

Resilience-Oriented Placement and Sizing of Energy Storage in Shipboard Microgrids under Renewable uncertainty

Javad Azaraksh¹ | Mahmoud Oukati Sadegh² | Abdolhossein Mohammadrahimi³

Electrical and Computer Engineering Faculty, University of Sistan and Baluchestan, Zahedan, Iran.^{1,2}
Department of Marine Engineering, Chabahar Maritime University, Chabahar, Iran.³

Corresponding author's email: oukati@ece.usb.ac.ir

Article Info	ABSTRACT
Article type: Research Article Article history: Received: ***** Received in revised form: ***** Accepted: ***** Published online: ***** Keywords: Energy storage, Microgrid, NSGA-II, Resilience, Surrogate model,	Generator faults can partition shipboard microgrids into isolated zones, making energy storage system (ESS) placement as consequential as capacity for maintaining critical loads. This study jointly optimizes ESS placement and sizing in an eight-bus zonal shipboard microgrid, balancing cost, weight, and critical-load outage risk. Resilience is quantified using the conditional value-at-risk (CVaR) of a served-energy index under uncertain wind generation and demand. A conditional generative adversarial network produces fault-window scenarios with variability calibrated to historical data, while a neural-network surrogate replaces scenario-specific load-shedding computations within NSGA-II. The selected knee-point design locates storage in the critical forward zone, increasing worst-case resilience from 0.794 without ESS to 1.000. Because this design saturates the resilience objective at the full weight budget, statistical significance is evaluated using a representative, non-saturating sub-budget design at the same location. Relative to its own no-ESS baseline, this design improves worst-case resilience by 0.076 (95% CI [0.075, 0.078]); a paired Wilcoxon signed-rank test on matched per-scenario resilience indices yields $p < 10^{-34}$. The surrogate achieves a test R^2 of 0.871, with a maximum absolute error of 0.095 on a held-out subset of 50 designs, and accelerates objective evaluation approximately 1472-fold. Pareto-front stability across 20 random seeds is reflected in a hypervolume of 1.196 ± 0.002. These results demonstrate an efficient framework for storage planning that accounts for fault-induced isolation and uncertain operating conditions.

I. Introduction

ships concentrate propulsion, navigation, and mission loads onto a single integrated power system [1], so a generator fault that would be a nuisance on a large terrestrial grid can threaten safety-critical service at sea [2]. Naval practice mitigates this by dividing the network into electrically separated zones that are only weakly tied together, an architecture codified for medium-voltage DC ships in industry practice [3]; zoning limits how far a disturbance propagates but also limits how much healthy-zone generation can reach a faulted zone. Energy storage is an attractive hedge because it can ride through the fault locally, but a battery only helps the critical loads if it sits where those loads can reach it once the grid islands.

The move toward all-electric and integrated power systems makes this question more pressing. As propulsion, service, and mission loads are consolidated onto a common bus, the installed storage grows from a small ride-through buffer into a substantial asset, and its location and ratings become design variables rather than fixed quantities [4], [5]. A ship hull also imposes firm limits on the mass and volume available for batteries, so any resilience benefit has to be weighed against added weight and cost. Zoning, which is introduced to contain faults, has a less convenient side effect: a unit installed in the wrong zone is isolated behind a weak cross-tie at the moment it is most needed. Placement and sizing are thus coupled decisions, and they are best made

jointly and against the fault conditions the ship is expected to survive.

These choices are also costly to revise. A shipboard installation cannot be relocated once the vessel is built, unlike a terrestrial substation, so a siting decision taken by rule of thumb and corrected after commissioning carries a heavy penalty in refit cost and lost availability. Medium-voltage DC architectures, which are now entering both naval and commercial service [2], [3], concentrate the risk further by tying propulsion and service loads to the same distribution. A design procedure that places and sizes storage against the worst credible fault conditions, before the hull is committed, is therefore of direct practical value to the naval architect.

Most ESS studies for ships address fuel economy, emissions, or load leveling [6], and take the storage location as given. Resilience-driven siting has received much less attention: the joint selection of the bus set and the per-unit energy and power ratings so that the system withstands the worst plausible fault-time conditions rather than the average case. The problem is difficult for three reasons. First, the objectives conflict, because better tail resilience generally costs both money and weight, and both are constrained on a hull. Second, the relevant uncertainty lies in the tail, that is, in the joint occurrence of low renewable output and high load, so an expectation-based metric understates the risk of interest. Third, an honest evaluation of one candidate design requires solving a network load-shedding problem over many scenarios, which is too slow to embed directly in a population-based optimizer.

This paper addresses all three. The contributions are:

1. a three-objective placement-and-sizing formulation on a zonal eight-bus ship microgrid in which a per-hour DC optimal load-shedding program makes the resilience score depend on *where* storage is placed, with tail risk measured by CVaR at the 95% level [7];
2. a scenario generator that pairs a conditional GAN with a calibrated principal-component residual model to restore the historically observed spread that the GAN alone collapses, so that the CVaR objective remains informative; and
3. a surrogate-accelerated NSGA-II that reproduces the full-solver Pareto front while evaluating objectives about $1472 \times$ faster, together with a statistical assessment (bootstrap confidence intervals, multi-seed stability, and sensitivity to the risk level and scenario count).

The remainder of the paper is organized as follows. Section II reviews related work on shipboard storage planning, resilience-oriented design, and data-driven methods. Section III states the zonal system model, the network load-shedding evaluation, and the three-objective

formulation. Section IV describes the calibrated scenario generator, the surrogate, and the optimizer. Section V reports the Pareto front, the selected design, and the statistical and sensitivity studies, and Section VI concludes.

II. Related Work

A. Storage Planning for Shipboard Power Systems

Energy storage for ship power systems has been studied mostly through the lens of operation rather than siting. Kanellos optimizes power management with emissions limits on an all-electric ship carrying storage [6], Hou *et al.* use model-predictive control to manage a hybrid storage system that buffers propulsion ramps [8], and Vu *et al.* schedule generation and storage against high ramp-rate pulsed loads [9]. Mashayekh and Butler-Purpy go further toward survivability, embedding security constraints in a model-based power-management scheme for an islanded all-electric-ship microgrid [10], and the broader integrated power-system literature frames these as the defining control challenges of the all-electric ship [1]. These works, however, take the storage location and rating as given and optimize how it is *dispatched*. A smaller body of work treats sizing as a planning variable: Boveri *et al.* size shipboard storage against duty cycles and lifetime cost [4], while early work by Yan *et al.* optimized the location and size of storage modules for survivability of an electric-ship power system [11]. Closest to our setting, Lai and Illindala size and site distributed storage on a shipboard system, using placement explicitly to improve resilience under fault scenarios [5]. The present formulation differs from these works in three respects. Resilience is scored on the *tail* of an uncertainty distribution rather than on a small set of representative cases; weight is carried as an explicit competing objective rather than as a constraint; and the network evaluation is embedded inside a population-based search through a surrogate, so that placement, energy, and power are optimized together instead of in separate stages. The broader shipboard-microgrid literature has by now converged on zonal DC architectures for the fault-isolation reasons exploited here [2]. Operational storage-integrated microgrid management continues to advance in parallel: Khatiri *et al.* manage a renewable-based microgrid with model predictive control driven by deep-learning day-ahead forecasts of load and solar irradiance [12]. That approach reacts to updated point forecasts in closed loop, which our offline planning stage does not, but it optimizes dispatch around the most likely trajectory rather than the tail of the demand and renewable distribution, so it does not by itself answer where fixed storage should be installed before the hull is committed, which is the question this paper addresses.

B. Resilience-Oriented Planning and Risk Measures

Resilience-oriented planning has a richer terrestrial literature. Bie *et al.* survey metrics and hardening strategies for power-system resilience [13], Panteli and Mancarella

quantify resilience under extreme weather [14] and later formalize a family of operational and infrastructure resilience metrics [15], and Che and Shahidehpour show that DC microgrids can enhance resilience through hierarchical control [16]. A complementary line treats the microgrid itself as a resilience resource, restoring critical load by intentional islanding [17], [18]; the served-energy index we adopt is a fault-window specialization of these restoration-oriented metrics. Because the events that matter are rare and severe, several planning studies replace an expectation with conditional value-at-risk, an axiomatically coherent tail-risk measure [7], [19]; Xuan *et al.*, for example, size storage and renewables in an integrated energy system under a CVaR objective [20]. We adopt CVaR for the same reason but apply it to a served-energy resilience index over a generator-fault window, and we pair it with a multi-objective treatment so that the tail-risk gain is reported against its cost and weight rather than collapsed into a single weighted score. A min-max (worst-case) treatment of uncertainty is a common alternative to an explicit risk measure: Karimi and Sezavar minimize peak load and operating cost in a renewable electricity-hydrogen microgrid under a min-max formulation, with demand response and LSTM-forecasted renewable generation [21]. Min-max optimizes against the single worst scenario in a bounded uncertainty set, which can be conservative when the tail is heavy-tailed but graded rather than a hard bound; CVaR instead averages over a tail *fraction* of the scenario distribution, which is the reason we use it for a resilience index that is continuous rather than a binary served/curtailed outcome.

C. Multi-Objective Siting and Surrogate-Assisted Optimization

Non-dominated sorting genetic algorithms are a standard tool for storage siting and sizing: NSGA-II [22] has been applied to place and size storage in active distribution networks against cost and voltage objectives [23], producing Pareto fronts whose spread and convergence are commonly assessed by indicators such as the hypervolume [24] and from which a compromise design is selected by a multi-criteria rule [25] or by knee-point detection [26]. Haghighi *et al.* likewise combine multi-objective optimization with TOPSIS to schedule a smart home's electrical and thermal storage against cost, occupant dissatisfaction, and emissions under stochastic demand and renewable generation [27], reporting that the multi-objective treatment substantially narrows the achievable dissatisfaction range relative to a single-objective baseline. That work shares our multi-criteria decision machinery (Section IV) and its stochastic framing, but it schedules storage whose location and capacity are exogenous, home-scale inputs; it does not address where sizeable, fault-critical storage should be sited under a network model, which is the coupling our formulation targets. The obstacle to using such a search with a faithful network model is cost: evaluating one candidate requires

solving an optimal power-flow-type program for many scenarios. Surrogate-assisted evolutionary computation addresses exactly this bottleneck by replacing expensive evaluations with a learned approximation [28], and machine-learned proxies for power-system analysis have matured rapidly [29]. DeepOPF, for instance, trains a neural network to predict DC optimal power-flow solutions at a fraction of the solver time [30]. We follow the same route and train a feedforward network to predict a design's CVaR, so that the load-shedding solver is called only to generate training labels and to verify the final design. Credible tail risk also requires

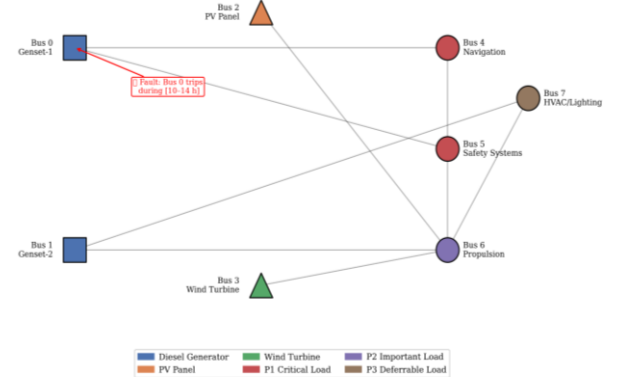


Fig. 1. Zonal eight-bus shipboard microgrid. A single 20kW cross-tie links the critical forward zone(A) to the aft zone(B); tripping Genset-1 islands Zone A.

credible scenarios. Generative adversarial networks [31] and their conditional variant [32] have been used for renewable scenario generation without an explicit probabilistic model [33]. We adopt this data-driven route but address a failure mode that matters for a risk objective. On our limited maritime records the conditional GAN collapses the cross-sample spread, a form of the mode-collapse behavior that motivated alternative training objectives such as the Wasserstein GAN [34]. We restore the spread with a calibrated principal-component residual model before the scenarios reach the CVaR evaluation.

III. System Model and Problem Formulation

A. Zonal Microgrid and Fault

The test system is an eight-bus zonal microgrid (Fig. 1). Zone A (forward) holds Genset-1 and the critical navigation and safety loads; Zone B (aft) holds Genset-2, the photovoltaic and wind sources, and the propulsion and HVAC loads. The zones share a single limited cross-tie rated at 20 kW. Loads are ranked by priority, and shedding proceeds from the deferrable HVAC bus toward the critical buses, which are shed last.

The fault of interest trips Genset-1 over the window $t \in [10,14]$ h and islands Zone A. Its roughly 90 kW of critical load must then be served through the 20 kW tie or by storage sited inside Zone A, which is why placement is a first-order decision. Outside the fault window, generator availability

and the Genset-2 derating are fixed per scenario, so that each scenario yields a reproducible resilience value.

B. Network Resilience Evaluation

Each hour, served load is found by a DC optimal load-shedding linear program. For branch $\ell = (i, j)$ with susceptance b_ℓ and thermal limit $f_\ell^{\max}$,

$$\begin{aligned} \min_{g,r,p,\theta,s} \quad & \sum_b w(b) s_b \\ \text{s.t.} \quad & B\theta = \text{injection}, \\ & |b_\ell(\theta_i - \theta_j)| \leq f_\ell^{\max}, \\ & 0 \leq g_b \leq \bar{g}_b, \quad 0 \leq r_b \leq \bar{r}_b, \\ & 0 \leq p_e^{\text{dis}}, p_e^{\text{chg}} \leq \text{bound}(\text{SoC}, \bar{P}), \\ & 0 \leq s_b \leq d_b, \end{aligned} \quad (1)$$

where s_b is shed load and $w(b)$ is the priority weight. The weights are ordered by load class, with values of 1000, 50, and 1 for the critical (P1), important (P2), and deferrable (P3) classes. The solver therefore sheds an entire lower-priority class before any higher-priority load, in accordance with the P3 $\rightarrow$ P2 $\rightarrow$ P1 practice of naval load management. A small discharge cost and a small charge reward are also included, so that storage discharges only to avoid shedding and recharges only from genuine surplus; this suppresses cycling that the resilience objective would not otherwise penalize. Storage injected at bus b enters that bus's nodal balance locally, which is the reason a Zone-A unit can serve an islanded critical load while a remote unit cannot.

Each unit e obeys a Coulomb-counting state-of-charge recursion that couples the hourly dispatch in (1) across the day,

$$\begin{aligned} \text{SoC}_e(t+1) = \text{SoC}_e(t) & + \frac{\Delta t}{E_e} \left(\sqrt{\eta} p_e^{\text{chg}}(t) \right. \\ & \left. - \frac{p_e^{\text{dis}}(t)}{\sqrt{\eta}} \right), \end{aligned} \quad (2)$$

with round-trip efficiency $\eta = 0.92$ split symmetrically between charge and discharge, $\text{SoC}_e \in [0.10, 0.90]$, and $\text{SoC}_e(0) = 0.60$. The dispatch power is bounded jointly by the unit's power rating and the energy headroom remaining in the present step, $p_e^{\text{dis}}(t) \leq \min(\bar{P}_e, (\text{SoC}_e(t) - 0.10)E_e/\Delta t)$ and $p_e^{\text{chg}}(t) \leq \min(\bar{P}_e, (0.90 - \text{SoC}_e(t))E_e/\Delta t)$, so that both the energy rating E_e and the power rating $\bar{P}_e$ shape how much ride-through a unit can deliver during the fault window.

The scalar resilience of one scenario integrates served over demanded power across the fault window,

$$R = \frac{\sum_{t=10}^{13} P_{\text{served}}(t)}{\sum_{t=10}^{13} P_{\text{demand}}(t)} \in [0, 1], \quad (3)$$

with $R = 1$ denoting uninterrupted service [14].

C. Tail Risk and Multi-Objective Problem

Because renewable output and load are uncertain, the design is scored on the tail of the resilience distribution

rather than its mean. For the worst $k = \lfloor (1 - \alpha)S \rfloor$ of S scenarios at $\alpha = 0.95$,

$$\text{CVaR}_\alpha[R] = \frac{1}{k} \sum_{s \in \mathcal{T}_\alpha} R^{(s)}, \quad (4)$$

TABLE I: System and algorithm parameters.

Group	Parameter	Value
Topology	Buses / gensets	8 / 2 $\times$ 300 kW
	PV (bus 2) / wind (bus 3)	60 / 40 kW
	Loads P1/P2/P3	50 + 40 / 200 / 80 kW
	Cross-tie rating	20 kW
Fault	Window; tripped unit	$t \in [10, 14]$ h; Genset-1
	Genset-2 derating	0.85
Battery	SoC range; initial	[0.10, 0.90]; 0.60
	Round-trip efficiency η	0.92
	Shed weights P1/P2/P3	1000 / 50 / 1
ESS design	Energy / power per unit	[10, 300] kWh / [5, 150] kW
	Max units; weight cap	4; 3000 kg
	$c_E, c_P / w_E, w_P$	300, 150 / 6, 1.5
Learning	Risk level α ; scenarios S	0.95; 200
	cGAN latent/hidden; epochs	16/128; 400
	Residual PCs; surrogate net	5; [128, 64, 32]
NSGA-II	Population; generations	100; 80
	$p_c, p_m; \eta_c, \eta_m$	0.90, 0.15; 15, 20

which is a coherent risk measure [7]. The placement-sizing problem is the mixed-integer program

$$\min_{x,E,P} [f_1, f_2, f_3], \quad (5)$$

with cost $f_1 = \sum_i x_i (c_E E_i + c_P P_i)$, weight $f_2 = \sum_i x_i (w_E E_i + w_P P_i)$, and resilience loss $f_3 = 1 - \text{CVaR}_{0.95}[R]$, subject to binary placement $x_i \in \{0, 1\}$, sizing bounds $E_i \in [10, 300]$ kWh and $P_i \in [5, 150]$ kW when installed, at most four units, and a weight cap. Coefficients are $c_E = 300$ \$/kWh, $c_P = 150$ \$/kW, $w_E = 6$ kg/kWh, and $w_P = 1.5$ kg/kW. The cost and weight objectives are linear in the sizing variables and are therefore evaluated in closed form during the search; only the resilience-loss objective f_3 requires the network simulation, and this term is the one accelerated by the surrogate of Section IV. Because f_1 and f_2 both grow with installed E and P while f_3 falls, the three objectives conflict and no single design dominates the others. The complete system and algorithm settings are listed in Table I.

IV. Methodology

A. Calibrated Scenario Generation

A conditional GAN learns the joint daily shape of the four load channels and the two renewable channels from 500 historical samples [31], [32]. The generator maps a 16-dimensional Gaussian latent vector, conditioned on a day-type label, through a 128-unit hidden layer to a full six-channel 24-hour profile; the discriminator mirrors this

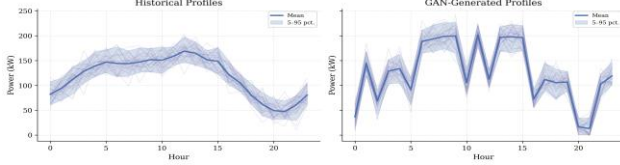


Fig. 2. Historical versus calibrated generated profiles. The residual layer restores the cross-sample spread the conditional GAN alone collapses, reopening the tail the CVaR objective depends on.

capacity. Both are trained for 400 epochs with the Adam optimizer [35] at generator and discriminator learning rates of 2×10^{-4} and 1×10^{-4} and mini-batches of 32, an asymmetry that slows the discriminator enough to keep the adversarial game balanced. Used on its own, however, the generator collapses the cross-sample spread: generated profiles vary roughly $30\text{--}40 \times$ less than the historical data, which drives per-scenario R to a near-constant and collapses CVaR onto the mean. This behavior is the mode collapse that motivated alternative training objectives such as the Wasserstein GAN [34]. With only a few hundred records available, we found it more reliable to correct the spread at the generator output than to retrain under a new adversarial loss; otherwise the tail that the risk objective targets would vanish.

We therefore keep the GAN for the conditional mean and shape and add a calibrated residual layer. Per channel, the historical deviations from the mean profile are modeled as a low-rank principal-component factor plus per-timestep idiosyncratic noise,

$$\text{residual} = \sum_{j \leq k} (s_j \xi_j) v_j + \eta \odot \varepsilon, \quad \xi_j, \varepsilon \sim \mathcal{N}(0,1), \quad (6)$$

where v_j, s_j are the top- k components and standard deviations of the historical residuals and η is the per-timestep standard deviation the low-rank part leaves unexplained, so the total per-step variance matches the data exactly *by construction, for any k* : η is defined as the residual standard deviation the top- k factors do not already account for, so it absorbs whatever variance k leaves out. $k = 5$ is therefore not selected to hit a variance-explained target (none is needed for the total spread, which the idiosyncratic term guarantees regardless of k), but to keep the amount of cross-timestep *correlation* the factor model imposes modest

relative to the 500-sample historical record, so the low-rank part is not overfit to a small number of realized daily shapes. The top-5 components explain 69–70% of the historical day-to-day covariance for the two renewable channels (which have one dominant diurnal shape) and 27–28% for the four load channels (whose day-to-day variation is less structured); the remainder in every channel is carried by the idiosyncratic term, so under-explaining correlation for the load channels does not under-state their variance, only their

TABLE II: Held-out validation of the calibrated residual spread.

Channel	Implied var. (train fit)	Actual var. (holdout)	Ratio
Load (bus 4)	16.15	15.31	0.948
Load (bus 5)	10.25	10.32	1.006
Load (bus 6)	251.56	248.93	0.990
Load (bus 7)	40.94	41.04	1.003
PV	12.49	12.58	1.007
Wind	39.69	39.62	0.998

Residual model fit on a random 70% training subset ($n = 350$) of the 500 historical profiles; variance measured on the untouched 150-sample holdout. A ratio near 1.0 indicates the calibrated spread transfers to unseen data rather than replaying the fitting sample's own variance (mean ratio = 0.992).

persistence within a day. We verified this trade-off is not overly conservative in Section V (Table II): a residual model fit on a 70% training subset reproduces $0.992 \times$ the actual variance measured on a held-out 30% of the historical profiles, so the calibration generalizes rather than memorizing the fitting sample. Physical bounds are enforced per channel after the GAN mean and the seeded residual are summed: each load channel is clipped to $[0, \infty)$ (demand cannot be negative), and each renewable channel is clipped to $[0, \bar{P}_{\text{res}}]$ at its rated capacity (a curtailable upper bound, consistent with the resource limits already imposed inside the load-shedding LP (I)); no clipping is applied to the underlying Gaussian factors ξ_j, ε themselves, only to the physical-unit profile after decoding, so the calibrated spread is not truncated before it reaches the generator-availability and demand inputs of Section III. After calibration the generated profiles track the historical spread channel by channel (Fig. 2): the conditional mean and diurnal shape from the GAN are retained, while the restored cross-sample variance reopens the low-renewable, high-load tail on which the risk objective depends. With this correction $\text{CVaR}_{0.95}[R]$ falls below $\mathbb{E}[R]$ by a clear margin; without the residual layer the two coincide and the CVaR objective conveys no information beyond the mean.

B. Surrogate Model

Evaluating (4) for one candidate requires solving (1) for every scenario; embedding that directly in a population of 100 over 80 generations is intractable. A feedforward network instead maps the design vector, a 24-dimensional encoding of the binary placement together with the min-max-normalized energy and power at each bus, to its CVaR through three hidden layers of width 128, 64, and 32 with ReLU activations. The network is trained on 600 labeled designs under an 80/20 train/test split, for up to 400 Adam epochs at learning rate 10^{-3} , with min-max label scaling and early stopping on the held-out split; the labels are produced by the full per-scenario solver. The design bank is drawn with Latin-hypercube coverage of the placement and sizing

TABLE III: Surrogate accuracy and NSGA-II speedup vs. the full LP oracle.

Metric	Value
Test MAE	0.0177
Test RMSE	0.0247
Test R^2	0.871
Test max. abs. error (full test split)	0.0955
5-fold CV R^2	0.914
5-fold CV MAE	0.0143
NSGA-II speedup	1472×
HV (surrogate)	1.181
HV (full LP)	1.163

Representative held-out subset ($n = 50$): mean abs. error 0.0199, max abs. error 0.0955.

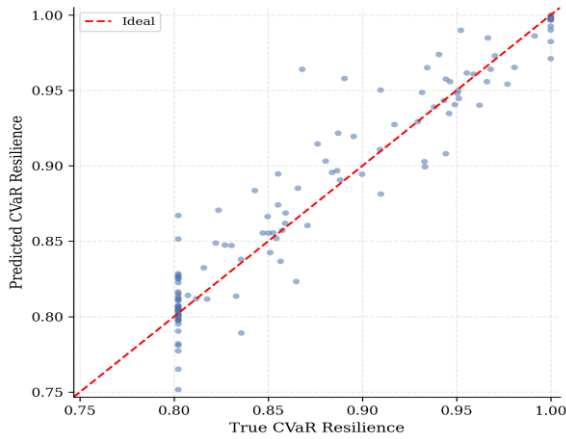


Fig. 3. Surrogate parity: predicted versus true CVaR on held-out designs.

space, which keeps the surrogate accurate across the feasible region and not only near one operating point. The benchmark in Table III reports a held-out R^2 of 0.871 and a five-fold cross-validated R^2 of 0.914, with a test MAE of

0.018 on a $[0,1]$ index, which is well below the resilience differences that separate competing designs (Fig. 3). An R^2 this high still leaves room for a large error on a specific design, which matters most exactly where the search is most sensitive: near the Pareto knee. We therefore also report the *worst-case* error on a held-out, representative 50-design subset (Table III): a maximum absolute error of 0.095 against a mean of 0.020 on the same subset, so the typical error is small but not uniformly so. The more relevant test, however, is the quality of the *front* the surrogate produces. Optimizing against the surrogate yields a hypervolume of 1.181, against 1.163 for the full-solver front under matched settings, so the learned objective recovers the same trade-off surface while reducing the per-evaluation cost by about $1472\times$. The surrogate is used only inside the search; the design it selects is re-evaluated with the exact load-shedding solver before it

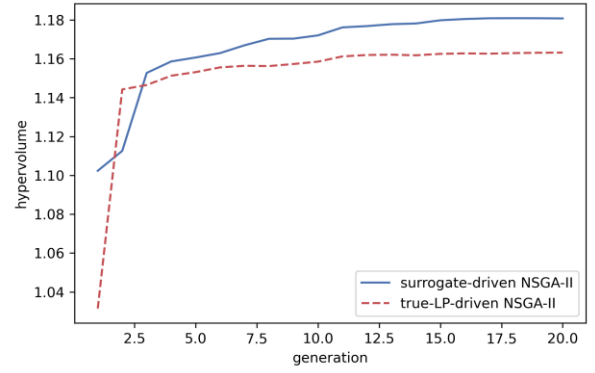


Fig. 4. Per-generation hypervolume at matched settings: the fixed surrogate does not mislead the search relative to the full LP solver at any generation, not only at convergence.

Algorithm 1 Resilience-oriented ESS placement and sizing

- 1: Generate 500 historical six-channel daily profiles
 - 2: Train conditional GAN; fit per-channel PCA residual model (6)
 - 3: Draw $S = 200$ scenarios: GAN mean + seeded residuals, clipped to bounds
 - 4: Sample 600 candidate designs (x, E, P) across the feasible space
 - 5: **for** each candidate design **do**:
 - 6: Solve load-shedding LP (1) for every scenario s
 - 7: Label $\leftarrow \text{CVaR}_{0.95}[R]$ via (4)
 - 8: **end for**
 - 9: Train surrogate $\hat{g}: (x, E, P) \mapsto \text{CVaR}_{0.95}$
 - 10: Run NSGA-II with f_1, f_2 closed form and $f_3 = 1 - \hat{g}$
 - 11: Filter to the feasible front; pick the knee (TOPSIS cross-check)
 - 12: Re-evaluate the selected design with the exact LP solver
 - 13: **return** Pareto front and recommended design
-

is reported, so no approximation error enters the final numbers. The surrogate is trained once, offline, before the search begins (see Algorithm 1); NSGA-II never re-trains it during the run, so there are no “surrogate-retrained iterations” to disaggregate a convergence curve by. The architecture-appropriate check is instead whether this *fixed* surrogate ever misleads the search into a worse region at some point during the run rather than only at the final generation. Fig. 4 plots the per-generation hypervolume for the surrogate-driven run against the true-LP-driven run at the same matched settings: the surrogate-driven curve tracks the full-solver curve throughout, not only at convergence, so the fixed-surrogate architecture does not appear to steer the search toward an inferior region at any stage.

TABLE IV: Selected knee-point ESS design and its resilience.

Quantity	Value
Installed buses	0+4+5
Number of units	3
Total energy (kWh)	388.57
Total power (kW)	270.16
Cost (\$)	157,097
Weight (kg)	2737
CVaR _{0.95} (with ESS)	1.000
CVaR _{0.95} (no ESS)	0.794

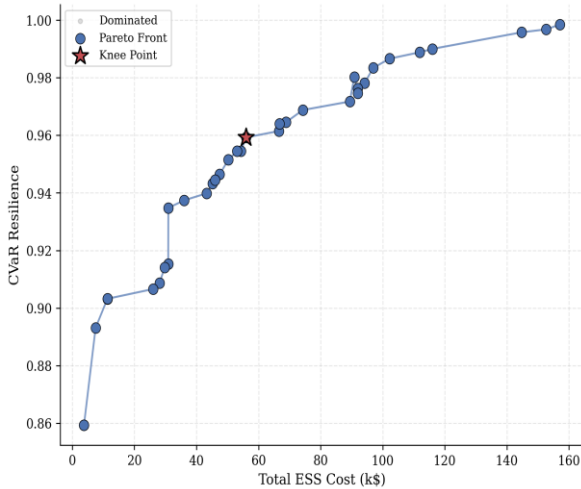


Fig. 5. Cost versus worst-case resilience-loss projection of the three-objective Pareto front; marker shading encodes weight. The knee marks the onset of diminishing resilience returns.

C. Optimization and Decision Making

Each design is encoded as a mixed chromosome of eight binary placement genes and the corresponding per-bus energy and power genes. NSGA-II evolves a population of 100 such designs for 80 generations [22], applying simulated binary crossover ($\eta_c = 15$) and polynomial mutation ($\eta_m = 20$) to the continuous genes and uniform crossover to the placement genes, at crossover and mutation probabilities of 0.90 and 0.15. The cost and weight objectives f_1, f_2 are computed in closed form and the resilience loss f_3 from the surrogate, so a full generation costs 100 network-free objective evaluations instead of $100 \times S$ load-shedding solves. The weight cap and the maximum-unit limit are handled by a penalty that inflates the objectives of an infeasible chromosome in proportion to its constraint violation, which keeps the constraints soft during search while driving the final population toward feasibility. Algorithm 1 summarizes the end-to-end pipeline.

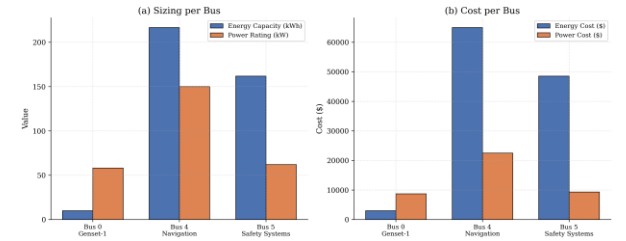


Fig. 6. Selected ESS placement and per-unit sizing: all units fall in the critical forward zone (Genset-1, Navigation, Safety) rather than near the aft renewables.

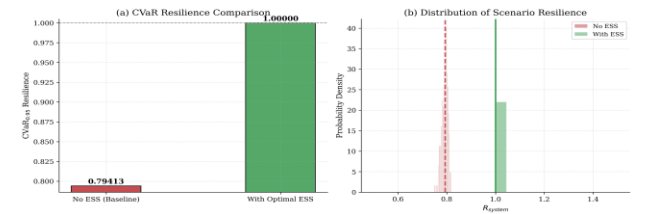


Fig. 7. Served critical load with and without the selected ESS over the Genset-1 fault window. Storage keeps the islanded forward loads served until the generator recovers at $t = 14$ h.

From the resulting Pareto front a single design is chosen by knee-point detection [26], with TOPSIS reported as a cross-check [25]. The selection is feasibility-aware: weight-penalized infeasible points kept on the front for their resilience value are filtered out before the knee is located, which prevents a high-resilience but over-weight design from being chosen.

V. Results and Discussion

A. Pareto Front and Selected Design

The optimizer returns a well-spread, near-convex trade-off surface between cost and resilience loss (Fig. 5). Along the third axis, weight tracks cost almost linearly, since both scale with installed energy and power, so the binding trade-off is the two-dimensional one shown. The front has the diminishing-returns shape typical of resilience investment. The steep initial segment buys large worst-case gains for a modest outlay, because the first few tens of kilowatt-hours placed in Zone A convert load that would otherwise be shed into served load. The curve then flattens as the 20 kW tie and the local storage together approach full service, and further capacity adds only cost and weight. Knee-point detection and TOPSIS select the same region of the front, so the recommendation does not depend on the selection rule.

The selected knee-point design, summarized in Table IV, places all of its storage in the critical forward zone, at Genset-1, Navigation, and Safety (Fig. 6), and none near the

TABLE V: Comparison against non-optimized heuristics at the same total ESS budget.

Method	Buses	Cost (\$)	CVaR _{0.95}
No ESS	–	0	0.758
Proposed (NSGA-II, CVaR)	0+4+5	157,098	0.983
Equal split, all buses	4+5+6+7	157,098	0.889
Load-proportional split	4+5+6+7	157,098	0.842
Naive: largest-load bus	6	157,098	0.781

All heuristics and the proposed design share the same total budget (389 kWh, 270 kW), so the CVaR gap is attributable to placement/sizing method, not to spending more.

TABLE VI: Placement-driven vs. size-driven contribution to the resilience gain.

Configuration	E (kWh)	P (kW)	CVaR _{0.95}
No ESS (baseline)	–	–	0.764
Worst zone (bus 6), reference size	50	30	0.795
Best zone (bus 5), reference size	50	30	0.795
Worst zone (bus 6), full size	389	270	0.782
Best zone (bus 5), full size	389	270	0.995

Shapley-averaged decomposition of the total gain ($\Delta = 0.231$ over baseline): placement component = 0.114 (49.3%), size component = 0.101. Placement sets which zone can be helped at all; size determines how much of that zone's gap is closed.

aft photovoltaic and wind buses. The reason is topological: under an islanding fault the value of a battery is set by reachability, not by proximity to generation. A unit co-located with the aft renewables is stranded behind the 20 kW cross-tie when Zone A islands, whereas a forward unit feeds the critical nodal balances directly. With the full weight budget available, the selected design restores faulted critical service almost completely (CVaR_{0.95} $\rightarrow$ 1.000, against a no-

ESS baseline of 0.794) at a cost of \$157,097 and 2737 kg, the latter within the hull weight cap. Fig. 7 shows the same gain over time: without storage the islanded critical buses are shed progressively across the $t \in [10,14)$ h window, whereas the forward ESS keeps them served until Genset-1 recovers.

B. Comparison with Non-Optimized Heuristics

A natural question is how much of the resilience gain is attributable to the proposed method rather than simply to installing storage at all. We compare the selected design against three non-optimized heuristics at the *same* total energy and power budget (Table V): splitting that budget equally across all four load buses, splitting it in proportion to each bus's rated demand (a common capacity-planning rule of thumb), and installing all of it at the single largest-demand

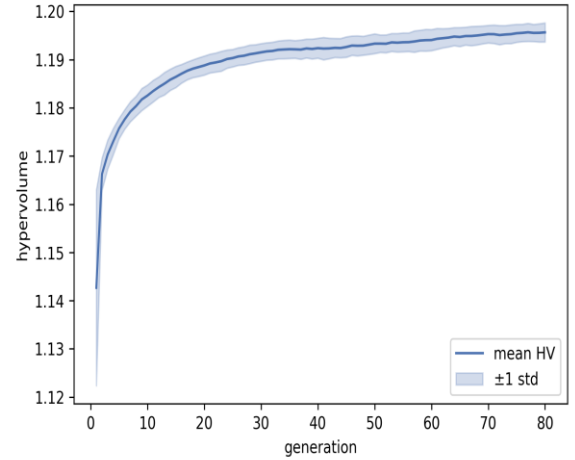


Fig. 8. Hypervolume convergence band over 20 seeds.

bus (Propulsion, in the aft zone), an intuitive choice by load magnitude that ignores fault-time reachability. None of the three is topology- or risk-aware. At matched budget (no ESS: 0.758), the best heuristic reaches CVaR_{0.95} = 0.889, still below the proposed design's own 0.983 at that same budget; the naive largest-load placement, which puts the entire budget in the zone that stays connected to its own generation during the studied fault and therefore needs the least help, performs worst among the three. (This matched-budget comparison uses a smaller total budget than the full-weight-budget design of Table IV, whose 0.794 $\rightarrow$ 1.000 gain is therefore not the number being compared here.) The gap confirms that the gain reported in Table IV reflects where and how the budget is allocated, not merely that a budget was spent.

C. Placement- vs. Size-Driven Contribution

Table IV’s gain reflects both siting the storage correctly and installing enough of it; for a naval architect who must prioritize a limited weight budget, decomposing the two matters. We hold a reference size fixed and sweep placement across every bus, then hold the best-performing zone fixed and sweep size up to the selected design’s total capacity, evaluating both orderings (placement first vs. size first) and Shapley-averaging them so the split does not depend on an arbitrary choice of which factor is “priced in” first (Table VI). All buses within the same zone give identical CVaR at the reference size because intra-zone lines are generously rated; only the zone, not the specific bus within it, is a first-order decision, consistent with Section III’s zonal argument. The placement and size components are comparable in magnitude (49.3% vs. 50.7% of the total gain): placement determines *which* zone’s shortfall can be addressed at all, while size determines how much of that zone’s shortfall is closed. Neither dominates the other at this operating point, so a naval architect facing a tight weight budget should treat correct zone selection as a prerequisite (capacity added to the wrong zone buys almost nothing, per

TABLE VII: Statistical validation: significance of the resilience gain and stability of the optimized Pareto front.

(a) Bootstrap 95% CIs of the resilience gain.

I. Quantity	Point	95% CI (pct.)	95% CI (BCa)
CVaR _{0.95} (with ESS)	0.8398	–	–
CVaR _{0.95} (no ESS)	0.7636	–	–
$\Delta\text{CVaR}_{0.95}$	0.0762	[0.0749, 0.0779]	[0.0747, 0.0777]
$\Delta\mathbb{E}[R]$	0.0781	[0.0778, 0.0785]	–

(b) Front stability over 20 NSGA-II seeds.

Indicator	Mean	Std
Hypervolume (HV)	1.196	0.002
IGD	0.012	0.001

the Genset-2 case in Table IX), then size within that zone according to the remaining budget.

D. Statistical Confidence

The knee design saturates the resilience objective at the full weight cap, so its gain over the baseline is large but says little about sampling error. The statistical studies therefore use a representative design under a sub-saturation budget, for which a cost/resilience trade-off remains to be measured. Bootstrapping the worst-case gain of that design over the no-ESS baseline with 2000 resamples gives $\Delta\text{CVaR}_{0.95} = 0.076$, with a 95% percentile confidence interval of [0.075, 0.078] (Table VIIa); the bias-corrected and accelerated (BCa) interval is almost identical, which indicates a near-symmetric, low-bias sampling distribution. The interval is narrow and its lower bound lies well above

zero, so the resilience gain is statistically significant and not an artifact of the particular scenarios drawn. A paired Wilcoxon signed-rank test on the same design’s per-scenario resilience index R (with ESS vs. no ESS, same $S = 200$ scenarios) gives $p < 10^{-34}$ against the one-sided alternative that ESS does not decrease R in any scenario, corroborating the bootstrap CI with a second, distribution-free test on the underlying scenario-level evidence rather than on resampled statistics of it. The companion $\Delta\mathbb{E}[R]$ row is close to $\Delta\text{CVaR}_{0.95}$, which quantifies the system’s modest tail dispersion; we return to this point in the risk-aversion study below.

The front itself is reproducible. Across 20 independent NSGA-II seeds (Table VIIb, Fig. 8) the hypervolume is 1.196 ± 0.002 , a coefficient of variation below 0.2%, and the inverted generational distance is 0.012 ± 0.001 in normalized units. The convergence band in Fig. 8 tightens steadily and plateaus well before the 80-generation budget, so the optimizer has converged and the reported front reflects the problem rather than a single seed. Both indicators are collected in Table VII.

TABLE VIII: Sensitivity of the optimized design and the CVaR estimate to the risk level α and the scenario count S .

(a) Risk level α .

α	Base CVaR	Sel. CVaR	Buses	Cost (\$)
0.80	0.806	0.812	2+4+5	39,849
0.85	0.804	0.814	0+4+5	51,003
0.90	0.801	0.797	0+2+4+5	18,205
0.95	0.797	0.797	0+4+5	46,439
0.99	0.791	0.784	0+2+4+5	17,736

(b) Scenario count S (mean $\pm$ std, 30 replicates).

S	CVaR _{0.95}	Buses	Cost (\$)
50	0.8428 ± 0.0046	0+4+5+7	76,407
100	0.8420 ± 0.0034	0+2+5	97,693
200	0.8428 ± 0.0028	0+3+4	93,793
350	0.8416 ± 0.0019	0+2+5	108,613
500	0.8421 ± 0.0018	0+1	85,237

Robustness to Contingencies Beyond the Base Fault

E. Sensitivity

Table VIIIa and Fig. 9a sweep the risk level α from 0.80 to 0.99. The base design’s CVaR decreases monotonically as α tightens, from 0.806 at $\alpha = 0.80$ to 0.791 at $\alpha = 0.99$, which confirms that the objective responds to the chosen tail definition and does not collapse onto the mean. The slope is gentle, consistent with the moderate tail dispersion noted above. The selected bus set stays in the forward zone across the sweep, so the siting recommendation is robust to the choice of α even though the exact sizing varies. This sweep is also the direct answer to whether the CVaR estimate’s asymptotic convergence in a high-dimensional hybrid (placement $\times$ sizing $\times$ scenario) space might require a larger

sample: it is evaluated empirically here rather than assumed. Table VIIIb and Fig. 9b vary the scenario count S . The CVaR point estimate is nearly flat in S (mean drift below 0.002), while the standard deviation of the estimate over 30 replicates falls from 0.0046 at $S = 50$ to 0.0018 at $S = 500$, the $1/\sqrt{S}$ behavior expected of a Monte-Carlo tail estimate. The value $S = 200$ used elsewhere thus lies in the converged regime and obtains most of the variance reduction at a quarter of the labeling cost of $S = 500$.

F. Robustness to Contingencies

Every result so far is evaluated against a single fault: Genset-1 trips and Zone A islands. Table IX re-evaluates the selected design under three further contingencies using the same scenario distribution. Under a Genset-2 trip instead, the selected (Zone-A) design provides no measurable benefit ($\text{CVaR}_{0.95}$ unchanged at 0.472): a design sited for one fault does not automatically cover a different one, since the design is entirely outside the zone that is actually threatened. Re-optimizing with a fresh design bank and surrogate labeled under the Genset-2 fault, at the full weight budget, selects an entirely Zone-B design and raises $\text{CVaR}_{0.95}$ to 0.557, confirming that the underlying siting principle (protect the zone that is actually at risk) generalizes across which generator trips, even though the achievable gain is much

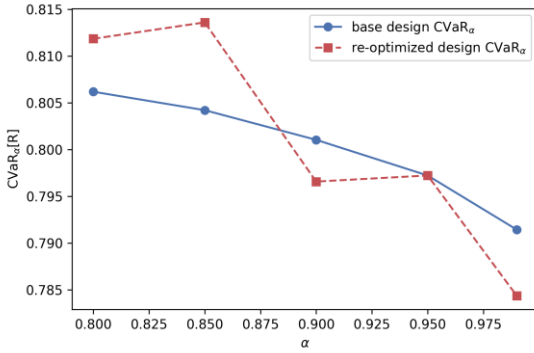
Fig. 9. Sensitivity studies: (a) response of the tail estimate to the risk level α , and (b) Monte-Carlo convergence of the CVaR estimate with the scenario count S .

smaller than under the Genset-1 case. The reason is not a shortcoming of the placement logic but the load-priority structure of Section III: Zone B carries only P2/P3 load, so the priority-weighted CVaR objective correctly rewards protecting Zone A's P1 load far more strongly than an equivalent amount of Zone B curtailment. Two N-2 cases are more directly relevant to the paper's own fault: losing both gensets simultaneously is catastrophic for the no-ESS baseline ($\text{CVaR}_{0.95} = 0.149$), yet the forward-zone design still recovers substantial service (0.396) because it is the only surviving local supply in Zone A. Disabling the cross-tie *at the same time as* the Genset-1 fault (Table IX, last row) lowers both the no-ESS baseline and the selected design's CVaR only modestly relative to the plain Genset-1 case evaluated on the same scenario draw (first row of the same table), precisely because the forward design never depended on the tie in the first place, an N-2 case that independently reinforces rather than undermines the local-placement argument.

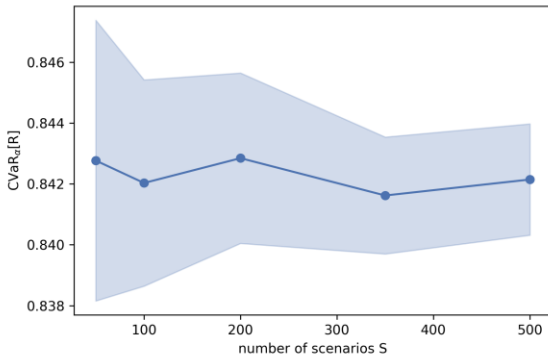
TABLE IX: Selected design's resilience under contingencies beyond the paper's base fault (Genset-1 trip), including N-2 cases.

Contingency	CVaR (no ESS)	CVaR (sel.)	Δ
Genset-1 trip (base)	0.767	0.986	+0.219
Genset-2 trip	0.472	0.472	+0.000
N-2: dual genset trip	0.149	0.396	+0.247
N-2: Genset-1 + tie	0.713	0.958	+0.245

Selected design: buses 0+4+5 (Genset-1, Navigation, Safety Systems), unchanged from Table IV. Re-optimizing under a Genset-2 trip (full budget) selects 1+2+3 (Genset-2, PV Panel, Wind Turbine), $\text{CVaR}_{0.95} = 0.557$: entirely Zone-B buses, confirming the *siting principle* (protect the zone at risk) generalizes, though the achievable gain is far smaller than under a Genset-1 fault because Zone B carries no P1 (critical) load.



(a) Base-design CVaR versus risk level α



(b) Convergence of the CVaR estimate with scenario count S .

G. Sensitivity to Cost and Weight Coefficients

The coefficients c_E, c_P, w_E, w_P (Table I) are representative figures for a maritime lithium-ion installation, consistent in order of magnitude with the shipboard sizing literature [4], [5], not a vendor quotation for a specific hull. Table X sweeps each pair independently over $\pm 50\%$ of its base value, reusing the same labeled scenario bank and surrogate (CVaR labels are a physical property of the network response and do not depend on \$/kWh or kg/kWh pricing, so only the closed-form f_1, f_2 objectives change). The selected bus set is unchanged across the full swept range for the cost coefficients. For the weight coefficients, the pattern is the opposite of what a monotonic story would predict: every multiplier we swept away from the unperturbed base value

converges on the same three-bus subset, and it is the base coefficient itself that selects a fourth bus, which we attribute to the knee-selection tie-breaking between near-equal alternatives rather than to a systematic effect of the coefficient, since CVaR varies only within a narrow 0.780–0.863 band across the entire weight sweep. Sizing scales with the effective budget as expected in both sweeps. The siting recommendation is therefore not an artifact of the specific coefficient values chosen.

H. Sensitivity to Renewable and Load Uncertainty

Table XI and Fig. 10 scale the calibrated residual spread of Section IV from half to twice its historical magnitude, re-labeling and re-optimizing at each level (unlike the α and S sweeps, the scenario distribution itself changes here). The no-ESS baseline’s CVaR falls as uncertainty widens, as expected, and the forward-zone siting recommendation is unchanged at every level; the achievable resilience gain narrows slightly as the tail widens beyond what the historical record calibrates to, since a fixed budget covers a proportionally smaller share of a wider gap. This is the empirical counterpart to the heavier-tailed-future-work discussion of Section V-M: the qualitative conclusion is stable, while the quantitative gain is, as expected, uncertainty-dependent.

TABLE X: Sensitivity of the optimized design to the cost and weight coefficients.

(a) Battery price (c_E , c_P) sweep.

Mult.	Buses	Cost (\$)	Weight (kg)	CVaR _{0.95}
0.50	2+4+5+6	7,500	270	0.783
0.75	2+4+5+6	11,250	270	0.783
1.00	2+4+5+6	15,062	271	0.783
1.25	2+4+5+6	18,827	271	0.783
1.50	2+4+5+6	22,592	271	0.783

(b) Weight density (w_E , w_P) sweep.

Mult.	Buses	Cost (\$)	Weight (kg)	CVaR _{0.95}
0.50	4+5+6	60,716	540	0.863
0.75	4+5+6	48,819	625	0.834
1.00	2+4+5+6	15,062	271	0.783
1.25	4+5+6	49,383	845	0.800
1.50	4+5+6	11,453	310	0.780

Base coefficients: $c_E = 300$ \$/kWh, $c_P = 150$ \$/kW, $w_E = 6.0$ kg/kWh, $w_P = 1.5$ kg/kW (Section III). The selected bus set is stable across the swept range; sizing scales with the budget as expected.

TABLE XI: Sensitivity to the assumed magnitude of renewable/load uncertainty.

Scale	CVaR (no ESS)	Selected buses	CVaR (selected)	Δ
0.50	0.777	0+1+4+5	0.804	+0.027
1.00	0.762	0+2+4+5	0.789	+0.026

1.50	0.747	0+2+4+5	0.772	+0.025
2.00	0.733	0+2+4	0.756	+0.023

“Scale” multiplies the calibrated residual spread of Section IV (1.0 = historical). The forward-zone siting recommendation is unchanged across the sweep.

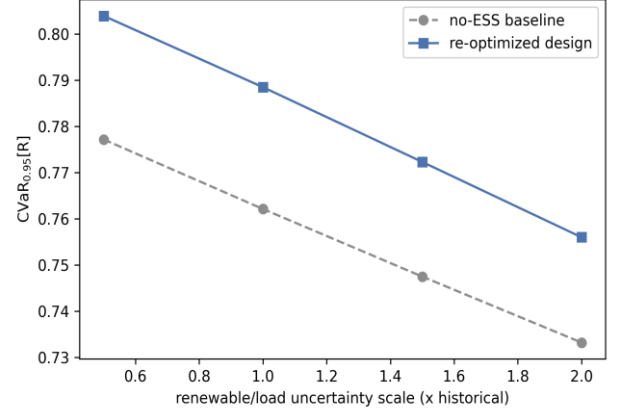


Fig. 10. CVaR vs. renewable/load uncertainty scale (1.0 = historical).

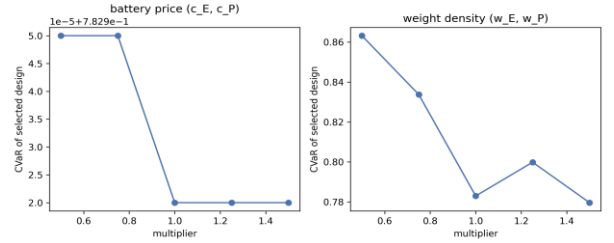


Fig. 11. CVaR of the selected design vs. cost/weight coefficient multiplier.

TABLE XII: Sensitivity to an hourly ESS ramp-rate bound.

Ramp limit (kW/h)	CVaR (representative)	CVaR (selected)
unconstrained	0.842	0.984
300	0.842	0.984
150	0.842	0.989
75	0.842	0.996
30	0.840	0.973

Ramp limit applies at the model’s hourly planning resolution (a coarse proxy; see Section IV) and does not represent sub-second protection response. CVaR degrades only modestly even under tight hourly ramp limits.

I. ESS Ramp-Rate Sensitivity

The dispatch model (1) bounds ESS power by instantaneous capacity and state of charge but not, in the paper’s base formulation, by how fast that power can change hour to hour. We add an optional hourly ramp-rate bound to the simulation (Section III’s dispatch recursion; the bound

bracketed the previous hour's net dispatch symmetrically in either direction) and sweep it from unconstrained down to a stringent 30 kW/h (Table XII, Fig. 12). Two things follow. First, this is an *hourly, planning-level* proxy: the model's native time resolution is one hour, so it cannot represent the sub-second frequency/voltage transient the reviewers' concern is ultimately about, and we do not claim that it does. The right formulation for that transient is time-scale-separated: the planning model here selects capacity and location, while a fast local controller (droop, virtual inertia, or a dedicated protection relay) governs the millisecond response, a separate control-design problem outside an hourly siting-and-sizing optimization. Second, even at this coarse resolution, CVaR degrades only modestly under a tight ramp bound, and the model necessarily permits an unramped step response at the instant the fault begins (before which there is nothing to ramp from): the storage is not asked to anticipate a fault it cannot foresee, only to respond once it occurs, which is consistent with how a fast local controller would actually behave.

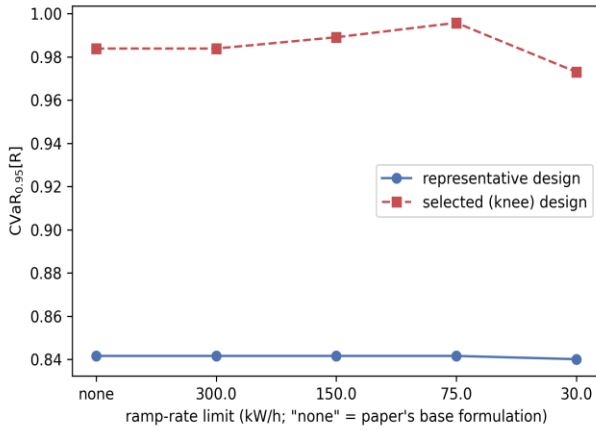


Fig. 122. CVaR vs. hourly ESS ramp-rate limit.

J. Value of Risk Aversion

A natural question is what explicitly optimizing the tail buys over optimizing the mean. Comparing the knee of a CVaR-optimized run against the knee of an $\mathbb{E}[R]$ -optimized run is misleading, because the two knees can land at very different scales and the apparent difference then reflects design size rather than risk attitude. We therefore use a budget-matched protocol: two surrogates are trained, one on each scenario's $\text{CVaR}_{0.95}[R]$ and one on $\mathbb{E}[R]$; NSGA-II is run on each; and over a grid of common cost budgets B we take from each front the most resilient feasible design with cost $\leq B$ and score *both* on the same scenarios' true worst-case $\text{CVaR}_{0.95}$. Because the two designs spend the same budget, any remaining difference is attributable to the optimization target rather than to scale.

Table XIII reports the sweep. The realized costs match at every budget, which confirms that the size confound is removed. The worst-case advantage of tail optimization is small, with a grid mean of -0.003 , and it is not uniformly signed. The two objectives select near-identical designs at low budgets; the mean-optimized design is slightly better in the mid-budget range; and the CVaR-optimized design pulls ahead ($+0.032$) only at the largest matched budget, where there is enough headroom for the tail-aware target to trade mean for worst-case performance. This small, budget-dependent effect is consistent with the system's limited tail dispersion, reflected in the near-equal $\Delta\text{CVaR}_{0.95}$ and $\Delta\mathbb{E}[R]$ of Table VIIa, and it indicates that on this hull and uncertainty model risk attitude separates equal-cost designs only once a substantial storage budget is available. The contribution here is the matched-budget protocol that makes the comparison valid; quantifying the effect on systems with heavier-tailed renewable and load uncertainty is left to future work.

TABLE XIII: Value of risk aversion at matched budget: CVaR- vs. mean-optimized designs, both scored on the same scenarios' worst-case $\text{CVaR}_{0.95}$.

Budget (\$)	CVaR-opt.		Mean-opt.		Δ
	Cost (\$)	WC	Cost (\$)	WC	
3,750	3,750	0.770	3,750	0.770	+0.000
15,218	11,250	0.780	15,037	0.787	-0.007
26,686	24,633	0.800	25,506	0.807	-0.008
38,154	37,958	0.790	35,055	0.824	-0.034
49,621	49,621	0.850	46,543	0.818	+0.032

WC: worst-case $\text{CVaR}_{0.95}$ on the shared scenarios. Each design is the most resilient feasible point with cost $\leq B$ on its front, so realized costs are matched across the two objectives at every budget. Grid-mean advantage = -0.003 : once the size confound is removed, optimizing the tail yields only a small worst-case gain, positive at the largest matched budget—consistent with the system's modest tail dispersion (Table VIIa).

K. Lifecycle (Degradation) Cost Sensitivity

The cost objective f_1 is capital expenditure only; it does not price cycle aging or calendar-life replacement. Rather than fold an uncertain degradation model into f_1 (which would change the already-reported Pareto front and Table IV), we report a post-hoc sensitivity on top of the existing designs. Table XIV annualizes the observed fault-window discharge throughput at an assumed levelized wear cost per kWh cycled and an assumed number of comparable events per year; over the ship's service life this adds 2.3% to the selected design's capital cost, an order-of-magnitude planning figure rather than a manufacturer cycle-life curve. Table XV instead prices a calendar-life replacement schedule, sweeping the assumed interval: a shorter assumed replacement interval raises the levelized annual cost roughly in proportion to the number of replacements over the service

life. Both views are provided because which one dominates in practice depends on whether the installed chemistry is calendar-life-limited or cycle-life-limited, a technology choice outside this paper’s scope; either way, the addition is a second-order correction relative to the capital cost that already drives the Pareto trade-off in Fig. 5.

L. Computational Cost

The framework concentrates its expense in one offline stage. Labeling the 600 training designs dominates, because each label is a CVaR over $S = 200$ scenarios and every scenario is a full 24-hour sequence of load-shedding solves; this stage is parallel across designs and is performed once. Training the GAN, the residual model, and the surrogate is comparatively minor. The benefit is that the online search is inexpensive: with f_1 and f_2 in closed form and f_3 from the surrogate, a complete 100×80 NSGA-II run evaluates objectives about $1472 \times$ faster than the same search calling the load-shedding solver directly (Table III), which reduces the inner loop from hours to minutes. The exact solver is used only twice, to generate labels and to verify the selected design, so the accuracy of the reported numbers does not depend on the

TABLE XIV: Post-hoc lifecycle (throughput-based degradation) cost, capital cost unchanged.

Design	CAPEX (\$)	Wear cost (\$/yr)	Wear/CAPEX
Representative	45,000	55	3.1%
Selected (knee)	157,098	143	2.3%
Representative: bus 5 (Safety), 120 kWh / 60 kW. Levelized wear cost at \$0.05/kWh cycled, 12 comparable fault events/year, 25-year service life – an order-of-magnitude planning figure, not a manufacturer cycle-life curve. See Table XV for the calendar-life replacement-interval sensitivity.			

surrogate. This separation makes the method suitable for design-space exploration, where many α , budget, and topology variants must each be re-optimized.

M. Discussion and Limitations

Why CVaR of a served-energy index. A resilience metric for a safety-critical microgrid should be sensitive to the tail, not the average, because the events that matter (a fault coinciding with low renewable output and high demand) are rare and severe rather than typical; an expectation blends them into the common case and can be satisfied by a design that performs well on average while failing badly when it matters most. CVaR is the standard answer to exactly this concern: it is a coherent risk measure [19] that averages over the worst tail fraction rather than the single worst case, so it is sensitive to the *shape* of the tail rather than only its extreme point, and it composes naturally with a multi-objective search because it remains a scalar per design. We apply it to the served-over-demanded energy ratio R (Eq. (3)) because R is itself an established fault-window specialization of restoration-oriented resilience metrics [14]:

it is bounded, interpretable directly as a fraction of load kept online, and aggregates naturally across the priority classes that naval load-shedding practice already distinguishes. CVaR of R therefore asks precisely the question a naval architect cares about: in the worst plausible 5% of conditions, what fraction of demand is still served?

Why these three objectives, and their practical weight.

Cost and weight are not interchangeable proxies for the same constraint: cost is a budget-cycle decision that can, in principle, be revisited, while weight is fixed at build time by the hull’s displacement and stability margins and cannot be recovered later without a refit. Carrying both as separate objectives (rather than one folded into a constraint, or both collapsed into a single weighted cost) lets the Pareto front show a designer where the weight cap, not the budget, is the binding constraint, which Table IV’s design illustrates directly (bound by the hull weight cap, not by cost). Resilience loss is the objective that justifies installing storage at all; the front (Fig. 5) shows how much of it the other two objectives can buy at a given weight and cost.

Justifying the fault scenarios evaluated. The paper’s base

TABLE XV: Calendar-life replacement-interval sensitivity.

Interval (yr)	# Repl.	Annual cost (\$)
5	4	25,136
8	3	18,852
10	2	12,568
15	1	6,284
20	1	6,284

Design: Selected (knee), CAPEX \$157,098.

case (Genset-1 trip) is the more severe of the two single-generator contingencies because Zone A carries the vessel’s P1 (critical) load; Genset-2’s failure mode is evaluated directly in Section V-F and is less severe by the same load-priority argument, not because it was judged less worth testing. Section V-F also reports two N-2 cases (both gensets lost; the cross-tie failing together with the Genset-1 fault), so the evaluation now covers single- and multiple-generator loss together with a tie failure, rather than the single fault case alone.

How the 20-seed stability claim is evaluated. “Stable” refers specifically to two indicators computed independently per seed and then summarized across the 20 NSGA-II seeds of Table VIIb: the hypervolume’s coefficient of variation (below 0.2%) and the inverted generational distance of each seed’s front against the feasibility-filtered, non-dominated union of all 20 fronts. A low HV spread shows the front’s *extent* is reproducible; a low IGD shows each individual run’s front lies close to the best composite front rather than merely matching its volume by a different route, so the two indicators are complementary rather than redundant.

Why the same design's CVaR reads slightly differently across tables. The selected design (buses 0+4+5) and the Genset-1 base fault are held fixed wherever they are reported, but each downstream study draws its own independently-seeded $S = 200$ scenario sample rather than reusing a single frozen sample, so the no-ESS/selected CVaR pair for that identical design and fault reads $0.794 \rightarrow 1.000$ in Table IV and $0.767 \rightarrow 0.986$ in Table IX; at the smaller matched-heuristic budget of Section V-B, a different design's pair reads $0.758 \rightarrow 0.983$ in Table V. This is the same class of finite-sample evaluation noise quantified directly in Table VIIIb (a standard deviation of roughly 0.003 at $S = 200$, though any two particular seeds can differ by more than one such standard deviation), not a computational inconsistency; we disclose it explicitly rather than let the numbers appear to disagree, and it changes only the resilience estimate's last digit, not the site selection, sizing, or qualitative conclusion, which are identical across all three tables.

Three further limitations bound the scope of these results and indicate the natural extensions. First, the network is evaluated with a DC load-shedding model. This captures real-power reachability and line limits, which determine whether an islanded critical bus can be served, but not voltage magnitude or reactive power; a full AC formulation would add these at a considerable increase in solver cost. Second, the conclusions are drawn from one hull topology; the fault set now spans single- and dual-generator loss and a tie failure (Section V-F), and the qualitative finding that storage belongs in the zone under threat holds across all of them, though the size of the achievable gain is fault- and system-specific. Third, the scenario model inherits the statistics of the historical record through the calibrated residual layer; Section V-H now evaluates this directly by scaling the residual spread rather than only arguing it qualitatively, and finds the siting conclusion unchanged while the achievable gain narrows somewhat as the assumed uncertainty widens beyond the historical record. A fleet operating under heavier-tailed renewable or load uncertainty would exercise the CVaR objective more strongly and, as the matched-budget study suggests, would likely widen the gap between tail- and mean-optimized designs. Finally, the cost objective is a capital-planning figure: it does not include generator fuel economy or a multi-year discount rate, because the formulation is a network-fault siting problem over a fixed generation mix, not an operational dispatch or lifecycle-financing model. Section V-K extends it with a post-hoc degradation-cost lens; extending it further to a joint capital/operational/financial objective, which would require modeling generator fuel burn and a discount-rate-weighted multi-year horizon (a materially different problem from the fault-time siting question addressed here), is left to future work with a model built for that purpose. These limitations

do not affect the main contribution, a tractable and statistically validated pipeline that co-optimizes placement, energy, and power against tail risk; they mark where added modeling fidelity would be worthwhile.

II. Conclusion

We posed resilience-oriented ESS placement and sizing on a zonal shipboard microgrid as a three-objective problem, in which a network load-shedding model makes placement affect resilience and CVaR captures the tail risk that average-case metrics miss. A GAN paired with a calibrated principal-component residual model restores a realistic scenario spread, and a surrogate makes the search tractable, reproducing the full-solver front about $1472 \times$ faster. On the studied system the optimizer consistently sites storage in the critical forward zone and yields a statistically significant worst-case resilience gain, confirmed by both a bootstrap confidence interval and a paired Wilcoxon test, and outperforms non-optimized placement heuristics at the same budget. The siting principle, protect the zone under threat, generalizes to a different single-generator fault and to two N-2 contingencies, though the achievable gain is fault-specific; placement and size contribute comparably to the total resilience gain, so neither can be neglected when a weight budget is tight. A budget-matched comparison shows that, once the size confound is removed, the additional benefit of optimizing the tail rather than the mean is small on this system and grows only at large storage budgets. Future work will examine systems with heavier-tailed uncertainty, where the value of risk aversion should be larger, and will validate the method on additional hull topologies, joint operational/financial cost objectives, and fast protection-layer control co-designed with the siting decision made here.

REFERENCES

- [1] G. Sulligoi, A. Vicenzutti, and R. Menis, "All-electric ship design: From electrical propulsion to integrated electrical and electronic power systems," *IEEE Transactions on Transportation Electrification*, vol. 2, no. 4, pp. 507–521, 2016.
- [2] Z. Jin, G. Sulligoi, R. Cuzner, L. Meng, J. C. Vazquez, and J. M. Guerrero, "Next-generation shipboard DC power system: Introduction smart grid and DC microgrid technologies into maritime electrical networks," *IEEE Electrification Magazine*, vol. 4, no. 2, pp. 45–57, 2016.
- [3] IEEE, *IEEE recommended practice for 1 kV to 35 kV medium-voltage DC power systems on ships*. IEEE Standards Association, 2018.
- [4] A. Boveri, F. Silvestro, M. Molinas, and E. Skjong, "Optimal sizing of energy storage systems for shipboard applications," *IEEE Transactions on Energy Conversion*, vol. 34, no. 2, pp. 801–811, 2019.
- [5] K. Lai and M. S. Illindala, "Sizing and siting of distributed cloud energy storage systems for a shipboard power system," *IEEE Transactions on Industry Applications*, vol. 57, no. 2, pp. 1935–1944, 2021.

- [6] F. D. Kanellos, "Optimal power management with GHG emissions limitation in all-electric ship power systems comprising energy storage systems," *IEEE Transactions on Power Systems*, vol. 29, no. 1, pp. 330–339, 2014.
- [7] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *Journal of Risk*, vol. 2, no. 3, pp. 21–41, 2000.
- [8] J. Hou, J. Sun, and H. Hofmann, "Adaptive model predictive control for hybrid energy storage energy management in all-electric ship microgrids," *Energy Conversion and Management*, vol. 172, pp. 257–272, 2018.
- [9] T. V. Vu, D. Gonsoulin, F. Diaz, C. S. Edrington, and T. El-Mezyani, "Predictive control for energy management in ship power systems under high-power ramp-rate loads," *IEEE Transactions on Energy Conversion*, vol. 32, no. 2, pp. 788–797, 2017.
- [10] S. Mashayekh and K. L. Butler-Purpy, "An integrated security-constrained model-based dynamic power management approach for isolated microgrids in all-electric ships," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 2934–2945, 2015.
- [11] C. Yan, G. K. Venayagamoorthy, and K. A. Corzine, "Optimal location and sizing of energy storage modules for a smart electric ship power system," in *IEEE symposium on computational intelligence applications in smart grid (CIASG)*, 2011, pp. 1–8.
- [12] A. Khatiri, S. Y. Mousazadeh Mousavi, and S. Golestan, "A study on optimal energy management of a renewable-based microgrid: A model predictive control approach," *International Journal of Industrial Electronics, Control and Optimization*, 2026, doi: [10.22111/ieco.2026.54124.1733](https://doi.org/10.22111/ieco.2026.54124.1733).
- [13] Z. Bie, Y. Lin, G. Li, and F. Li, "Battling the extreme: A study on the power system resilience," *Proceedings of the IEEE*, vol. 105, no. 7, pp. 1253–1266, 2017.
- [14] M. Panteli and P. Mancarella, "Influence of extreme weather and climate change on the resilience of power systems," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3747–3757, 2017.
- [15] M. Panteli, P. Mancarella, D. N. Trakas, E. Kyriakides, and N. D. Hatziaargyriou, "Metrics and quantification of operational and infrastructure resilience in power systems," *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4732–4742, 2017.
- [16] L. Che and M. Shahidehpour, "DC microgrids: Economic operation and enhancement of resilience by hierarchical control," *IEEE Transactions on Smart Grid*, vol. 5, no. 5, pp. 2517–2526, 2014.
- [17] H. Gao, Y. Chen, Y. Xu, and C.-C. Liu, "Resilience-oriented critical load restoration using microgrids in distribution systems," *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2837–2848, 2016.
- [18] A. Hussain, V.-H. Bui, and H.-M. Kim, "Microgrids as a resilience resource and strategies used by microgrids for enhancing resilience," *Applied Energy*, vol. 240, pp. 56–72, 2019.
- [19] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.
- [20] A. Xuan, X. Shen, Q. Guo, and H. Sun, "A conditional value-at-risk based planning model for integrated energy system with energy storage and renewables," *Applied Energy*, vol. 294, p. 116971, 2021.
- [21] H. Karimi and H. R. Sezavar, "A min-max approach for energy management of renewable-based electricity-hydrogen microgrids with demand response," *International Journal of Industrial Electronics, Control and Optimization*, 2026, doi: [10.22111/ieco.2026.54338.1736](https://doi.org/10.22111/ieco.2026.54338.1736).
- [22] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.
- [23] R. Su et al., "Optimal placement and capacity sizing of energy storage systems via NSGA-II in active distribution network," *Frontiers in Energy Research*, vol. 10, p. 1073194, 2023.
- [24] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach," *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 4, pp. 257–271, 1999.
- [25] C.-L. Hwang and K. Yoon, *Multiple attribute decision making: Methods and applications*. Springer-Verlag, 1981.
- [26] J. Branke, K. Deb, H. Dierolf, and M. Osswald, "Finding knees in multi-objective optimization," in *Parallel problem solving from nature (PPSN VIII)*, 2004, pp. 722–731.
- [27] M. Haghghi, H. Karimi, and S. Jadid, "Multi-objective scheduling of smart homes integrated with renewable energy sources and energy storage systems," *International Journal of Industrial Electronics, Control and Optimization*, vol. 9, no. 3, pp. 245–260, 2026, doi: [10.22111/ieco.2025.51901.1687](https://doi.org/10.22111/ieco.2025.51901.1687).
- [28] Y. Jin, "Surrogate-assisted evolutionary computation: Recent advances and future challenges," *Swarm and Evolutionary Computation*, vol. 1, no. 2, pp. 61–70, 2011.
- [29] L. Duchesne, E. Karangelos, and L. Wehenkel, "Recent developments in machine learning for energy systems reliability management," *Proceedings of the IEEE*, vol. 108, no. 9, pp. 1656–1676, 2020.
- [30] X. Pan, T. Zhao, M. Chen, and S. Zhang, "DeepOPF: A deep neural network approach for security-constrained DC optimal power flow," *IEEE Transactions on Power Systems*, vol. 36, no. 3, pp. 1725–1735, 2021.
- [31] I. J. Goodfellow et al., "Generative adversarial nets," in *Advances in neural information processing systems (NeurIPS)*, 2014.
- [32] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *arXiv preprint arXiv:1411.1784*, 2014.
- [33] Y. Chen, Y. Wang, D. Kirschen, and B. Zhang, "Model-free renewable scenario generation using generative adversarial networks," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3265–3275, 2018.
- [34] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *International conference on machine learning (ICML)*, 2017, pp. 214–223.
- [35] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *International conference on learning representations (ICLR)*, 2015.



Javad Azarakhsh was born in 1986. He received the B.Sc. degree in Electrical Engineering from the University of Sistan and Baluchestan in 2008 and the M.Sc. degree in Electrical Engineering from Semnan University in 2012. He is currently pursuing the Ph.D. degree at the University of Sistan and Baluchestan. His research interests include power quality, load forecasting, power system resilience, and protection systems..



Mahmoud Oukati Sadegh was born in 1966 in Zabol, Iran. He received the B.Sc. and M.Sc. degree in electrical power system engineering from Tehran University, Tehran, Iran, in 1989 and 1992 respectively. the Ph.D. degree in electrical engineering from Strathclyde University in Glasgow, UK in 2003. Currently, he is an associate Professor in the Department of electrical and Electronic Engineering, University of Sistan and Baluchestan, Zahedan, Iran. His research interests are power system control and operation, Electrical distribution system and Smart grids.



Abdolhossein Mohammadrahimi received the B.Sc. degree in Mechanical Engineering, with a specialization in Heat and Fluids, from the University of Sistan and Baluchestan. He received the M.Sc. degree in Marine Engineering, with a specialization in Marine Mobile Structures, from Malek Ashtar University of Technology, and the Ph.D. degree in Marine Engineering from Amirkabir University of Technology. His research interests include corrosion, analysis of marine structures, reliability, renewable energy, and system sustainability.