, this methodology does not include mitigation strategies that can further enhance the system's resilience. Furthermore, [13] presents a cost-based optimization model aimed at enhancing system resilience against cascading failures by reducing demand not served (DNS). Nevertheless, the formulation of contingencies within this model is randomly generated and does not capture the spatial and temporal characteristics of specific types of natural hazards.

From the literature, it is apparent that most methodologies neglect to consider the cascading effect of initiating contingencies in the decision-making on investment portfolios. Furthermore, they quantify resilience primarily based on more reliability-oriented metrics like EENS (i.e., averaged across all simulated scenarios), which may not effectively capture the unique characteristics of HILP events observed in the tail end of stochastic scenario distributions. Therefore, it becomes vital to quantify and reinforce resilience using cascading simulators that can output and seamlessly integrate CVaR-based metrics in the identification of resilience investment portfolios. The confidence level of CVaR, denoted as α, varies in the range [0, 1], where zero represents a risk neutral approach, and α=1 a risk-averse approach; it is often selected as 0.95 or 0.99 [14]. However, in the context of tail risk optimization utilizing CVaR, there is notable interest from the finance and business sector (where mainly the CVaR-driven optimization was originated) in investigating and varying the confidence level for finding its suitable value as it has a substantial influence on the optimization process [15]. This contrasts with existing CVaR-driven optimization approaches in power systems where the confidence level α is considered fixed and pre-determined. However, determining the value of α proves to be a challenging task, particularly when dealing with distributions exhibiting heavy tails [16]. In a hypothetical scenario, a loss function related to DNS within a network is characterized by a value-at-risk (VaR) of 1000 MW and a CVaR of 1100 MW at α=0.95. The objective is to minimize the CVaR of this loss function. During the stochastic optimization process, various scenarios ranging from 1000 MW up to the worst-case scenario (α=1) are investigated for this purpose. There is hence a potential for an event with 950 MW demand loss (or very close, but smaller than VaR=1000 MW) to be overlooked during the optimization process. Therefore, it becomes imperative to identify a suitable CVaR confidence level to address this issue effectively and guarantee the consideration of tail events of high impact in the stochastic CVaR optimization process.

This paper presents a novel investment planning framework for enhancing resilience against cascading failures from HILP events. Initially, a stochastic event simulator is applied to generate a wide range of hazard scenarios. In line with maintaining generality, this paper addresses scenarios involving critical outages resulting from natural hazards. Subsequently, a cascading failure model proposed in [17] assesses the impacts of the event scenarios on the network and quantifies cascading quantification metrics (CQM). Given the potential computational challenges stemming from the extensive scenario set, a scenario reduction technique employing random sampling is implemented. This aims to streamline scenarios while retaining key attributes. The study targets cascading failures triggered by over/under frequency and line overload protection mechanisms, crucial for CQM establishment. The framework optimizes CVaR linked with the selected CQM, adhering to a confidence level determined through an unsupervised clustering algorithm. This algorithm encapsulates a set of worst-case scenarios, yielding optimal investment portfolios. Several case studies were conducted on the IEEE 39-bus system to demonstrate the effectiveness of the proposed framework.

The main contributions of this paper include:

- Coupling a spatial and temporal simulator of extreme weather events with cascading failure analysis for the quantification of their impact using various cascading metrics.
- Seamless integration of the cascading failure analysis into investment planning for the identification of resilience enhancement portfolios against cascading events initiated by natural hazards and extreme weather.
- Application of a machine learning (ML) algorithm, specifically an unsupervised clustering algorithm, to determine the confidence level of CVaR, providing tailored risk assessment to the system operators and planners.

The remainder of this paper is organized as follows: Section II provides an elaboration of the proposed framework for investment planning, addressing the mitigation of cascading failures caused by natural hazards. Following that, Section III demonstrates the outcomes from several case studies undertaken and offers corresponding discussions. Lastly, Section IV presents the conclusions derived from this study.

II. Proposed Framework

The proposed framework, as illustrated in Fig. 1, seamlessly integrates advanced spatiotemporal event and cascading simulators with ML-driven risk-informed investment decision-making optimizer. In this framework, a fragility-based event simulator initially generates spatiotemporal hazard scenarios for a defined time horizon. Subsequently, the initiating events are derived from these scenarios, reflecting the time at which the first hazard-induced cascading outage occurs in the network. Furthermore, the network is subjected to these initiating events to perform cascading failure analysis, resulting in the

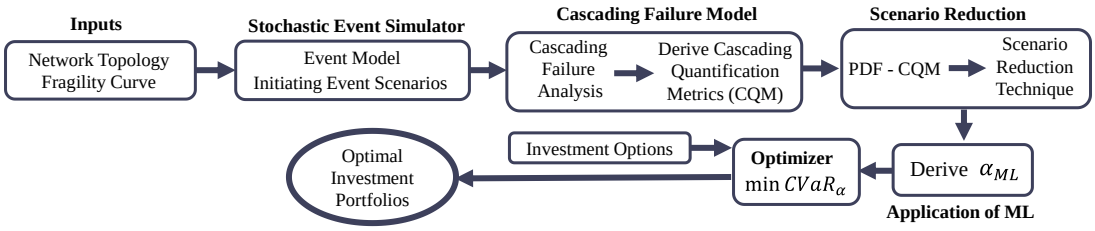

Figure 1. Overview of the Proposed Framework.

derivation of CQMs. A key component within CQM comprises metrics related to DNS resulting from frequency and line active power limit breaches, collectively referred to as 'total DNS' (TDNS) when summed together. Additionally, CQM includes metrics that reflect topological impacts, encompassing quantities such as the number of active islands (N_{AI}), number of passive islands (N_{PI}), and number of tripped assets (N_{TA}). With these CQMs, an unsupervised clustering algorithm determines α_{ML} to group scenarios with similar impact. Then, constrained optimization minimizes TDNS's CVaR based on α, identifying optimal asset portfolios. This paper focuses on windstorms to reflect extreme events; however, the proposed framework has the potential to expand to different types of extreme weather events.

A. *Stochastic Event Simulator and Fragility-driven System Impact Assessment*

The input data for this event simulator includes the network topology and fragility curves, which depict the relationship between the hazard intensity and the probability of damage to network assets. In the case of transmission lines, wind fragility curves are deployed, as explained in [5], which ultimately provide the time- and wind-dependent outage probabilities of the assets exposed to the generated wind scenarios. To simulate realistic windstorm events, this model derives windstorm characteristics from a historical database [18], specifically the levels of wind gust speed, windstorm radius, and windstorm directions. The spatial modeling of windstorms is performed by randomly selecting the starting coordinates of the windstorm within the region of study and simulating its movement via spatiotemporal modeling based on parameters such as windstorm direction, storm duration, and windstorm radius. Further, a large set of windstorm scenarios is generated to encompass both known (historical) and unknown, potential future windstorm patterns within a defined time horizon. For the sake of simplicity, this paper adopts a 24-hour time horizon. Furthermore, the spatial intersection between the windstorm trajectories and the transmission lines is analyzed to determine the status of affected lines based on their wind-dependent failure probability derived by the wind fragility curves. In this paper, the decision regarding tripped lines within this overlap is made using (1). This equation demonstrates that if the wind-dependent failure probability, denoted as $P_L(w_t)$, surpasses a randomly generated number r between 0 and 1, the line will be tripped. Conversely, if $P_L(w_t)$ is less than r, the line status remains operational (status 1).

$$L_S(w_t, l) = \begin{cases} 1 & \text{if } P_L(w_t) < r \\ 0 & \text{if } P_L(w_t) > r \end{cases} \quad (1)$$

where, w_t is the wind speed at step-time t. After generating the time- and wind-dependent status of the network lines within the defined time horizon, the time at which the first wind-induced line outage occurs, denoted as t_i, is determined. The line statuses corresponding to this t_i are designated as the initiating event, i.e., $L_S(w_{t_i}, l)$. These line statuses, for all the scenarios, represent the set of stochastic initiating events.

B. *Cascading Failure Model*

This paper employs the cascading failure model from [17] for tailored cascading failure analysis in power network resilience assessment. In this paper, the analysis utilizes DC power flow, referred to as DC-CFM, and is implemented using the MATPOWER toolkit [19] within MATLAB. As aforementioned, the protective mechanisms considered in this research for cascading analysis encompass over/under frequency and line overload. The implementations of these mechanisms are as follows:

1) Over/Under Frequency Protection: With the presence of synchronous generators in the network, sudden discrepancies or imbalances between electrical and mechanical power can result in changes in rotor frequency. Consequently, protective schemes such as over-frequency load shedding (OFLS) or under-frequency load shedding (UFLS) are triggered to restore balance and maintain the frequency within acceptable limits. In the event of a power mismatch within the allowable change (±) in power generation, the OFLS/UFLS follows redistributing slack generation across generation proportionally, ensuring that each generator operates within its capacity. If the generation increases (+) by an allowable mismatch level but does not exceed the available capacity, UFLS activates the load reduction. On the other hand, if the generation decreases (−) within the allowable mismatch level, the OFLS activates to trip the generation. For a more comprehensive understanding, refer to Algorithms 1 and 2 outlined in [17].

2) Line Overload Protection: In the case of line overload, the overload protection (OLP) initiates the tripping of lines that surpass their predefined ratings. This proactive measure prevents lines from sustaining damage due to overheating, thereby ensuring safe network operation once the issues are resolved.

Each initiating event from the stochastic event simulator (those with $L_S(w_{t_i}, l)$) can trigger the operation of these protection

measures. These can then create cascading isolated network segments, potentially causing DNS, quantified with CQM.

C. Scenario Reduction

The CQM is calculated for all initiating events within a large scenario set (LS). It is crucial to note that utilizing all these scenarios in the optimization process could lead to a computational slowdown. Therefore, this paper employs a scenario reduction technique based on random sampling method like [10], as shown in Fig. 2. While metrics other than TDNS provide insights into initiating event consequences, it is essential to recognize that the primary impact is predominantly influenced by the demand-based metric, TDNS derived by the cascading simulator. For example, a non-zero N_{TA} with zero TDNS indicates no cascading propagation, while a non-zero TDNS corresponds to cascading propagation, reflected in active/passive islands and non-zero N_{TA}. Due to its significance, TDNS is evaluated for all the original scenarios and used in the scenario reduction process. During scenario reduction, an initial desired number of reduced scenarios (e.g., $r_i = 50$) is established. Subsequently, $r = r_i$ scenarios are randomly selected from LS to form a reduced scenario set (RS). Comparison between the mean and standard deviation of the original large scenario set (LS_μ, LS_σ) and the randomly sampled reduced set (RS_μ, RS_σ) is performed. If their means' and standard deviations' percentage error is within a predefined tolerance ($tol = 0.1\%$), the reduced set is acceptable. Otherwise, another set of r reduced scenarios are sampled with $r = r + r_i$. This process continues until the error is within tolerance. This iterative approach ensures that reduced scenarios retain their original hazard scenario characteristics.

D. Application of ML to derive 'α'

In the application of unsupervised clustering to determine α representing scenarios with similar impacts, all four CQMs are employed. TDNS addresses demand impact, while N_{AI}, N_{PI}, and N_{TA} represent topological impact.

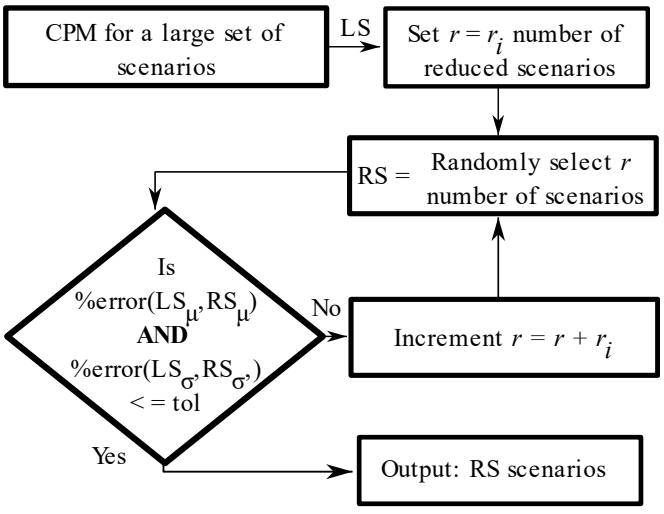

Figure 2. Flowchart of Proposed Scenario Reduction Technique.

These metrics enrich the ML analysis, providing a comprehensive view of initiating event scenarios covering both demand and network topological impact. Considering the substantial scale of data involved and the advantage of not needing a predetermined cluster count, this study employs a hierarchical clustering algorithm [20]. The actual count of scenarios constituting the identified set of worst-case scenarios is a guiding factor in determining the value of α. In this paper, the cophenetic correlation coefficient (c) is applied to measure the quality of hierarchical clustering, which assesses how well the clustering preserves the original pairwise distances between data points. A c value close to 1 indicates a high correlation, signifying meaningful clustering [20]. The study involves experimenting with various linkage methods, including 'single', 'ward,' 'average,' and 'complete,' in combination with the Euclidean distance metric. Subsequently, the value of c is used to evaluate the quality of these clustering approaches. Following this procedure with the CQMs derived for all the scenarios, a set of similar impact scenarios with a count ε is derived. Now, the value of $\alpha = \alpha_{ML}$ is calculated using ε and the number of reduced scenarios (N_{RS}) as shown in (2).

$$\alpha_{ML} = \left(1 - \frac{\varepsilon}{N_{RS}}\right) \tag{2}$$

The α_{ML} plays a pivotal role in steering the risk aversion level of the system planner, and hence the optimal asset portfolios obtained from the optimizer. This systematic and comprehensive approach ensures the coverage of a wide array of potential scenarios to reveal the most critical scenarios of similar behavior lying within the reduced set.

E. The Optimizer

The mathematical formulation of the optimizer, encompassing the objective function and its associated constraints, is detailed in this section. The primary goal is to mitigate the demand curtailment caused by cascading failure propagation. To achieve this, the TDNS obtained from the cascading simulator is utilized in the objective function, rather than considering all CQMs. The objective function formulated will identify the optimal assets that lead to a reduction in the CVaR of TDNS as shown in (3).

$$\min CVaR_\alpha\left(f_{CFM}(\gamma, \varsigma)\right) \tag{3}$$

$$CVaR_\alpha\left(f_{CFM}(\gamma, \varsigma)\right) = \nu + \frac{1}{1-\alpha}\sum_{\varsigma \in r} \varrho_\varsigma \tag{4}$$

where, γ represents the decision variables linked to network enhancements for various assets namely OHLs and ESSs, ς denotes scenarios within the reduced scenario set RS, f_{CFM} represents the loss function, evaluated using DC-CFM, which quantifies the CQM across all scenarios ς with the prescribed network enhancement γ. In (4), ν refers to value-at-risk (VaR) at $\alpha \in [0,1)$ of $f_{CFM}(\gamma, \varsigma)$, ϱ_ς represents an excess variable that ensures that CVaR_α is calculated only for values beyond VaR_α in ς scenarios, and $1 - \alpha$ refers to the scenarios representing HILP events.

1) Constraints: The objective function is constrained with power flow equations that apply protection mechanisms as explained in Section II.B while evaluating CQM for the given

hazard scenario using DC-CFM. Moreover, the objective function is subject to asset and budget constraints as follows:

$$y^{\varsigma,l} = \begin{cases} 1 & if\ w_{t_i}^{\varsigma,l} > w_{critical} \\ 0 & Otherwise \end{cases} \quad (5)$$

$$ls_{\varsigma,l} = \begin{cases} (1 - y^{\varsigma,l})\ if\ b_{OHL}^{l} = 1 \\ L_S\left(w_{t_i}, l\right)\ Otherwise \end{cases} \quad (6)$$

$$\sum_l C_{OHL}.b_{OHL}^{l}.ll_{OHL} + \sum_n C_{ESS}.b_{ESS}^{n}.P_{ESS}^{max} \leq Budget_{max} \quad (7)$$

In the above equations, b_{OHL}^{l} and b_{ESS}^{n} are the binary decision variables that represents the hardening of OHL at line l and installation of ESS at node n, respectively when equal to 1. Notably, despite the hardening of OHL, the possibility of failure remains when these lines are exposed to wind speeds surpassing the critical threshold ($w_{critical}$). This requirement is ensured through (5) and (6), where $ls_{\varsigma,l}$ signifies the line status of line l during scenario ς, $y^{\varsigma,l}$ is a binary variable used to implement the windspeed condition. The total cost of the investment through OHL hardening and the installation of ESS is calculated using the left-hand side of (7), which is constrained to be less/equal to the maximum allocated budget. Here, C_{OHL} and C_{ESS} represent the cost of OHL in €/km and the cost of ESS in €/kW, respectively. Additionally, ll_{OHL} denotes the length of the OHL line and P_{ESS}^{max} represents the maximum capacity of ESS. The updated line status, which considers the line hardening ($ls_{\varsigma,l}$), the location or node where the ESS is installed, and its maximum size (P_{ESS}^{max}), is provided as input to the DC-CFM function ($f_{CFM}(\gamma, \varsigma)$) for evaluating CQM.

III. Simulation Results and Discussion

The effectiveness of the proposed investment planning framework, incorporating α_{ML}, is evaluated using the IEEE 39-bus system with fictional geographical coordinates. For illustrative purposes, this paper applies hazard scenarios corresponding to windstorms, generated based on a range of parameters derived from historical data [18]. The investment options considered in this work involve hardening OHLs (i.e., making them more robust to the windstorm) and installing ESSs to explore the trade-offs between infrastructure options (OHL) and flexibility providers (ESS). The cost of OHL considered is 300 €/m, as derived from [21], while the maximum capacity of each ESS is assumed to be 100 MW, with a cost of 150 €/kW [22]. Firstly, the outcome of the scenario reduction process, representing the base case (i.e., without planning), is presented to highlight its effectiveness in capturing the characteristics of the original scenarios. Furthermore, the importance of cascading failures in the context of investment planning is emphasized through a comprehensive comparison between the performance of the proposed framework, the conventional approach, and the base case condition. This analysis serves to underscore the resilience of assets identified within both planning methodologies when exposed to cascading failures. For the initial illustration, a standard CVaR confidence level of α=0.95 is applied. Additionally, the influence of α_{ML} in decision-making and its performance against cascading failures is illustrated. To highlight its effectiveness, its performance is compared with the outcomes of α=0.95. For demonstration purposes, based on the cost of assets, €50 million is assumed for investments in both cases. To effectively illustrate these conditions, the following case studies are presented below:

- *Case A*: Define and compare the investment portfolios for CVaR with α=0.95 using the proposed approach considering the cascading outages as illustrated in (3), and without considering the cascading outages (i.e., the conventional planning approach). Evaluate the effectiveness of the defined portfolios from the conventional approach when exposed to cascading failures initiated by extreme events.
- *Case B*: Perform the same steps as Case A but with α_{ML}.

A. Results from Scenario Reduction: Base Case

A substantial set of 1000 windstorm scenarios are generated by employing the stochastic event simulator over a 24-hour time horizon. Each scenario contributes to deriving the initiating set of line outages ($L_S\left(w_{t_i}, l\right)$), as explained in Section II.A. Subsequently, these initiating events are used to assess TDNS through cascading failure model. Simultaneously, all tripped lines throughout the complete time horizon are determined, and their impact is measured in terms of DNS following the conventional approach. Probability distributions for the original scenarios leading to TDNS and DNS are illustrated in Fig. 3a and Fig. 3c, respectively. The scenario reduction process, as depicted in Fig. 2, is employed to identify reduced scenarios for both TDNS and DNS. Through this process, 200 reduced scenarios (r) are obtained for both TDNS and DNS, with the corresponding probability distributions presented in Fig. 3b and Fig. 3d, respectively. Notably, Fig. 3 visualizes the mean and standard deviation of all distributions, indicating a close similarity between the mean and standard deviation of both TDNS and DNS. It is worth mentioning that these reduced scenarios of TDNS and DNS are applied in the optimization and used to derive α_{ML} for both proposed and conventional approach.

B. Case A: Comparing investment porftolios and CQMs with and without considering cascading outages

In this case, optimization of (3) while considering cascading outages (i.e., the proposed approach) and optimization of CVaR while ignoring cascading outages (i.e., the conventional

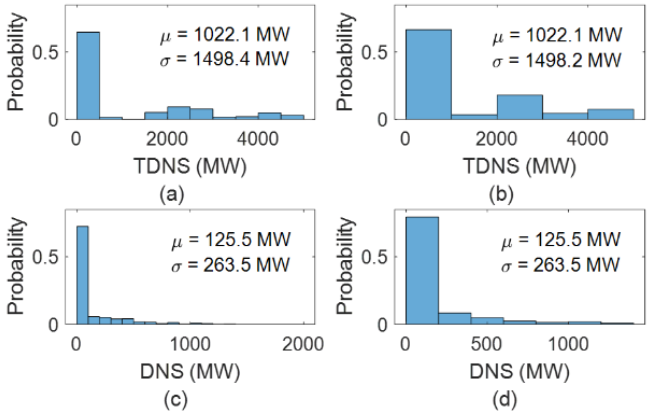

Figure 3. Probability Distribution of (a) LS of TDNS (b) RS of TDNS (c) LS of DNS (d) RS of DNS.

approach) with α=0.95 are performed. Then, the investment portfolios from the latter approach are exposed to the cascading simulator. The proposed approach resulted in two ESSs and six OHLs, while the conventional approach resulted in one ESS and eight OHLs. Both these portfolios are exposed to the cascading simulator, and the resulting CQMs are tabulated in Table I (including the Base Case without any reinforcement). The table shows that the optimal assets obtained from the proposed approach significantly improve all CQMs compared to the conventional approach from the base case. For instance, the conventional approach reduces the mean of TDNS ($\overline{TDNS}$) from 1022.1 MW to 945.7 MW with 7.5% reduction from base case. In contrast, the proposed approach reduces the $\overline{TDNS}$ from 1022.1 MW to 933.1 MW, signifying an 8.7% reduction from the base case.

Similarly, in the case of CVaR of TDNS, the conventional approach reduces it from 4570.9 MW to 4181.3 MW, signifying an 8.5% reduction from the base case and the proposed approach reduces it from 4570.9 MW to 4103.2 MW, signifying a 10.2% reduction from the base case. This shows that the proposed approach effectively reduces $\overline{TDNS}$ and CVaR of TDNS further by 1.3% and 1.9%, respectively. In the case of number of active islands, the average active islands ($\overline{N_{AI}}$) and the CVaR of active islands ($CVaR(N_{AI})$) is significantly improved from 2.1 to 2.2 and 5.0 to 5.3 by conventional approach and from 2.2 to 2.3 and 5.0 to 6.0 by proposed approach, respectively. Further, in the case of number of passive islands, the average passive islands ($\overline{N_{PI}}$) and the CVaR of active islands ($CVaR(N_{PI})$) is significantly reduced from 2.2 to 1.9 and 10.9 to 8.9 by conventional approach and from 2.2 to 1.6 and 10.9 to 7.7 by proposed approach, respectively. This demonstrates that the proposed approach effectively allocates optimal asset portfolios which reduces the blacked-out islands. Furthermore, in the case of number of tripped assets, the average tripped assets ($\overline{N_{TA}}$) and the CVaR of tripped assets ($CVaR(N_{TA})$) is significantly reduced from 8.6 to 7.7 and 38.7 to 33.8 by conventional approach and from 8.6 to 6.9 and 38.7 to 27.1 by proposed approach, respectively. From these findings, it is evident that the proposed approach effectively mitigates cascading propagation compared to the conventional approach.

C. Case B: Application of α_{ML} in investment decision-making

In this case, optimization of (3) while considering cascading outages (i.e., the proposed approach) and optimization of CVaR while ignoring cascading outages (i.e., the conventional approach) is performed with $\alpha = \alpha_{ML}$. In the proposed approach, the 200 scenarios corresponding to TDNS, as shown in Fig. 3b, are utilized to derive CQMs representing various scenarios. Employing these CQMs as input for the hierarchical clustering algorithm results in the identification of 23 worst-case scenarios, with α_{ML} = 0.885 derived for optimizing equation (3) using (2). Similarly, for optimizing the CVaR of DNS, the clustering algorithm is applied to the 200 DNS scenarios shown in Fig. 3d, leading to the identification of 19 worst-case scenarios, with α_{ML} = 0.905 obtained for optimization.

TABLE I. COMPARISON OF MEAN AND CVAR VALUES OF CQMS: CASE – A

CQM	Base case	Conventional Approach	Proposed Approach
$\overline{TDNS}$ (MW)	1022.1	945.7	933.1
$CVaR(TDNS)$ (MW)	4570.9	4181.3	4103.2
$\overline{N_{AI}}$	2.1	2.2	2.3
$CVaR(N_{AI})$	5.0	5.3	6.0
$\overline{N_{PI}}$	2.2	1.9	1.6
$CVaR(N_{PI})$	10.9	8.9	7.7
$\overline{N_{TA}}$	8.6	7.7	6.9
$CVaR(N_{TA})$	38.7	33.8	27.1

The proposed approach yielded two ESSs and five OHLs, whereas the conventional approach resulted in one ESS and nine OHLs. Like Case A, the impact of these optimal assets against cascading failures is evaluated using mean and CVaR values of CQMs calculated across all reduced scenarios. As Case A clearly demonstrates improvement over the base case, a comparison of CQMs for this case is conducted between the conventional and proposed approaches, presented in Table II. This table shows that the optimal assets obtained from the proposed approach with α_{ML} significantly improved most CQMs compared to the conventional approach with α_{ML}. For instance, $\overline{TDNS}$ is reduced to 891.7 MW with the proposed approach, compared to 965.8 MW with the conventional approach, representing a 7.7% improvement. Likewise, for CVaR of TDNS, the proposed approach achieves a reduction to 3718.5 MW. In contrast, the conventional approach reduces it to 3990.1 MW, demonstrating a 6.8% performance advantage. While both approaches perform similarly in active islands, there is a significant improvement from the proposed approach in other CQMs reflecting the network topology's impact. For example, the average number of passive islands ($\overline{N_{PI}}$) is reduced to 1.8 using the proposed approach, in contrast to 2.1 under the conventional approach, indicating a 14.3% enhancement. Similarly, the CVaR of passive islands ($CVaR(N_{PI})$) is reduced to 8.2 by the proposed approach, compared to 9.1 with the conventional approach, demonstrating a 9.9% improvement. Additionally, regarding the number of tripped assets, both the average number of tripped assets ($\overline{N_{TA}}$) and the CVaR of tripped assets ($CVaR(N_{TA})$) show notable reductions. Specifically, the proposed approach achieves reductions to 7.6 and 30.1, compared to 8.3 and 32.9 observed with the conventional approach.

TABLE II. COMPARISON OF MEAN AND CVAR VALUES OF CQMS: CASE – B

CQM	Conventional Approach	Proposed Approach
$\overline{TDNS}$ (MW)	965.8	891.7
$CVaR(TDNS)$ (MW)	3990.1	3718.5
$\overline{N_{AI}}$	2.3	2.3
$CVaR(N_{AI})$	5	5
$\overline{N_{PI}}$	2.1	1.8
$CVaR(N_{PI})$	9.1	8.2
$\overline{N_{TA}}$	8.3	7.6
$CVaR(N_{TA})$	32.9	30.1

These findings emphasize the proposed approach's ability to significantly mitigate cascading failures compared to the conventional approach.

D. Discussion on CVaR Confidence Level 'α'

The comparison in cases A and B as illustrated in Tables I and II highlights the advantage of integrating the cascading failure model into the resilience planning framework. To determine a suitable α value for resilience investment planning, this section examines results from the proposed approach using both traditional α=0.95 and α_{ML} (i.e., comparing Tables I and II). The analysis employs the mean and CVaR values of CQMs presented in both cases related to the proposed approach.

In terms of active islands, both α=0.95 and α_{ML} result in an average of 2.3 active islands ($\overline{N_{AI}}$). However, when it comes to the CVaR of active islands ($CVaR(N_{AI})$), α=0.95 improves it to 6, while α_{ML} achieves 5. For passive islands, α=0.95 elevates the average passive islands ($\overline{N_{PI}}$) to 1.6 and the CVaR of passive islands ($CVaR(N_{PI})$) to 7.7. In contrast, α_{ML} results in 1.8 and 8.2, respectively. Similarly, α=0.95 enhances the number of tripped assets compared to α_{ML}. However, the key CQM component, TDNS, significantly improves with α_{ML}, reducing ($\overline{TDNS}$) to 891.7 MW compared to 933.1 MW with α=0.95, signifying a 4.4% improvement. Likewise, the CVaR of TDNS reduces to 3718.5 MW with α_{ML}, in contrast to 4103.2 MW with α=0.95, signifying a 9.4% improvement. This clearly demonstrates that the proposed approach with α_{ML} offers a significant improvement in the mean and CVaR values of CQMs compared to the α=0.95 approach.

To further analyze the tail behavior of the proposed approach with both α values, the complementary cumulative distribution function (CCDF) of TDNS is plotted in Fig. 4. This plot distinctly illustrates that, in most scenarios, the proposed approach with α_{ML} exhibits lower probabilities compared to the α=0.95 approach. Specifically, the tail-end probability reveals superior performance with the proposed approach using α_{ML}. For example, the probability of a 3900 MW loss with α=0.95 is 3%, reduced to 2% with α_{ML}. Moreover, it is worth noting that the probability of a loss exceeding 4100 MW is less than 0.6% in the case of the proposed approach with α_{ML}, while it is significant with α=0.95. These findings emphasize that a customized α selection approach performs notably better, particularly in handling tail-end events, compared to the traditional α=0.95 approach.

IV. Conclusion

This paper presents a novel investment planning framework designed to mitigate cascading failures while maintaining its generality. To illustrate the framework's effectiveness, it is tested with failures triggered by initiating events caused by natural hazards, with windstorms used as an example. The framework identifies optimal assets by minimizing tail-risk, specifically the CVaR of the loss function, which effectively reduces cascading propagation as quantified by cascading quantification metrics.

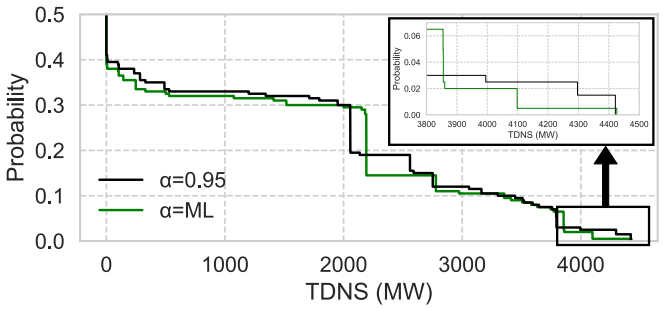

Figure 4. Comparison of CCDF of TDNS between proposed approach with α=0.95 and α_{ML}

Initially, the effectiveness of integrating cascading failure analysis with the investment planning framework is demonstrated with α=0.95. The outcomes show that the proposed approach significantly reduces the mean TDNS by 1.3% and the CVaR of TDNS by 1.9% compared to the conventional approach. Moreover, similar improvements are observed in other cascading quantification metrics when comparing the proposed approach to the conventional approach. Furthermore, the paper investigates the effectiveness of selecting a customized α in the resilient investment planning framework through the application of a machine learning algorithm. To ensure a fair comparison between the conventional and proposed approaches with α selection, both methods are examined using α derived from the machine learning algorithm. The outcomes reveal that the proposed approach outperforms the conventional approach with α_{ML}. Finally, the performance of the proposed approach with α=0.95 and α_{ML} is assessed. The results show that the proposed approach with α_{ML} effectively manages tail-risk by reducing tail-end loss probability. Additionally, it significantly reduces the mean and CVaR of TDNS by 4.4% and 9.4%, respectively, compared to the approach with α=0.95. These findings suggest that system planners should consider adopting a customized confidence level for CVaR calculations and account for cascading failures in their investment planning.

Acknowledgment

This work was funded by H2020 and Horizon Europe program through the projects "EUniversal" (Grant agreement ID: 864334), "Reliability, Resilience, and Defense Technology for the Grid" (R2D2) (Grant agreement ID: 101075714), and "HVDC-based Grid Architectures for Reliable and Resilient WideSpread Hybrid AC/DC Transmission Systems" (HVDC-WISE) (Grant agreement ID: 101075424).

References

[1] "Power Outages after Hurricane Ian." Accessed: May 17, 2023. [Online]. Available: https://earthobservatory.nasa.gov/images/150431/power-outages-after-hurricane-ian

[2] 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, doi: 10.1109/JSYST.2015.2389272.

[3] T. Lagos *et al.*, "Identifying Optimal Portfolios of Resilient Network Investments against Natural Hazards, with Applications to Earthquakes," *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 1411–1421, Mar. 2020, doi: 10.1109/TPWRS.2019.2945316.

[4] B. Venkateswaran V, D. K. Saini, and M. Sharma, "Techno-economic hardening strategies to enhance distribution system resilience against earthquake," *Reliab Eng Syst Saf*, vol. 213, 2021, doi: 10.1016/j.ress.2021.107682.

[5] M. Panteli, C. Pickering, S. Wilkinson, R. Dawson, and P. Mancarella, "Power System Resilience to Extreme Weather: Fragility Modeling, Probabilistic Impact Assessment, and Adaptation Measures," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3747–3757, 2017, doi: 10.1109/TPWRS.2016.2641463.

[6] H. Zhang, S. Ma, T. Ding, Y. Lin, and M. Shahidehpour, "Multi-Stage Multi-Zone Defender-Attacker-Defender Model for Optimal Resilience Strategy with Distribution Line Hardening and Energy Storage System Deployment," *IEEE Trans Smart Grid*, vol. 12, no. 2, pp. 1194–1205, Mar. 2021, doi: 10.1109/TSG.2020.3027767.

[7] A. Nasri, A. Abdollahi, and M. Rashidinejad, "Multi-stage and resilience-based distribution network expansion planning against hurricanes based on vulnerability and resiliency metrics," *International Journal of Electrical Power and Energy Systems*, vol. 136, Mar. 2022, doi: 10.1016/j.ijepes.2021.107640.

[8] Y. Shen, C. Gu, Z. Ma, X. Yang, and P. Zhao, "A Two-Stage Resilience Enhancement for Distribution Systems under Hurricane Attacks," *IEEE Syst J*, vol. 15, no. 1, pp. 653–661, 2021, doi: 10.1109/JSYST.2020.2997186.

[9] D. Trakas and N. D. Hatziargyriou, "Strengthening Transmission System Resilience Against Extreme Weather Events by Undergrounding Selected Lines," *IEEE Transactions on Power Systems*, pp. 1–1, Nov. 2021, doi: 10.1109/tpwrs.2021.3128020.

[10] A. Poudyal, S. Poudel, and A. Dubey, "Risk-Based Active Distribution System Planning for Resilience Against Extreme Weather Events," *IEEE Trans Sustain Energy*, pp. 1–14, Nov. 2022, doi: 10.1109/tste.2022.3220561.

[11] M. Qorbani and T. Amraee, "Long term transmission expansion planning to improve power system resilience against cascading outages," *Electric Power Systems Research*, vol. 192, Mar. 2021, doi: 10.1016/j.epsr.2020.106972.

[12] X. Lian, T. Qian, Z. Li, X. Chen, and W. Tang, "Resilience assessment for power system based on cascading failure graph under disturbances caused by extreme weather events," *International Journal of Electrical Power and Energy Systems*, vol. 145, Feb. 2023, doi: 10.1016/j.ijepes.2022.108616.

[13] J. Zhou, D. W. Coit, F. A. Felder, and S. Tsianikas, "Combined optimization of system reliability improvement and resilience with mixed cascading failures in dependent network systems," *Reliab Eng Syst Saf*, vol. 237, p. 109376, Sep. 2023, doi: 10.1016/j.ress.2023.109376.

[14] R. T. Rockafellar, S. Uryasev, and others, "Optimization of conditional value-at-risk," *Journal of risk*, vol. 2, pp. 21–42, 2000.

[15] E. Anderson, H. Xu, and D. Zhang, "Varying confidence levels for CVaR risk measures and minimax limits," *Math Program*, vol. 180, pp. 327–370, 2020.

[16] A. Deo and K. Murthy, "Optimizing tail risks using an importance sampling based extrapolation for heavy-tailed objectives," in *2020 59th IEEE Conference on Decision and Control (CDC)*, IEEE, Dec. 2020, pp. 1070–1077. doi: 10.1109/CDC42340.2020.9304026.

[17] M. Noebels, R. Preece, and M. Panteli, "AC Cascading Failure Model for Resilience Analysis in Power Networks," *IEEE Syst J*, vol. 16, no. 1, pp. 374–385, Mar. 2022, doi: 10.1109/JSYST.2020.3037400.

[18] "Extreme Wind Storms Catalogue." Accessed: Feb. 24, 2023. [Online]. Available: http://www.europeanwindstorms.org/cgi-bin/storms/storms.cgi

[19] R. D. Zimmerman and C. E. Murillo-Sánchez, "MATPOWER User's Manual Version 7.1," 2020.

[20] M. Santini, "Advantages & disadvantages of k-means and hierarchical clustering (unsupervised learning)," *URL: http://santini. se/teaching/ml/2016/Lect_10/10c_Unsupervise dMeth ods. pdf (Accesed 17.04. 2019)*.

[21] A. Tokombayev and G. T. Heydt, "High temperature low sag (HTLS) technologies as upgrades for overhead transmission systems," in *2013 North American Power Symposium (NAPS)*, IEEE, Sep. 2013, pp. 1–6. doi: 10.1109/NAPS.2013.6666834.

[22] W. Cole and A. W. Frazier, "Cost Projections for Utility- Scale Battery Storage," *National Renewable Energy Laboratory*, no. June, p. NREL/TP-6A20-73222, 2019, [Online]. Available: https://www.nrel.gov/docs/fy19osti/73222.pdf