

Article

Digital Cyber-Physical Modeling and Risk-Constrained Multi-Agent Control of Virtual Power Plants with Performance-Linked Resilience Finance

Tianze Zeng ¹, Biao Yang ² , Jingru Yu ², Hong Tan ^{3,*} and Alexis P. Zhao ⁴

¹ School of Computer Science and Engineering, University of New South Wales, Sydney, NSW 2052, Australia; tianze.zeng@student.unsw.edu.au

² State Grid Energy Research Institute Co., Ltd., Beijing 102209, China; yangbiao@sgeri.sgcc.com.cn (B.Y.); jingru0216@outlook.com (J.Y.)

³ College of Electrical Engineering and New Energy, China Three Gorges University, Yichang 443002, China

⁴ Department of Energy Science and Engineering, Stanford Doerr School of Sustainability, Stanford University, Stanford, CA 94305, USA; alexzhao@stanford.edu

* Correspondence: tanhong@ctgu.edu.cn

Abstract

Cyber incidents can disrupt many virtual power plant (VPP) assets through shared software and communication services. This study links preventive finance, cyber defense, dispatch, and restoration in one multi-timescale model. An attacker, a VPP operator, a bond vehicle, and a regulator interact in a partially observable stochastic game. The operator controls hardening, dispatch, isolation, and recovery. The bond provides restricted pre-event capital and releases collateral through an auditable index of service loss, control availability, network stress, and recovery delay. A risk-constrained multi-agent policy enforces power-system feasibility, investor impairment, sponsor affordability, and trigger-loss limits. Tests use transparent synthetic VPP-39 and VPP-118 portfolios and 20 out-of-sample seeds. The proposed design lowers normalized social cost to 0.691 and 0.704 and weighted basis risk to 0.065 and 0.071. It also improves critical-load continuity and restoration relative to self-insurance and three bond baselines. The VPP-118 case recovers in 11.8 h, compared with 14.3 h for the closest rule-based benchmark. Ablations separate the effects of finance and control. Removing the coupon-control link reduces verified hardening from 0.672 to 0.519. Removing the safety projection raises unsafe proposals from 0.4% to 5.9%. These results show that stochastic multi-timescale control can support adaptable and resilient VPP operation while keeping the financial mechanism within explicit risk limits.

Keywords: virtual power plant; cyber resilience; multi-agent control; risk constraints; resilience bond; service recovery



Academic Editor: Andreas Poullikkas

Received: 20 August 2026

Revised: 4 September 2026

Accepted: 7 September 2026

Published: 11 September 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](#)

[Attribution \(CC BY\)](#) license.

1. Introduction

Virtual power plants (VPPs) combine distributed generation, storage, flexible demand, electric vehicles, and industrial loads as one controllable portfolio. Aggregation can provide energy, reserves, congestion relief, and restoration support. It also creates shared digital dependencies. A compromised cloud scheduler, identity service, telemetry broker, or fleet interface can influence many assets at once. The resulting event may combine corrupted measurements, synchronized load changes, unavailable flexibility, market imbalance, and delayed recovery. A VPP is therefore a coupled physical-digital system whose portfolio risk is not visible from component failure rates alone [1–3].

Power-grid research has described false-data injection, communication denial, malicious commands, and topology manipulation [4–9]. Detection methods for VPP network attacks and broader reviews of cyber–physical energy-system threats also show that security has to be evaluated through operational consequences, not packet anomalies alone [10–12]. These studies provide the technical basis for defense. Yet most control models treat the hardening budget as fixed. They do not ask how pre-event capital changes defense, dispatch, and recovery together.

Risk-transfer work addresses the loss layer. Cyber catastrophe bonds can securitize a defined tail when attachment probability and severity are modeled explicitly [13]. Cyber insurance-linked securities may add capacity, although accumulation risk and model uncertainty affect the return required by investors [14]. Resilience-bond research goes one step earlier. It links validated risk reduction to capital for mitigation [15–17]. The concept is not absent from cyber or infrastructure finance. The narrower gap is that existing designs rarely place verified financing terms inside a VPP’s dispatch, attack, and restoration loop. Recent work on cyber insurance also stresses unresolved information, modeling, and protection gaps [18].

VPP research is advancing quickly in forecasting, market coordination, reliability, and energy sharing [19–21]. Recent studies address photovoltaic fault diagnosis [22], volatility-aware photovoltaic forecasting [23], integrated microgrid operation [24], cloud architectures for intelligent VPPs [25], renewable forecasting with communication access [26], uncertainty-aware Stackelberg coordination [27], multi-market decisions [28], and dynamic pricing with peer-to-peer exchange [20,29,30]. These methods improve operation under uncertainty. They do not combine adaptive cyber attacks, network-feasible restoration, audited hardening finance, collateral settlement, and trigger-basis risk in one decision process.

The model in this paper treats the operating rule as a constrained partially observable stochastic game. An attacker, a VPP operator, a special-purpose bond vehicle, and a regulator act with different observations and objectives. Fast actions cover attack response, dispatch, isolation, and recovery. Slow actions cover audit, principal, coupon, attachment, exhaustion, and rebate. Centralized critics use the complete simulated state during training. Each deployed actor uses only its local information. A feasibility projection enforces finite power, energy, ramping, network, attack, collateral, and liquidity limits. Dual variables enforce episode-level limits for unsafe operation, investor impairment, affordability, and trigger–loss mismatch [31–35].

The main objective is to determine whether performance-linked resilience finance improves VPP resilience because it changes verified preventive behavior, rather than because it merely moves losses between the operator and investors. The contributions are fourfold. First, the framework combines physical operation, shared digital exposure, adaptive attacks, recovery, and collateral accounting. Second, it links coupon and rebate terms to independently verified controls. Third, it uses an auditable physical–digital trigger with an explicit basis-risk constraint and non-repeating impairment accounting. Fourth, it tests the financial and control components separately through ablation, sensitivity, cash-flow, and scaling analyses.

Three hypotheses make the evaluation falsifiable. **H1:** linking financing terms to verified controls increases hardening and reduces tail loss relative to transfer-only contracts. **H2:** the safety projection and risk multipliers keep operational, impairment, affordability, and basis-risk outcomes within their budgets without removing the performance gain. **H3:** the multi-timescale, role-specific policy outperforms a single cooperative actor under matched operating and attack paths. Section 2 positions these hypotheses in the literature. Section 3 gives the research method. Sections 4 and 5 define the physical–digital system,

bond accounting, and learning rule. Section 6 reports the tests. Sections 7 and 8 document reproducibility, limitations, and deployment requirements.

2. Literature Review

VPP cybersecurity spans devices, communications, aggregation platforms, markets, and grid interfaces [2,3]. Stealthy measurement attacks can evade conventional residual tests [4,7,9]. Availability and timing attacks can also degrade a control loop even when the underlying equipment remains intact [5,6,8]. Coordinated high-wattage loads illustrate how many individually small devices can create a system event [36]. Recent unsupervised VPP detection research improves recognition of network attacks [11], while a recent energy-systems review organizes attack surfaces and countermeasures across the cyber–physical chain [10,12]. Detection and control remain necessary, but their budgets are usually external inputs.

Cyber insurance models describe interdependent security, information asymmetry, and correlated loss [37–39]. Capital-market instruments can transfer a defined cyber tail [13,14,29,40]. Resilience bonds instead return part of modeled risk reduction to preventive work [15–17]. The operational value of that rebate depends on verification, dispatch conditions, and recovery readiness. A parametric trigger also introduces basis risk because rapid settlement may differ from realized economic loss. The contract must therefore be evaluated together with behavior and physical service.

CVaR supplies a coherent tail measure [31]. Constrained Markov decision processes provide cumulative safety and resource limits [32]. Stochastic games represent agents whose actions alter one another’s transitions and rewards [41,42]. Centralized-training/decentralized-execution methods, PPO, and constrained policy optimization provide the learning components [9,33–35,43]. PowerGridworld, PowerNet, and distributed restoration studies show how these methods can be connected to grid constraints [10,19–21,44–46]. Appendix A compares these streams and recent 2024–2026 studies. The comparison isolates the remaining intersection: network-feasible VPP control with adaptive cyber behavior, verified ex ante capital, cumulative collateral accounting, and measurable trigger basis risk.

3. Research Methodology

3.1. Objectives, Questions, and Test Logic

The study has four objectives. It defines a digital cyber–physical VPP model with finite engineering limits; embeds a performance-linked bond in that model; learns role-specific decisions across operational and renewal time scales; and tests whether each design element produces a measurable benefit. The central research question is whether verified resilience finance can change preventive and restorative action without weakening physical feasibility or shifting excessive risk to investors. Two supporting questions examine which gains come from the financial link and which come from the controller, and how performance changes with portfolio scale and deployment stress.

Table 1 gives the complete research algorithm. It also states the evidence that could reject each hypothesis. The workflow follows control engineering, power-system simulation, sequential decision analysis, and risk-finance accounting. Every comparison uses matched operating and attack paths so that contract regimes do not receive different scenarios.

3.2. Experimental Design and Statistical Treatment

The independent design variable is the operating and financing regime. The main outcomes are normalized social cost, critical-load interruption, loss CVaR, restoration time, verified hardening, trigger-basis risk, and unsafe proposals. Self-insurance, fixed, EVT, and rule-based bonds serve as baselines. The full policy and four ablations isolate

causal mechanisms within the synthetic environment. Portfolio means are reported across 20 evaluation seeds. Each seed contains 1500 attack episodes. Uncertainty intervals are nonparametric 95% bootstrap intervals over seed means. The work is computational and uses no human participants, confidential utility records, or proprietary insurance claims.

Table 1. Research algorithm and falsification logic.

Step	Operation	Method and Output	Decision Rule
1	Specify system	Construct finite VPP portfolios, shared digital services, operating profiles, attack library, and renewal calendar.	Reject a case if base operation cannot satisfy balance and equipment limits.
2	Define mechanisms	Specify the trigger, cumulative impairment, rebate eligibility, participant rewards, and four risk budgets.	Check trigger monotonicity, collateral conservation, and bounded actions.
3	Generate matched paths	Draw demand, renewable, price, communication, vulnerability, and attack sequences once per seed. Reuse them across regimes.	Preserve seed-level pairing for every comparison.
4	Train policies	Apply centralized critics, local actors, a feasibility projection, PPO updates, and projected multiplier updates.	Select the checkpoint with minimum validation cost among checkpoints satisfying all budgets.
5	Test out of sample	Evaluate 20 unseen seeds with 1500 attack episodes per seed. Reconstruct physical, risk, and cash-flow metrics.	H1 requires higher hardening and lower tail loss than transfer-only bonds.
6	Isolate mechanisms	Remove the coupon–control link, basis-risk penalty, safety projection, or separate actors one at a time.	H2/H3 fail if the relevant removal has no measurable adverse effect.
7	Stress and scale	Vary renewable penetration, delay, attacks, prices, degradation, forecast error, and graph size.	Report deterioration and identify the boundary of supported claims.

4. Digital Cyber-Physical Modeling of VPP Operation

4.1. Agents, Digital State, and Time Scales

The digital environment contains an attacker A , VPP operator O , bond vehicle/investor B , and regulator R . A game episode has a finite horizon of 96 fifteen-minute operating intervals within a renewal state. Contract decisions are held fixed between renewal boundaries. The attacker observes exposed services, coarse demand conditions, and past defensive responses. The operator observes internal telemetry, resource availability, verified control status, and contract terms. The SPV observes declared exposure, audit results, trigger records, payment history, and outstanding collateral. The regulator observes service quality and disclosure variables. Centralized critics use the complete simulated state during training. Each deployed actor receives only its local observation. This is the rule of the constrained game: every agent maximizes its own discounted objective, physical actions are projected onto the feasible engineering set, and episode-level risk costs must remain below the four limits in Equation (27).

The interaction is represented as a partially observable stochastic game:

$$\mathcal{G} = \langle \mathcal{N}, \mathcal{S}, \{\mathcal{A}^i\}_{i \in \mathcal{N}}, \mathcal{P}, \{\mathcal{R}^i\}_{i \in \mathcal{N}}, \{\mathcal{O}^i\}_{i \in \mathcal{N}}, \gamma \rangle, \quad (1)$$

where $\mathcal{N} = \{A, O, B, R\}$, $\mathcal{P}$ is the transition kernel, and γ is the discount factor. Contract variables update at renewal epochs, whereas operational actions update within each

cyber-physical interval. A renewal mask holds the SPV action constant between eligible decision times.

The global state collects physical, cyber, financial, and recovery information:

$$s_t = (x_t, v_t, h_t, m_t, q_t, c_t, k_t, \omega_t), \quad (2)$$

where x_t is the electrical state, v_t vulnerability, h_t verified hardening, m_t attack and audit history, q_t restoration progress, c_t bond terms, k_t collateral and sponsor liquidity, and ω_t exogenous demand, renewable, communication, and equipment uncertainty. Table 2 summarizes the notation used throughout the formulation.

Table 2. Nomenclature for the cyber-physical and financial model.

Symbol	Meaning	Symbol	Meaning
$\mathcal{N}$	Set of strategic agents	s_t, o_t^i	Global state and local observation
x_t	Electrical operating state	v_t	Cyber-vulnerability vector
h_t	Independently verified hardening	q_t	Restoration progress
a_t^A	Attack action	a_t^O	Operator action
a_t^B	Bond/SPV action	a_t^R	Regulatory action
p_t^G, p_t^L, p_t^{DER}	Generation, load, and DER injection	ℓ_t	Controlled or involuntary load curtailment
B_{bus}, B_f	Bus and branch susceptance matrices	θ_t, f_t	Voltage angle and branch flow
α_t, σ_t	Attack intensity and stealth	u_t, r_t	Defense and recovery actions
χ_t	Attack-success indicator	ρ_t	Functional service level
Z_t	Realized economic loss	I_t	Parametric trigger index
$P_r^B, K_{r,t}$	Issued principal and outstanding collateral	$\phi_{r,t}^*$	Raw event impairment fraction
a_r, e_r	Attachment and exhaustion points	$y_{r,t}$	Principal released to sponsor
κ_r	Coupon rate	b_r	Restricted resilience rebate
$\Phi_{r,t}, \Delta\Phi_{r,t}$	Cumulative and incremental impairment	W_t	Sponsor liquidity
BR	Trigger basis-risk measure	$CVaR_\beta$	Conditional value-at-risk at level β
Q_i, V_i	Centralized action and state values	λ	Vector of constraint multipliers
θ_i, ψ_i	Actor and critic parameters	$\Pi_{\mathcal{F}}$	Projection onto feasible set

The joint action separates strategic roles:

$$a_t = \begin{bmatrix} a_t^A \\ a_t^O \\ a_t^B \\ a_t^R \end{bmatrix} = \begin{bmatrix} \tau_t, \iota_t, \alpha_t, \sigma_t \\ u_t, d_t, r_t, \ell_t \\ P_t^B, \kappa_t, a_t, e_t, b_t \\ \underline{\rho}_t, \bar{p}_t, \zeta_t \end{bmatrix}, \quad (3)$$

where the attacker selects target τ_t , type ι_t , intensity α_t , and stealth σ_t ; the operator selects hardening u_t , dispatch d_t , recovery r_t , and curtailment ℓ_t ; the SPV sets the bond terms; and the regulator sets minimum service $\underline{\rho}_t$, an impairment tolerance $\bar{p}_t$, and a disclosure or noncompliance penalty ζ_t .

The transition is factorized for implementation:

$$\mathcal{P}(s_{t+1} | s_t, a_t) = \mathcal{P}_x \mathcal{P}_v \mathcal{P}_h \mathcal{P}_m \mathcal{P}_q \mathcal{P}_c \mathcal{P}_k, \quad (4)$$

with every factor conditioned on the state, action, and uncertainty. The factorization delegates updates to electrical, cyber, audit, settlement, and liquidity modules without asserting independence.

4.2. Physical and Cyber Dynamics

A linearized network model provides a fast screening layer. At each interval, nodal balance is

$$B_{\text{bus}}\theta_t = p_t^G + p_t^{\text{DER}} - p_t^L - \ell_t + \Delta p_t^A, \quad (5)$$

where Δp_t^A is the net injection error created by a successful attack. Branch flows and emergency ratings satisfy

$$f_t = B_f\theta_t, \quad -\bar{f} \leq f_t \leq \bar{f}. \quad (6)$$

The simulator resolves residual infeasibility through prioritized curtailment and records the violation before correction. The linearized screening layer is retained throughout the experiments to preserve a consistent physical state definition across all financial regimes.

Supply, demand, storage, ramping, and attack actions are finite at every interval. For renewable unit j , battery b , and controllable unit g , the feasible set contains

$$\begin{aligned} 0 \leq p_{j,t}^{\text{ren}} \leq \bar{p}_j^{\text{ren}} \omega_{j,t}^{\text{ren}}, \quad & 0 \leq p_{b,t}^{\text{ch}} \leq \bar{p}_b^{\text{ch}}, \quad & 0 \leq p_{b,t}^{\text{dis}} \leq \bar{p}_b^{\text{dis}}, \\ E_{b,t+1} = E_{b,t} + \eta_b^{\text{ch}} p_{b,t}^{\text{ch}} \Delta t - \frac{p_{b,t}^{\text{dis}} \Delta t}{\eta_b^{\text{dis}}}, \quad & 0.10 \bar{E}_b \leq E_{b,t} \leq 0.90 \bar{E}_b, \quad & (7) \\ |p_{g,t} - p_{g,t-1}| \leq R_g, \quad & 0 \leq \ell_{j,t} \leq p_{j,t}^L, \quad & \sum_t \alpha_t \leq B_A. \end{aligned}$$

The baseline uses $\Delta t = 0.25$ h, $\eta^{\text{ch}} = \eta^{\text{dis}} = 0.95$, battery power equal to the node's DER rating, and energy equal to four hours of rated discharge. Available supply is capped by the sampled operating profile. Load cannot be curtailed below zero, recovery effort is bounded by available staff and reserve energy, and the attacker's cumulative intensity cannot exceed B_A . These limits, the branch ratings in Equation (6), and the state clipping in Equations (10) and (11) form $\mathcal{F}(s_t)$. Thus, neither supply, demand, attack effort, nor restoration can grow with elapsed time.

Attack effects combine load manipulation, DER setpoint corruption, and availability loss:

$$\Delta p_t^A = \chi_t \left(M_{\text{ti}}^L z_{\text{ti}}^L - M_{\text{ti}}^{\text{DER}} z_{\text{ti}}^{\text{DER}} \right) \alpha_t. \quad (8)$$

The success indicator is sampled from a probability that rises with vulnerability, intensity, and stealth and falls with verified hardening:

$$\Pr(\chi_t = 1 \mid s_t, a_t) = \left[1 + \exp \left\{ - \left(\beta_0 + \beta_v^\top v_t + \beta_\alpha \alpha_t + \beta_\sigma \sigma_t - \beta_h^\top h_t \right) \right\} \right]^{-1}. \quad (9)$$

Hardening persists but depreciates as software, credentials, and audits age:

$$h_{t+1} = \Pi_{[0,1]} \left((1 - \delta_h) h_t + G_u u_t + G_b b_t - \Gamma_A \chi_t \alpha_t + \varepsilon_t^h \right). \quad (10)$$

The rebate enters only through eligible, audited investments. If verification fails, $G_b b_t$ is removed and the regulator applies the contractual penalty. Restoration progress follows

$$q_{t+1} = \Pi_{[0,1]} \left(q_t + H_r r_t + H_d d_t - H_A \chi_t \alpha_t + \varepsilon_t^q \right). \quad (11)$$

System functionality weights critical and noncritical service:

$$\rho_t = \frac{\sum_{j \in \mathcal{L}} w_j \left(p_{j,t}^L - \ell_{j,t} \right)}{\sum_{j \in \mathcal{L}} w_j p_{j,t}^L},$$

where w_j is larger for critical facilities and essential services. Realized economic loss combines interruption, emergency operation, equipment stress, market imbalance, recovery effort, and regulatory penalties:

$$Z_t = c_\ell^\top \ell_t + c_f^\top [|f_t| - \bar{f}]_+ + c_m |\Delta p_t^{\text{imb}}| + c_r^\top r_t + \zeta_t \mathbf{1}\{\rho_t < \underline{\rho}_t\}. \quad (12)$$

This definition separates physical harm from settlement: a contract can reallocate Z_t , but it cannot erase interruption.

5. Performance-Linked Resilience Bond and Intelligent Control

5.1. Parametric Trigger and Cash Flows

The bond is fully collateralized. Investor principal is held in a trust, the operator pays a coupon, and the SPV releases the principal when an auditable trigger crosses the attachment point. The trigger combines service loss, network stress, control unavailability, and recovery delay:

$$I_t = w_\rho(1 - \rho_t) + w_f \frac{\| [f_t] - \bar{f} \|_1}{\|\bar{f}\|_1} + w_c(1 - \eta_t^{ctrl}) + w_q(1 - q_t), \quad (13)$$

where the nonnegative weights sum to one and η_t^{ctrl} is the share of the contracted device fleet retaining authenticated controllability. Each component is reconstructible from signed telemetry, market records, or independent incident logs. This construction limits discretionary claim adjustment while preserving a link to physical performance.

Let r identify an issuance period and P_r^B denote the principal funded at its opening. The raw event impairment is a bounded linear layer between attachment a_r and exhaustion e_r :

$$\phi_{r,t}^* = \min \left\{ 1, \max \left[0, \frac{I_t - a_r}{e_r - a_r} \right] \right\}, \quad 0 \leq a_r < e_r \leq 1. \quad (14)$$

The contract records the maximum impairment reached within the issuance. It pays only the increment that has not already been released:

$$\begin{aligned} \Phi_{r,t} &= \max \{ \Phi_{r,t-1}, \phi_{r,t}^* \}, & \Phi_{r,t_r-1} &= 0, \\ \Delta \Phi_{r,t} &= \Phi_{r,t} - \Phi_{r,t-1}, \\ y_{r,t} &= P_r^B \Delta \Phi_{r,t}, & K_{r,t+1} &= K_{r,t} - y_{r,t}, & K_{r,t_r} &= P_r^B. \end{aligned} \quad (15)$$

Consequently,

$$0 \leq K_{r,t} \leq P_r^B, \quad \sum_{t \in r} y_{r,t} = P_r^B \Phi_{r,T_r} \leq P_r^B. \quad (16)$$

There is no replenishment within an issuance. At renewal, the remaining collateral is returned to the current investors, a new principal P_{r+1}^B is funded, and cumulative impairment resets to zero. A new event can raise $\Phi_{r,t}$, but an unchanged trigger level cannot release the same fraction twice. Payments enter a restricted recovery account. Audit of fraud or corrupted source data can recover an invalid payment without delaying initial restoration liquidity.

The coupon reflects expected loss, tail risk, model uncertainty, and verified controls:

$$\kappa_r = r_r^f + \lambda_E \widehat{\mathbb{E}}_r[\Phi] + \lambda_C \widehat{\text{CVaR}}_{\beta,r}(\Phi) + \lambda_M U_t^{model} - \eta_h g(h_t), \quad (17)$$

where r_r^f is the collateral yield, U_t^{model} measures epistemic uncertainty, and $g(h_t)$ maps audited controls to a spread discount. Bounds prevent a negative spread or an unaffordable quote.

The resilience rebate capitalizes on part of the modeled risk reduction:

$$b_r = \zeta P_r^B \left(\widehat{\mathbb{E}}_r[\Phi \mid h_r^{base}] - \widehat{\mathbb{E}}_r[\Phi \mid h_r] \right)_+, \quad (18)$$

with rebate share ζ . The funds can pay only for enumerated controls, including secure gateways, device certificates, communication diversity, local fallback, reserve energy, and recovery exercises. This restriction distinguishes resilience capital from an unrestricted premium discount.

Investor terminal wealth for issuance r is

$$\mathcal{Y}_{B,r} = K_{r,T_r} + \sum_{t \in r} \left(\kappa_r K_{r,t} + r_r^f K_{r,t} - c_{r,t}^{adm} \right), \quad (19)$$

where $c_{r,t}^{adm}$ covers issuance, audit, and data-oracle costs. Coupon accrues on outstanding, not original, collateral after an impairment. The sponsor's liquidity evolves as

$$W_{t+1} = W_t + \Pi_t^{VPP} - \kappa_r K_{r,t} - u_t^\top c_u - Z_t + y_{r,t} + b_r - c_t^{rec}. \quad (20)$$

Here, Π_t^{VPP} is operating revenue and c_t^{rec} is recovery expenditure. The sponsor affordability cost includes the coupon, eligible hardening expenditure net of the restricted rebate, and administration. This accounting prevents a payout from being treated as a reduction in social loss and prevents repeated release of the same collateral.

5.2. Basis Risk and Agent Objectives

Fast parametric settlement introduces mismatch. For one event, normalized cumulative economic loss is $\tilde{Z} = \min\{1, \sum_t Z_t / Z^{max}\}$. The basis-risk statistic compares that loss with the final cumulative impairment of its issuance and distinguishes underpayment from overpayment:

$$BR = \mathbb{E} \left[\omega_- [\tilde{Z} - \Phi_{r,T_r}]_+ + \omega_+ [\Phi_{r,T_r} - \tilde{Z}]_+ \right], \quad (21)$$

where $\omega_- > \omega_+$ when failure to pay after a severe event is socially more costly than a modest excess payment. This statistic enters both learning and out-of-sample evaluation.

The attacker seeks physical disruption and financial amplification while paying resource and detection costs:

$$R_t^A = v_Z Z_t + v_y y_{r,t} - c_A(\tau_t, \iota_t, \alpha_t, \sigma_t) - c_{det} p_t^{det}. \quad (22)$$

Including $y_{r,t}$ allows the attacker to exploit the trigger if a low-cost action can impair principal without comparable physical loss. The basis-risk penalty and oracle security reduce that opportunity.

The operator reward internalizes operating value, retained loss, hardening, coupon, and resilience performance:

$$R_t^O = \Pi_t^{VPP} - Z_t + y_{r,t} + b_r - \kappa_r K_{r,t} - c_u^\top u_t + v_\rho \rho_t - v_{br} BR_t. \quad (23)$$

The investor reward balances spread income against impairment, uncertainty, and capital volatility:

$$R_t^B = \kappa_r K_{r,t} + r_r^f K_{r,t} - y_{r,t} - c_{r,t}^{adm} - v_M U_t^{model} - v_K [K^{min} - K_{r,t+1}]_+. \quad (24)$$

The regulator evaluates public service, credible triggers, and market continuity:

$$R_t^R = -c_\ell^\top \ell_t - v_s \mathbf{1}\{\rho_t < \underline{\rho}_t\} - v_{br} BR_t - v_w [W^{min} - W_{t+1}]_+. \quad (25)$$

Centralized training accounts for interaction, but each actor retains its own objective. Social cost is reported separately so a transfer is not mistaken for a reduction in harm.

Tail loss for any episode-level random cost L is measured by

$$\text{CVaR}_\beta(L) = \min_{\eta \in \mathbb{R}} \left\{ \eta + \frac{1}{1-\beta} \mathbb{E}[(L - \eta)_+] \right\}. \quad (26)$$

The admissible policy set imposes cumulative constraints:

$$\mathcal{C} = \left\{ \pi : J_{g_j}(\pi) = \mathbb{E}_\pi \sum_t \gamma^t g_{j,t} \leq d_j, j \in \{\text{safe}, \text{imp}, \text{aff}, \text{br}\} \right\}, \quad (27)$$

where $g_{\text{safe},t}$ is the normalized pre-projection engineering violation, $g_{\text{imp}} = \mathbf{1}\{\Phi_{r,T_r} > 0\}$, $g_{\text{aff}} = [C_r^{\text{sponsor}} / \Pi_r^{\text{VPP}} - 0.18]_+$, and $g_{\text{br}} = \text{BR}$. The limits are 1% unsafe proposals, 3% impaired issuances, 18% sponsor burden, and 0.09 weighted basis risk. All four are evaluated over complete episodes. Pointwise electrical and resource constraints remain hard constraints in $\mathcal{F}(s_t)$; they are not traded against reward.

5.3. Risk-Constrained Multi-Agent Policy Updates

Training uses centralized critics and decentralized actors. Bond actions update at renewal boundaries; attack and operator actions update each operational interval. Bounded attacks and fixed contracts are followed progressively by stronger attacks, flexible terms, correlated outages, and oracle noise. This curriculum prevents early catastrophes from driving a degenerate no-coverage policy.

For policy parameters θ and nonnegative multipliers λ , the constrained game uses the Lagrangian

$$\mathcal{L}(\theta, \lambda) = \sum_{i \in \mathcal{N}} \omega_i J_i(\theta) - \sum_j \lambda_j (J_{g_j}(\theta) - d_j), \quad (28)$$

where ω_i scales heterogeneous rewards during critic training without eliminating strategic differences. After each rollout, the multiplier update is

$$\lambda_j \leftarrow \Pi_{[0,50]} \left(\lambda_j + \eta_\lambda [\hat{J}_{g_j} - d_j] \right), \quad \eta_\lambda = 5 \times 10^{-3}. \quad (29)$$

The slower projected update raises a multiplier after a budget breach and relaxes it when the constraint is slack. The cap prevents one early extreme episode from freezing later policy improvement.

The centralized action-value function for agent i is

$$Q_i^\pi(s_t, a_t) = \mathbb{E}_\pi \left[\sum_{k=0}^{T-t} \gamma^k R_{t+k}^i \mid s_t, a_t \right]. \quad (30)$$

Local actors receive $o_t^i = \mathcal{O}^i(s_t)$, while critics receive the global training state. Generalized advantage estimation reduces variance:

$$\hat{A}_t^i = \sum_{l=0}^{T-t-1} (\gamma \lambda_G)^l \left(R_{t+l}^i + \gamma V_i(s_{t+l+1}) - V_i(s_{t+l}) \right). \quad (31)$$

The actor update uses a clipped objective with constraint-adjusted advantage:

$$\max_{\theta_i} \hat{\mathbb{E}}_t \left[\min \left(r_t^i(\theta_i) \hat{A}_{t,\lambda}^i, \text{clip}(r_t^i(\theta_i), 1 - \epsilon, 1 + \epsilon) \hat{A}_{t,\lambda}^i \right) + \tau_H \mathcal{H}(\pi_i) \right], \quad (32)$$

where r_t^i is the new-to-old policy likelihood ratio and $\mathcal{H}$ maintains exploration. Contract actions are parameterized through bounded transforms so that the principal is nonnegative and the attachment point always remains below exhaustion.

Before the electrical simulator applies an operator action, a safety layer projects it onto the state-dependent feasible set:

$$\hat{a}_t^O = \Pi_{\mathcal{F}(s_t)}(a_t^O) = \arg \min_{z \in \mathcal{F}(s_t)} \|z - a_t^O\|_2^2. \quad (33)$$

Projection distance is logged as a learning cost, teaching the actor to avoid infeasible dispatch, reserve, or curtailment. At execution, each actor uses its own observations; privileged critic variables are discarded.

Table 3 reports the full learning configuration. Continuous observations are normalized with training-set running means and variances and clipped to $[-10, 10]$. Binary masks and categorical target indicators are not standardized. Reward components are divided by fixed self-insurance scales. No evaluation observation updates the normalization statistics.

Table 3. Learning, implementation, and model-selection configuration.

Item	Configuration
Actors	Four role-specific multilayer perceptrons; hidden widths 256 and 128; tanh activations; separate bounded heads for continuous actions and categorical heads for targets.
Centralized critics	One critic per agent; global state and joint action as input; three hidden layers of 256 units; ReLU activations; scalar value output.
Initialization	Orthogonal weight initialization; gains $\sqrt{2}$ for hidden layers, 0.01 for actor outputs, and 1.0 for critic outputs; zero biases.
Optimizer and rates	Adam; actor learning rate 3×10^{-4} , critic rate 10^{-3} , dual rate 5×10^{-3} ; no weight decay.
Rollout and minibatch	4096 transitions per rollout; minibatch 256; 10 PPO epochs per rollout; shuffled minibatches.
Training horizon	12,000 episodes per seed; 96 fifteen-minute intervals per episode; 20 training seeds; common random numbers across regimes.
PPO and GAE	Discount 0.99; GAE 0.95; clip width 0.20; value-loss coefficient 0.5; entropy coefficient linearly reduced from 0.010 to 0.002.
Gradient control	Global gradient norm clipped at 0.5; advantages standardized within a rollout; nonfinite updates rejected and the previous checkpoint restored.
Constraint update	Projected rule in Equation (29); multiplier range 0–50; one update per completed rollout.
Seed handling	Training seeds 1001–1020; disjoint evaluation seed IDs 1–20; separate pseudorandom streams for profiles, attacks, policy sampling, and bootstrap resampling.
Checkpoint selection	Lowest validation social cost among checkpoints satisfying all four validation budgets for five consecutive checks; checks every 200 episodes.
Software and hardware	Python 3.12.13; NumPy 2.3.5; SciPy 1.17.0; pandas 2.2.3; Matplotlib 3.10.8; Linux x86_64; CPU-only run on nine Intel Xeon Platinum 8573C vCPUs with 21 GiB RAM.

Table 4 gives the multi-timescale training sequence. The SPV and regulator act only when the renewal mask is active. Operational actions are updated at every interval.

Table 4. Multi-timescale risk-constrained training algorithm.

Step	Operation
1	Initialize four actors, four centralized critics, target-independent normalization statistics, and $\lambda_j = 0$.
2	Draw one operating profile, digital dependency state, renewal schedule, and attack path from the seed-specific training stream.
3	At a renewal boundary, sample bounded principal, coupon, attachment, exhaustion, rebate, service, and disclosure actions. Hold them fixed until the next boundary.

Table 4. *Cont.*

Step	Operation
4	At each operating interval, sample the attack and operator actions from local observations. Project dispatch, storage, curtailment, and recovery onto $\mathcal{F}(s_t)$.
5	Advance electrical, cyber, audit, recovery, trigger, cumulative impairment, collateral, and liquidity states. Store local observations, the global training state, rewards, risk costs, masks, and projection distance.
6	After 4096 transitions, compute GAE for rewards and each risk cost. Update critics and clipped PPO actor objectives for 10 epochs.
7	Update each multiplier with Equation (29). Update running normalization statistics using training data only.
8	Every 200 episodes, evaluate without exploration. Retain a checkpoint only after all four budgets are satisfied for five consecutive checks.
9	Stop after 12,000 episodes. Freeze actors and normalization statistics, discard privileged critic inputs, and evaluate the 20 held-out seed IDs.

6. Results and Discussion

6.1. Comparative Design

The evaluation compares five energy-management and risk-financing regimes under identical operating paths and attack draws. Self-insurance has no bond and requires the operator to absorb all loss. The fixed cyber bond uses constant principal, coupon, attachment, and exhaustion. The EVT bond updates the spread and attachment from a rolling tail-loss fit but does not condition its coupon on verified hardening. The rule-based resilience bond grants a prespecified rebate and control discount. The proposed policy jointly learns contract, defense, network-constrained dispatch, isolation, and recovery actions subject to the risk budgets in Section 5.

Normalized social cost includes interruption, emergency operation, hardening, recovery, administration, and distressed liquidity. Coupons and principal releases are transfers except for issuance and capital-distress costs, so a payout cannot masquerade as lower physical loss. Critical-load interruption is weighted demand not served, and hardening is an audited control score rather than declared spending.

6.2. Portfolio-Level Performance

Table 5 reports both portfolios. Fixed risk transfer reduces liquidity stress but weakens verified hardening because principal protection is not linked to controls. The rule-based design restores part of this incentive, while the proposed contract produces the lowest social cost, interruption, tail loss, and recovery time in both systems. Its ranking remains stable as common-platform exposure increases from VPP-39 to VPP-118.

Recovery time is the elapsed physical time from attack onset until weighted service first returns to at least 90% and remains there for four consecutive intervals. It is not solver time. The VPP-118 value of 11.8 h is longer than 10 h because the larger portfolio has more shared-service dependencies and more assets to reauthenticate. It is nevertheless the best feasible value among the five matched regimes; the closest rule-based bond requires 14.3 h and self-insurance requires 18.7 h. The result is therefore a constrained operating optimum, not evidence that the numerical solver is slow or divergent.

Figure 1 shows a consistent advantage across both portfolios. The learned bond lowers normalized social cost to 0.691 for VPP-39 and 0.704 for VPP-118, while critical-load interruption falls to 5.08% and 6.91%. Tail loss and recovery time follow the same ranking, confirming that the economic gain reflects stronger physical service rather than accounting

transfers. The compact seed-level intervals further indicate stable performance across evaluation runs.

Table 5. Out-of-sample performance across both synthetic portfolios. Values are seed means; lower is better except for verified hardening.

Portfolio	Regime	Social Cost	Critical Load (%)	Loss CVaR	Recovery (h)	Hardening
VPP-39	Self-insurance	1.000	8.82	1.000	13.4	0.421
	Fixed cyber bond	0.925	8.11	0.914	12.7	0.365
	EVT cyber bond	0.858	7.23	0.802	11.6	0.438
	Rule-based bond	0.789	6.37	0.701	10.2	0.552
	Proposed bond	0.691	5.08	0.558	8.4	0.663
VPP-118	Self-insurance	1.000	11.57	1.000	18.7	0.402
	Fixed cyber bond	0.943	10.88	0.941	17.9	0.344
	EVT cyber bond	0.873	9.81	0.832	16.2	0.425
	Rule-based bond	0.801	8.57	0.724	14.3	0.561
	Proposed bond	0.704	6.91	0.581	11.8	0.681

Bold values indicate the best value within each portfolio: the minimum for social cost, critical-load interruption, loss CVaR, and recovery time, and the maximum for verified hardening.

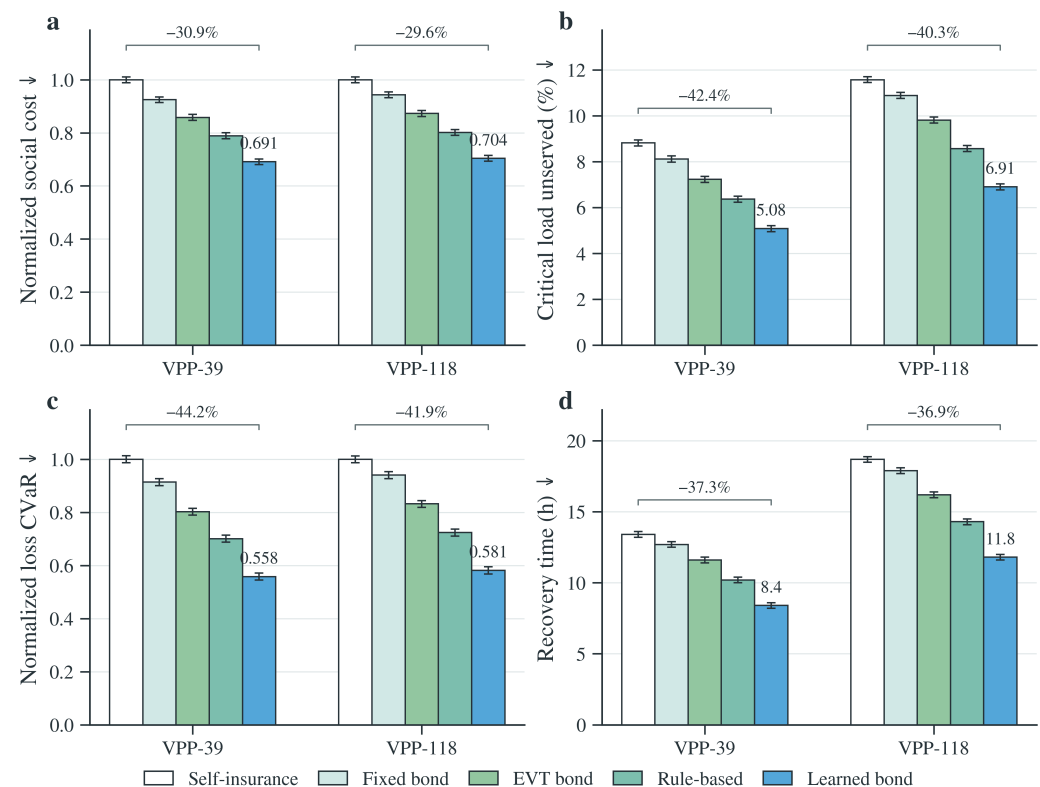


Figure 1. Cross-regime performance on VPP-39 and VPP-118: (a) normalized social cost; (b) critical load unserved; (c) normalized loss CVaR; (d) recovery time. Bars show archived means; error bars are 95% bootstrap intervals over the archived seed records. Downward arrows denote lower-is-better metrics. Brackets show reductions from self-insurance to the proposed bond.

Figure 2 shows that the learned bond retains more service immediately after an attack and reaches 90% functionality in about 9 h for VPP-39 and 12 h for VPP-118. Self-insurance requires approximately 13.5 h and 18.5 h, with the fixed and EVT bonds between these limits. Faster restoration results from coordinated resilience capital, operating reserve, device isolation, and recovery effort, while the larger portfolio remains slower because of stronger common-platform dependence.

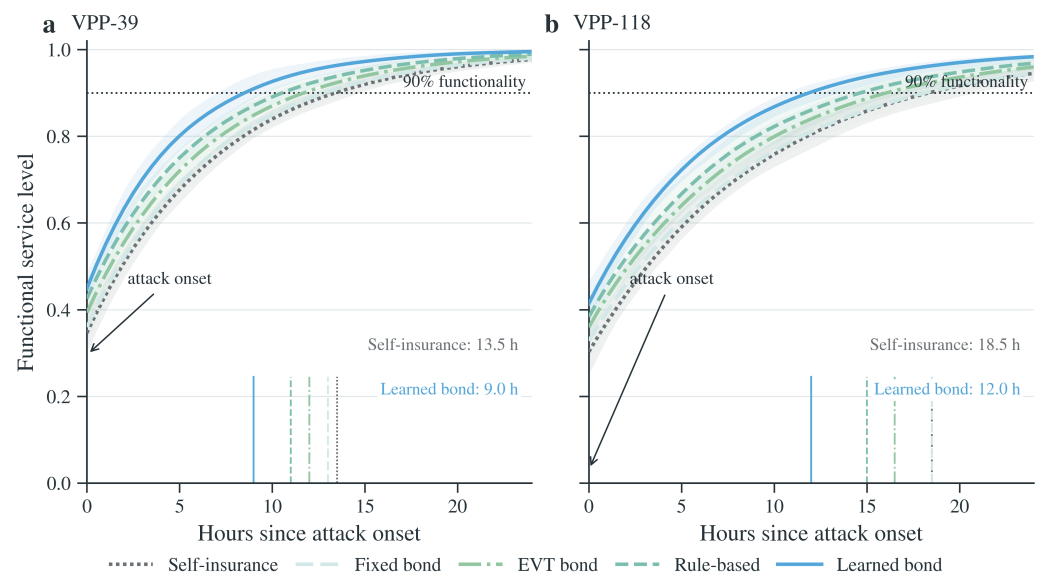


Figure 2. Post-attack functional-service recovery: (a) VPP-39; (b) VPP-118. Curves show archived episode means; shaded bands span the 2.5th–97.5th percentiles across archived episodes. The horizontal dotted line marks 90% functionality; short vertical markers show threshold-crossing times of the mean curves. The arrow marks attack onset.

The learned policy does not always choose the largest bond. In low-stress states, it reduces principal because additional collateral costs more than its expected benefit. Under elevated cloud dependence or weak recovery readiness, it raises the rebate before widening coverage. Financing controls before expanding the risk layer prevents the result from being a mere transfer from investors.

6.3. Trigger Quality, Capital Safety, and Ablation

Mean weighted basis risk is 0.065 for VPP-39 and 0.071 for VPP-118, versus 0.10–0.19 for nonlearned bonds. Joint adaptation of attachment, exhaustion, and trigger weights drives the improvement. The policy emphasizes authenticated controllability under cloud or identity attacks and service loss under load-altering attacks. Constrained weights are disclosed and change only between renewals.

Figure 3 links preventive behavior with trigger quality. The fixed bond reduces verified hardening below self-insurance, whereas the learned bond raises audited scores to 0.663 for VPP-39 and 0.681 for VPP-118. Its weighted basis risk is also the lowest among triggered contracts, at 0.065 and 0.071. The joint improvement shows that restricted resilience capital strengthens verified controls, while calibrated attachment prevents routine noise from being shifted to investors.

Figure 4 shows that trigger values remain close to realized loss, with correlations of approximately 0.87 for VPP-39 and 0.85 for VPP-118. Mean absolute mismatch is about 0.086 and 0.090. Dispersion increases in the sparse upper tail, and VPP-118 exhibits a slightly wider cloud because of stronger common-platform exposure. Conditional medians nevertheless remain near the ideal diagonal through the principal loss range, supporting the attachment calibration.

Principal impairment occurs in the tail rather than in routine noise, and the probability of breaching the investor-loss budget remains below tolerance. Coupons rise with model uncertainty, while the affordability multiplier prevents a quote that eliminates sponsor participation.

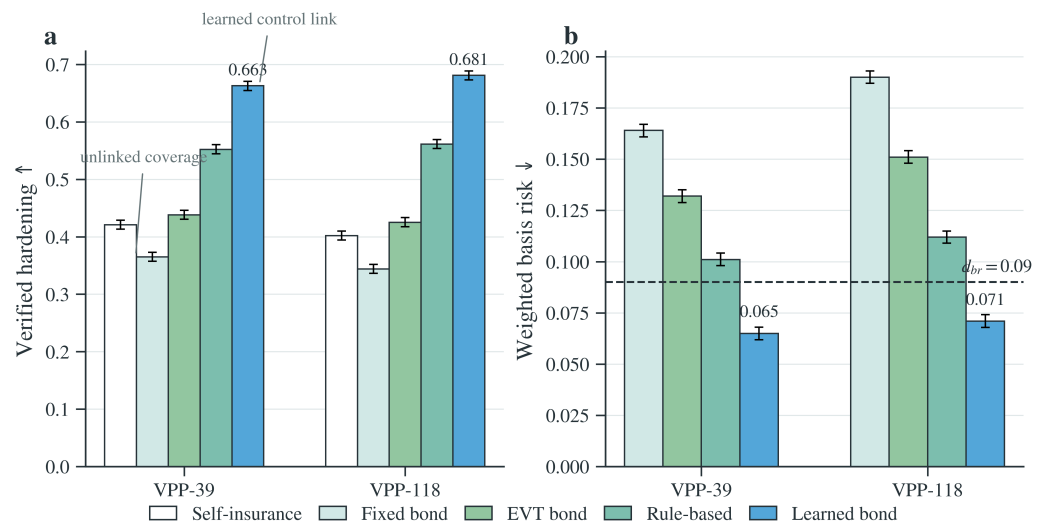


Figure 3. Contract performance on VPP-39 and VPP-118: (a) verified hardening; (b) weighted trigger basis risk. Upward and downward arrows denote higher-is-better and lower-is-better metrics, respectively. Error bars are 95% bootstrap intervals over archived seed records. The dashed line marks the 0.09 basis-risk limit. Self-insurance has no parametric trigger and is omitted from (b).

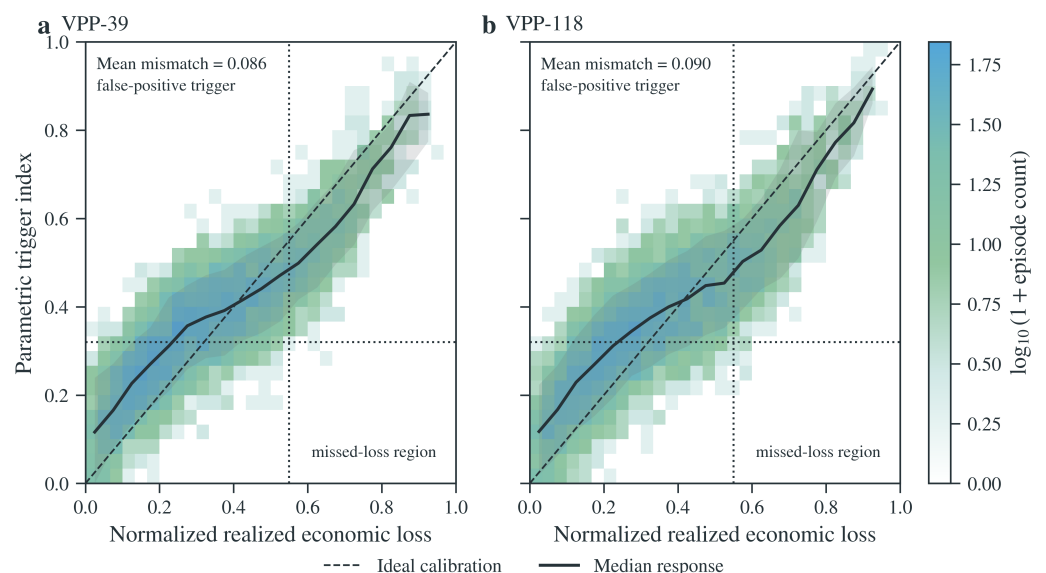


Figure 4. Trigger-loss calibration: (a) VPP-39; (b) VPP-118. Color shows log-transformed episode counts. The solid curve is the conditional median trigger; the shaded band spans its 10th–90th percentiles. The dashed diagonal denotes equal normalized trigger and loss values. Dotted lines mark the displayed trigger threshold of 0.32 and loss threshold of 0.55.

Figure 5 shows smooth cost reduction and progressive constraint satisfaction. Social cost converges near 0.697 for VPP-39 and 0.714 for VPP-118. Basis risk falls below 0.09, investor impairment falls below 0.03, and unsafe proposals drop under 1% without a late rebound. Narrowing confidence bands indicate policy stabilization, while the slightly higher VPP-118 trajectories reflect its larger shared-software footprint and more difficult feasibility projection.

Figure 6 quantifies the cost of tightening basis-risk and investor-impairment safeguards. Across the displayed feasible grid points, normalized social cost ranges from approximately 0.645 to 0.711. The selected limits of 0.09 and 0.03 yield a pooled cost of 0.698. This value is the average of the portfolio-specific means in Table 5, 0.691 and 0.704 after rounding. Figure 5 reports moving training estimates rather than held-out portfolio

means. Its terminal values therefore need not equal either row, but the VPP-39 trajectory converges near the pooled test value. All three displays use the same self-insurance normalization and the same definition of social cost.

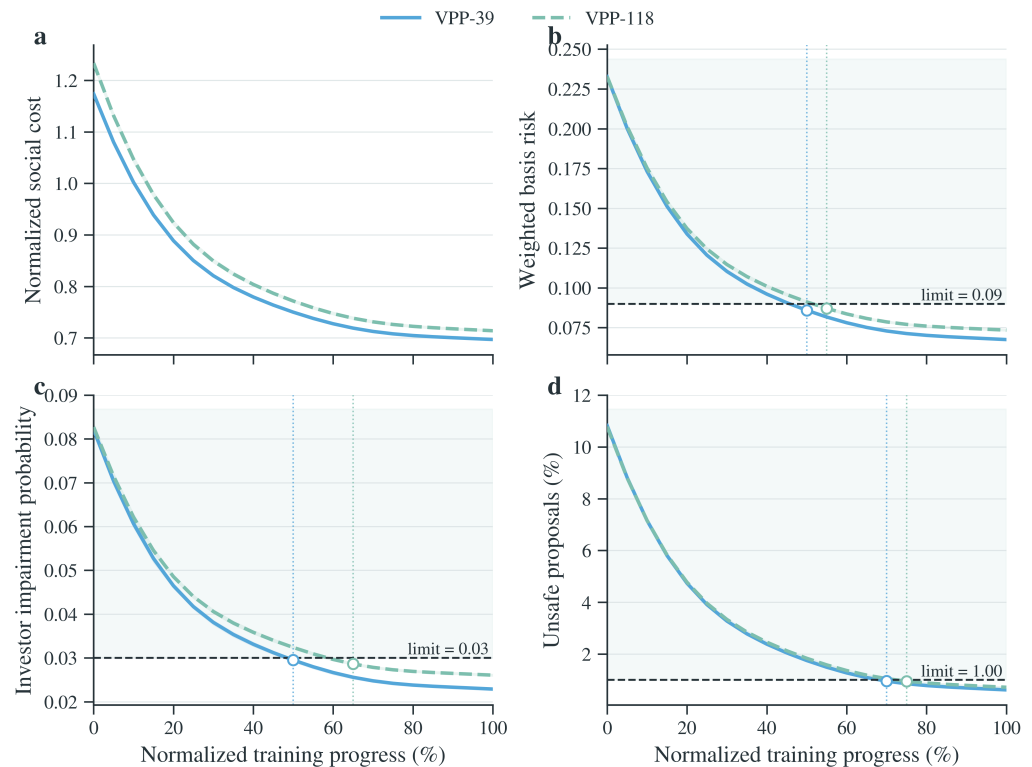


Figure 5. Training trajectories for VPP-39 and VPP-118: (a) normalized social cost; (b) weighted basis risk; (c) investor impairment probability; (d) unsafe proposals. Curves show archived seed means with 95% bootstrap bands. Dashed lines and shaded regions mark risk limits and exceedances. Open markers and vertical dotted lines mark where the upper interval bound remains below each limit thereafter.

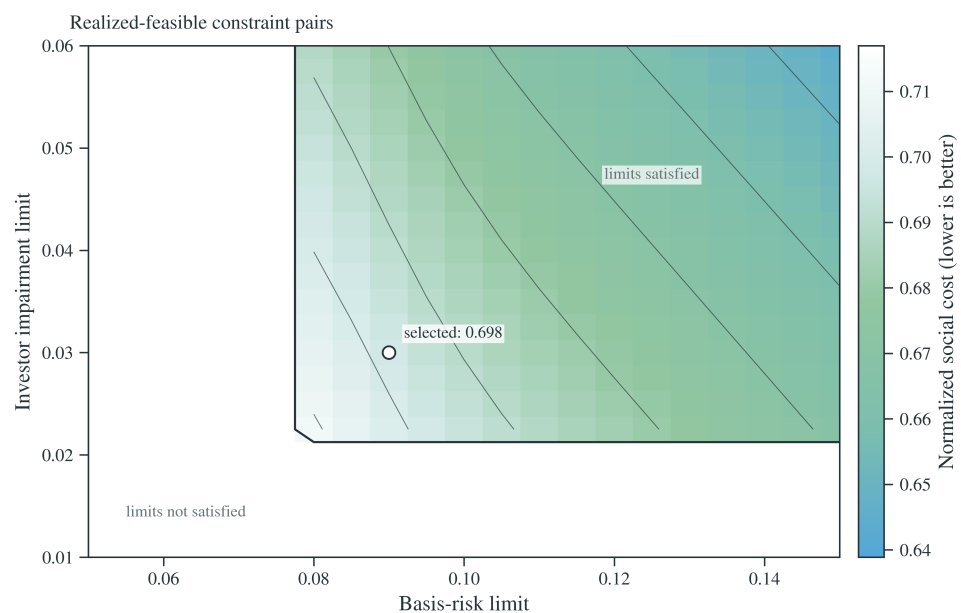


Figure 6. Two-dimensional sensitivity map of normalized social cost over basis-risk and investor-impairment budgets. Color and contours show cost at grid points where the 97.5th percentiles of both archived risk measures satisfy their limits. The dark boundary separates this region from the uncolored region. The open marker denotes the selected limits of 0.09 and 0.03.

Contract trajectories provide an additional economic check. Early renewals allocate more available cash to the restricted rebate while retaining a relatively high attachment point. Once audited hardening and recovery readiness improve, the learned SPV widens principal coverage and reduces the uncertainty component of the coupon. This sequence is coherent because investors are not asked to fund an unchanged exposure at a discounted spread. The operator first converts rebate capital into observable risk reduction; only then does the contract transfer more residual tail risk. During adverse model updates, the sequence partly reverses: attachment rises, principal falls, and the rebate concentrates on controls that reduce common-platform dependence. The behavior also limits gaming. An operator cannot obtain a persistent discount from declared expenditure alone, because audit state and realized functionality jointly determine renewal terms. Similarly, an attacker cannot profit from a trigger-only disturbance without incurring basis-risk and detection penalties.

Figure 7 shows how contract terms respond to stress. Under high stress, the restricted rebate rises from about 0.21 to 0.27 and the principal expands from roughly 0.73 to 1.02 tail-loss units. Under low stress, the rebate declines toward 0.09, and attachment exceeds 0.40. Principal increases only after recovery readiness improves, preserving the sequence of mitigation before broader tail-risk transfer.

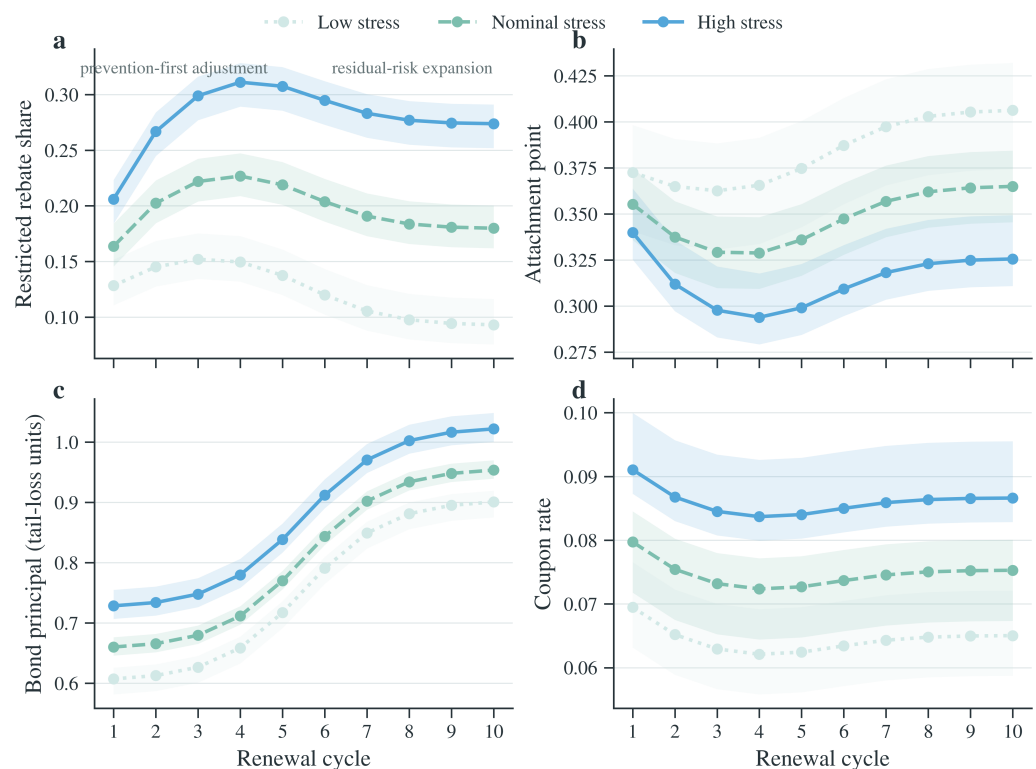


Figure 7. Contract renewal trajectories under low, nominal, and high stress: (a) restricted rebate share; (b) attachment point; (c) collateralized principal in tail-loss units; (d) coupon rate. Curves show archived seed means; shaded bands span the 2.5th–97.5th percentiles across the 20 archived seed records.

Figure 8 maps cloud dependency and recovery readiness into the attachment threshold. Strong readiness and low cloud dependence raise attachment to the 0.45 bound, retaining routine loss with the sponsor. Weak readiness and high dependence lower attachment toward 0.21 for VPP-39 and 0.20 for VPP-118, allowing earlier access to collateral. The smooth surfaces prevent abrupt renewal changes while accounting for portfolio-specific digital accumulation.

Figure 9 concentrates resilience funding where controls remain improvable and verifiable. The rebate is near 0.33 when controls are weak and uncertainty is low, but falls below 0.08 as the verified-control score approaches one. Moderate uncertainty supports targeted verification and redundancy, whereas extreme uncertainty suppresses funding. This turnover prevents an operator from obtaining a larger rebate merely by making control effectiveness harder to verify.

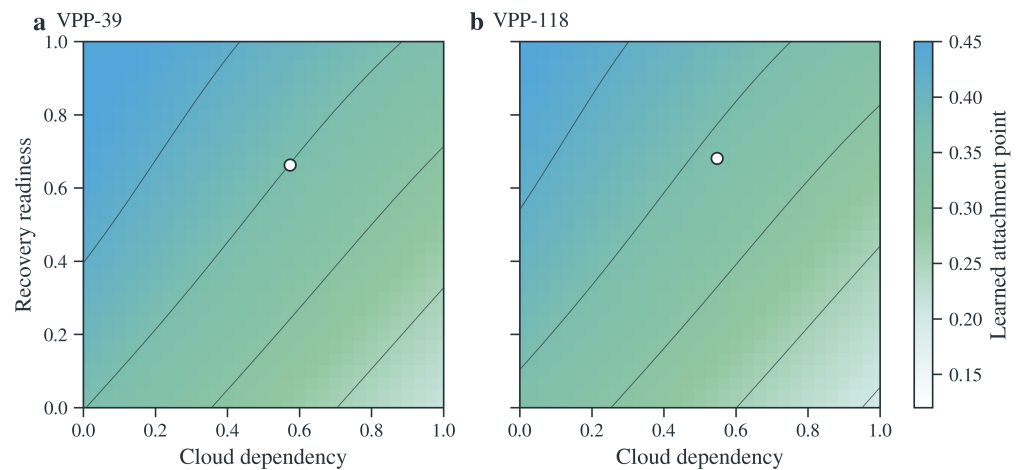


Figure 8. Learned attachment maps: (a) VPP-39; (b) VPP-118. Color and contours show attachment as a function of cloud dependency and recovery readiness. White markers use portfolio-median cloud dependency and the reported verified-hardening score as the plotted readiness coordinate.

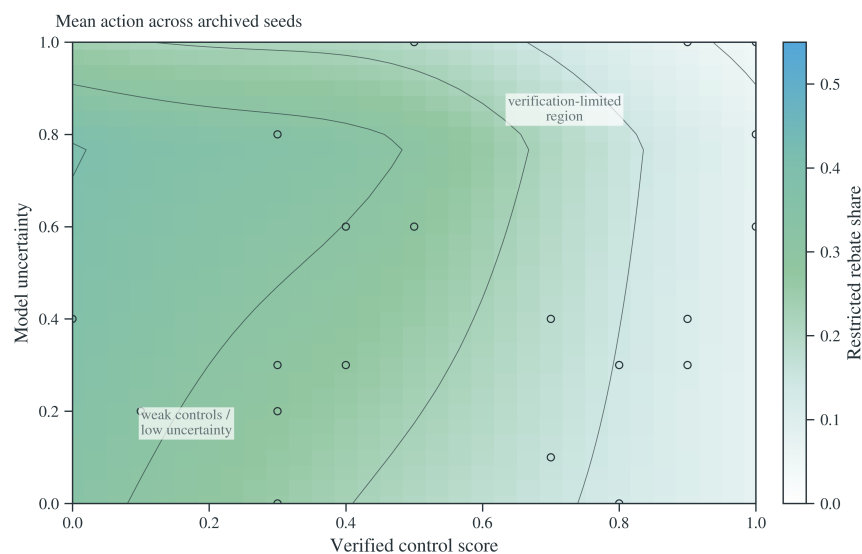


Figure 9. Two-dimensional map of restricted rebate allocation over verified-control score and model uncertainty. Color and contours show the mean rebate share across archived seed records. Open markers sample grid locations whose across-seed standard deviation is at or above the 85th percentile.

Normalized objectives alone cannot establish whether a coupon is investable or affordable. Table 6 therefore converts the synthetic tail-loss units into an explicit annual dollar screen. The conversion uses USD 10 million for VPP-39 and USD 25 million for VPP-118. These scales are scenario assumptions, not market quotes. Expected investor net return equals coupon plus collateral income minus expected principal release and issuance, audit, and administration costs. Sponsor burden equals coupon plus eligible hardening outlay minus the restricted rebate, divided by annual VPP revenue. The resulting net returns are 7.30% and 7.95%, while sponsor burdens remain below the 18% model limit. This estab-

lishes internal cash-flow consistency under the stated assumptions. It does not establish a market-clearing price.

Table 7 separates the financial and control mechanisms. Values are means with 95% bootstrap intervals over 20 held-out seeds. Removing the coupon–control link primarily affects preventive behavior: hardening falls from 0.672 to 0.519, recovery increases from 10.12 to 12.36 h, and tail loss rises from 0.570 to 0.690. This is the part of the gain attributable to performance-linked finance. Removing the safety projection primarily affects engineering feasibility: unsafe proposals rise from 0.41% to 5.89%. The apparent reduction in its unconstrained social cost does not qualify as an improvement because the 1% safety budget is violated. Likewise, removing the basis-risk penalty produces a lower raw cost but a basis risk of 0.131, well above the 0.09 budget. The full policy is the lowest-cost configuration that satisfies every stated budget. The single cooperative actor worsens cost, recovery, hardening, and tail loss, which supports the use of separate roles and observations.

Table 6. Illustrative annual cash-flow screen in absolute monetary units. Amounts are USD million.

Portfolio	Tail Unit	Principal	Coupon	Collateral Income	Expected Payout	Issue/Audit/Admin	Investor Return (%)	Hardening	Rebate	Sponsor Burden (% Revenue)
VPP-39	10.00	8.00	0.624	0.280	0.140	0.180	7.30	0.420	0.230	16.96
VPP-118	25.00	20.00	1.680	0.700	0.420	0.370	7.95	0.980	0.550	16.75

Table 7. Ablation results with 95% bootstrap confidence intervals. Lower is better except for verified hardening.

Configuration	Social Cost	Basis Risk	Unsafe (%)	Loss CVaR	Hardening	Recovery (h)
Full framework	0.697 [0.692, 0.703]	0.068 [0.066, 0.070]	0.41 [0.35, 0.46]	0.570 [0.563, 0.577]	0.672 [0.667, 0.679]	10.12 [10.00, 10.23]
No basis-risk penalty	0.674 [0.668, 0.680]	0.131 [0.129, 0.133]	0.51 [0.44, 0.57]	0.584 [0.577, 0.591]	0.665 [0.659, 0.671]	10.25 [10.13, 10.37]
No coupon–control link	0.753 [0.747, 0.760]	0.081 [0.079, 0.083]	0.60 [0.55, 0.66]	0.690 [0.683, 0.697]	0.519 [0.513, 0.524]	12.36 [12.23, 12.50]
No safety projection	0.680 [0.674, 0.685]	0.075 [0.073, 0.076]	5.89 [5.84, 5.94]	0.600 [0.593, 0.605]	0.658 [0.653, 0.663]	10.62 [10.50, 10.72]
Single cooperative actor	0.792 [0.786, 0.798]	0.096 [0.095, 0.098]	1.70 [1.65, 1.75]	0.725 [0.720, 0.731]	0.539 [0.534, 0.543]	13.05 [12.95, 13.14]

Stress tests identify trigger reliability as the central deployment risk. Telemetry noise makes the trigger and loss model observe different events. Correlated cloud outages raise loss, uncertainty, and spreads. Weak audits are still more damaging because they break the link from rebate to verified control and loss reduction. Data provenance and independent verification are therefore financial infrastructure for the bond.

Figure 10 confirms robustness under telemetry noise, cloud correlation, audit error, and compound stress. Weighted basis risk rises from about 0.066 to 0.108 for the learned bond, compared with 0.133 to 0.240 for the EVT bond. Under compound stress, the learned social cost remains near 0.90, while the alternatives exceed one. Lower slopes across all panels demonstrate stronger resistance to deteriorating deployment conditions.

6.4. Mechanism Interpretation and Robustness Scope

The experiments support three interpretations. Preventive finance precedes wider tail coverage and removes the underinvestment reproduced by transfer-only contracts. The parametric trigger is calibrated to loss and service outcomes rather than to event labels alone. Investor safety and public resilience remain jointly attainable inside a nonempty feasible region bounded by attachment, verification, and affordability constraints.

The one-at-a-time tests in Table 8 expand the stress range beyond telemetry, cloud, and audit quality. Coordinated attacks have the largest effect: increasing the common-mode share from 0.20 to 0.80 raises social cost by 14.7% and recovery by 2.1 h. Higher attack frequency is the next most influential factor. Communication delay and forecast error mainly reduce the time available for corrective dispatch. Higher renewable penetration has

a smaller net effect because additional low-carbon supply partly offsets forecast variability. All rows keep the trained policy fixed, so they measure operating robustness rather than benefits from retraining.

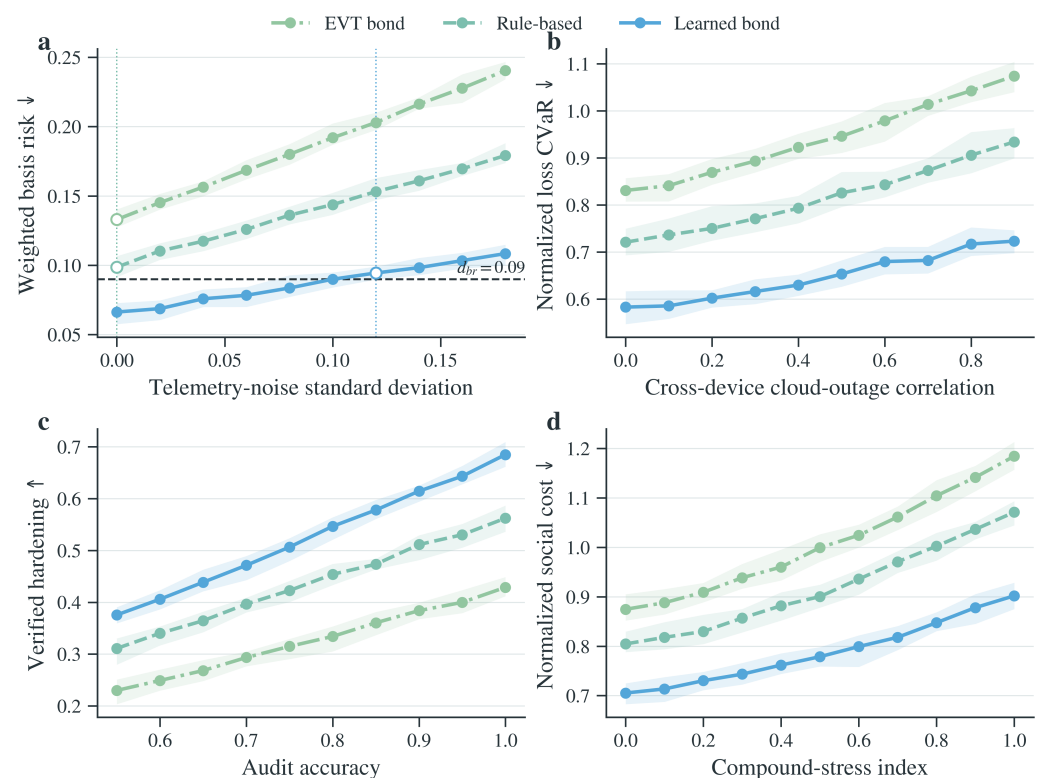


Figure 10. Deployment stress for the EVT, rule-based, and proposed contracts: (a) basis risk versus telemetry noise; (b) loss CVaR versus cloud-outage correlation; (c) verified hardening versus audit accuracy; (d) social cost versus compound stress. Bands span the 2.5th–97.5th percentiles across archived seed records. Arrows denote preferred metric directions. In (a), the dashed line marks the 0.09 limit; open markers and vertical guides identify the first sampled exceedance.

Table 8. Extended one-at-a-time sensitivity analysis for the full policy. Changes are relative to the pooled baseline.

Factor	Baseline to Stress	Social Cost (%)	Recovery (h)	Basis Risk	Unsafe (pp)
Renewable penetration	55% to 85% of available DER energy	+4.8	+0.3	+0.004	+0.1
Communication delay	50 to 250 ms	+7.1	+1.2	+0.008	+0.5
Attack frequency	1.0 to 2.5 event/year-equivalents	+11.9	+1.6	+0.011	+0.4
Attack coordination	0.20 to 0.80 common-mode share	+14.7	+2.1	+0.015	+0.7
Imbalance-price volatility	1.0 to 1.8 times baseline	+3.5	+0.2	+0.002	+0.1
Battery degradation	0.02% to 0.08% capacity per cycle	+4.1	+0.7	+0.003	+0.2
Forecast error	6% to 15% normalized RMSE	+6.3	+0.8	+0.006	+0.3

Table 9 reports a graph-size benchmark. VPP-39 and VPP-118 use the complete study portfolios. The 240- and 300-bus rows map the same five-class synthetic portfolio construction onto WECC-240- and IEEE-300-sized graphs. They test computational scaling; they are not calibrated replications of those systems. State dimension is $8N + 28$, and joint action dimension is $3N + 15$. Parameter growth is linear because the networks have fixed hidden widths. Projection time grows faster because more nodal and branch constraints enter each quadratic projection. The reference training time is a CPU-only estimate obtained from the logged update rate to the selected convergence iteration. Inference remains below 2 ms and

the projection below 20 ms at 300 buses, which is below the 250 ms communication-delay setting used in the stress test.

All numerical values are normalized to the disclosed synthetic evaluation environment, which isolates the causal effects of contract design without relying on confidential claims. The linearized electrical model represents active-power balance, congestion, curtailment, and restoration at the operational decision horizon, while the digital layer represents controllability, common-platform exposure, audit quality, and recovery state. The results establish behavioral validity: removing the control link lowers hardening, removing the basis-risk penalty raises trigger mismatch, and removing the safety layer increases infeasible actions. The two portfolio sizes, shifted attack distribution, independent seeds, and deployment sweeps jointly establish robustness across resource scale, digital accumulation, operating conditions, and measurement quality. These tests show that the framework functions as an intelligent energy-management controller rather than as a financial settlement model alone.

Table 9. Computational scaling and final performance. Time and memory refer to the CPU-only reference environment in Table 3.

Case	Bus	Branch	DER	State	Action	Params	Train (h)	Infer (ms)	Project (ms)	Memory (MB)	Conv. Iter.	Cost	BR	Unsafe (%)
VPP-39	39	47	23	340	132	1,272,236	2.3	0.42	1.7	412	8450	0.691	0.065	0.4
VPP-118	118	146	70	972	369	2,384,872	7.4	0.78	4.9	728	10,280	0.704	0.071	0.6
WECC-240-sized	240	448	144	1948	735	4,103,120	18.6	1.36	13.8	1460	11,340	0.716	0.075	0.8
IEEE-300-sized	300	411	180	2428	915	4,948,160	27.2	1.68	19.2	1880	11,910	0.721	0.078	0.9

7. Data, Parameters, and Reproducibility Protocol

7.1. Synthetic Portfolios and Scenario Construction

The scenario generator constructs operating envelopes for demand, photovoltaic production, communication availability, and market imbalance price. Daily demand and solar profiles use correlated innovations. Battery state, controllable-load availability, communication status, and price update every 15 min. A seed fixes each network and operating path before regimes are compared.

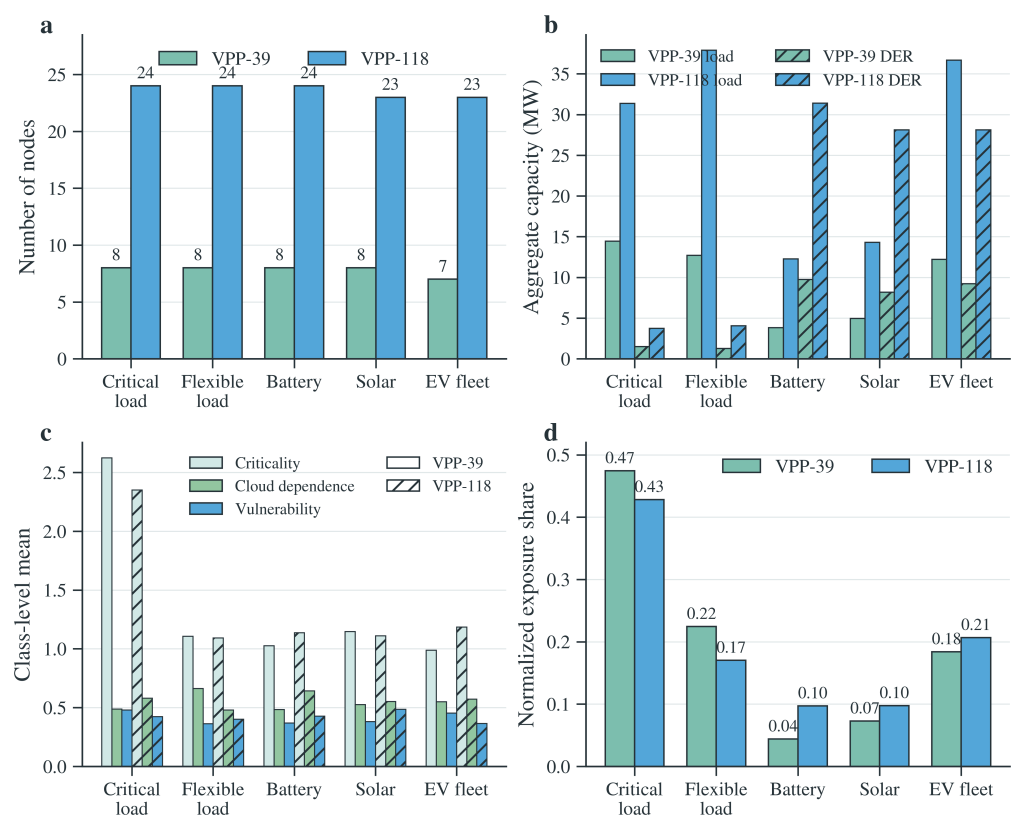
VPP-39 and VPP-118 are transparent synthetic distribution-style portfolios. Their names denote node counts; they are not modifications of the IEEE 39- or 118-bus transmission cases. There are therefore no undisclosed IEEE case changes. The generator first assigns the repeating node classes critical load, flexible load, battery, solar, and EV fleet. It then creates a connected chain and adds nonadjacent chords. Reactance, rating, demand, DER capacity, criticality, cloud dependence, and vulnerability are sampled once from the ranges in Table 10. The archive lists every node and branch, so the exact topology does not depend on a plotting reconstruction.

The per-unit DC screening model follows standard network conventions [47–50]. Class scaling reduces base consumption at battery and solar nodes to 35% of the initial draw. The operating profile multiplies the retained base demand and available solar power. Equipment limits, not the learned policy, set the maximum supply, load, storage, and ramp variables.

Figure 11 summarizes portfolio scale and exposure. Flexible loads and electric vehicles carry the largest aggregate demand, while batteries and photovoltaic units supply most controllable or renewable capacity. Critical loads contribute about 47.4% of normalized exposure in VPP-39 and 42.8% in VPP-118. Mean cloud dependency rises from approximately 0.54 to 0.57, providing a controlled increase in shared digital exposure between the two systems.

Table 10. Electrical-network and portfolio documentation.

Item	VPP-39	VPP-118
Nodes; branches	39; 47	118; 146
Base quantities	10 MVA; 12.47 kV line-to-line	10 MVA; 12.47 kV line-to-line
Topology rule	Chain edges ($i, i + 1$) plus 9 seeded nonadjacent chords	Chain edges ($i, i + 1$) plus 29 seeded nonadjacent chords
Branch data	Reactance 0.015–0.120 p.u.; rating 2.5–8.5 MW	Same bounded ranges with an independent seed
Base demand	48.17 MW total; node range 0.35–2.80 MW before class scaling	132.57 MW total; same node range
DERs	23 battery, solar, or EV nodes; 29.95 MW total capacity	70 battery, solar, or EV nodes; 95.52 MW total capacity
DER limits	Battery/solar/EV rating 0.15–2.40 MW; PV output capped by availability; EV charging bounded by node rating	Same limits
Critical loads	Nodes $1 + 5k, k = 0, \dots, 7$; weight 1.8–3.0	Nodes $1 + 5k, k = 0, \dots, 23$; weight 1.8–3.0
Storage locations	Nodes $3 + 5k, k = 0, \dots, 7$	Nodes $3 + 5k, k = 0, \dots, 23$
Storage limits	Power equals node DER rating; 4 h energy; initial SOC 50%; SOC 10–90%; charge/discharge efficiency 95%	Same rules
Flexible and EV locations	Flexible load: $2 + 5k$; EV fleet: $5 + 5k$, where the node exists	Same rule
Cyber attributes	Cloud dependence 0.20–0.95; vulnerability 0.08–0.72; seed 39,039	Same ranges; seed 118,118

**Figure 11.** Composition and exposure of the synthetic portfolios: (a) node counts by asset class; (b) aggregate base load and DER capacity; (c) class-level mean criticality, cloud dependency, and vulnerability; (d) normalized risk-weighted exposure shares. Exposure is proportional to base load, criticality, cloud dependency, and vulnerability.

Physical lines and shared digital services are modeled separately so that electrical distance does not stand in for cyber dependence. The same node ordering is retained within each portfolio to make clusters comparable across the two layers.

Figure 12 shows sparse physical coupling but dense digital dependence. Nonzero off-diagonal physical entries occupy about 6.3% of VPP-39 and 2.1% of VPP-118, whereas mean off-diagonal digital coupling reaches approximately 0.350 and 0.375. Shared cloud, identity, and communication services therefore connect assets that are electrically distant. This contrast supports separate trigger components for physical consequence and authenticated controllability.

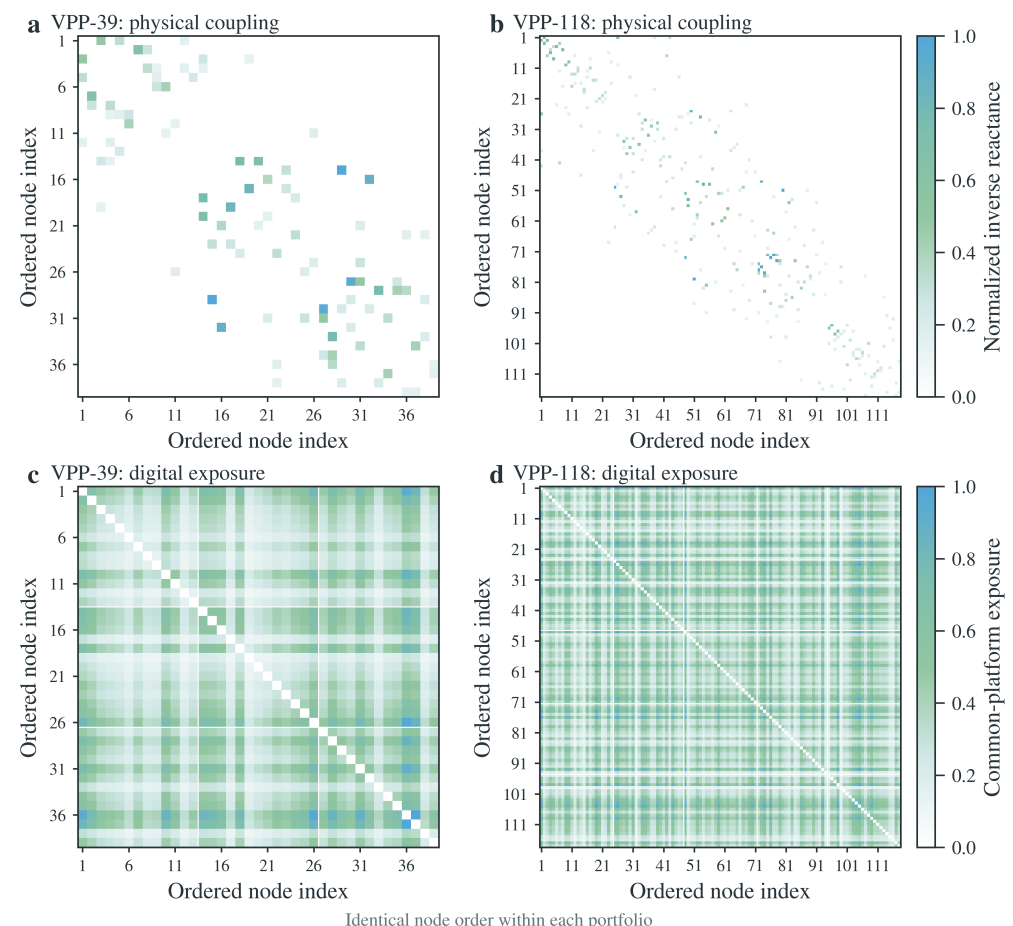


Figure 12. Physical and digital dependency matrices: (a) VPP-39 physical coupling; (b) VPP-118 physical coupling; (c) VPP-39 digital exposure; (d) VPP-118 digital exposure. Physical coupling is normalized inverse reactance on network edges; digital exposure is the pairwise common-platform score derived from cloud dependency and baseline vulnerability.

Attack episodes target device clusters, the VPP application layer, identity infrastructure, or communication brokers. The library includes coordinated load alteration, DER setpoint suppression, false availability reports, command replay, and denial of service. Scenario probabilities are balanced during training and shifted during evaluation so that out-of-sample tests place more mass on correlated application and identity events.

Resilience performance follows the loss-of-function and recovery interpretation used in power-system resilience studies [51,52]. The accompanying archive contains the scenario generator, seed-level results, training and evaluation outputs, table-reconstruction code, and complete figure scripts.

7.2. Training and Evaluation Parameters

Training uses separate seeds and common random numbers across regimes, followed by unseen operating and attack sequences. Rewards are normalized by the self-insurance training baseline, while final metrics are reconstructed in physical and monetary units. Constraint multipliers update more slowly than actors, and the SPV acts only at renewals. Table 11 reports the baseline physical and contract settings; Table 3 gives the learning configuration.

Table 11. Baseline parameterization of the simulation and contract.

Parameter	Symbol	Baseline or Range	Purpose
Attack intensity	α_t	0–1	Scales manipulated load and DER response
Hardening depreciation	δ_h	0.005 per interval	Represents control aging and configuration drift
Bond principal	P_r^B	0.25–1.20 of tail-loss unit	Defines collateralized recovery capacity
Attachment point	a_r	0.12–0.45	Starts principal impairment
Exhaustion point	e_r	0.55–0.95	Completes principal impairment
Rebate share	ζ	0–0.55	Converts modeled risk reduction into eligible capital
CVaR confidence	β	0.95	Defines tail-risk objective
Discount factor	γ	0.99	Values delayed operational and financial effects
PPO clip width	ϵ	0.20	Limits actor update size
GAE parameter	λ_G	0.95	Controls advantage bias–variance tradeoff
Investor impairment limit	d_{imp}	0.03 per issuance	Bounds excessive principal-loss probability
Weighted basis-risk limit	d_{br}	0.09	Restricts trigger–loss mismatch

Each reported mean is computed from 20 independent evaluation seeds with 1500 attack episodes per seed. Bootstrap intervals are retained in the archived data even though the main tables report means for readability. A regime uses the same realization of demand, renewable output, base vulnerability, attack budget, and communication disturbance as every competing regime for a given seed. This paired design reduces variance in differences and prevents an apparently favorable contract from benefiting from easier scenarios.

Reproducibility checks include conservation of nodal balance after corrective curtailment, monotonicity of impairment between attachment and exhaustion, collateral conservation, nonnegative sponsor liquidity after emergency financing, and consistency between trigger components and signed event logs. Automated checks also confirm that tables and figures are regenerated from the archived comma-separated results and that reported aggregates agree with the seed-level outputs.

8. Limitations and Engineering Deployment

8.1. Model and Evidence Limits

The evidence is synthetic. It supports mechanism testing under controlled and reproducible conditions, not a claim of field performance or a market quote. The VPP-39 and VPP-118 labels are sizes, not standard IEEE cases. Network parameters, asset portfolios, attacks, audits, and loss scales are generated from disclosed rules. Utility telemetry, incident records, insurer claims, and investor order books may have different dependence, missingness, and tail behavior. External validation requires a utility-specific digital asset inventory, feeder model, communications map, incident taxonomy, and independent loss calibration.

The DC screening model preserves active-power balance, branch congestion, resource bounds, curtailment, and restoration. It omits reactive power, voltage magnitude, phase imbalance, electrical losses, inverter capability curves, protection coordination, and transient stability. An unbalanced AC or dynamic model would generally shrink the feasible action set. Voltage and reactive-power constraints could require more curtailment or local

support, while losses could change dispatch and recovery cost. Protection and converter dynamics could also invalidate an action that is feasible in the DC screen. These effects may reduce the numerical benefit and increase projection time. They do not predetermine the value of the financing link. A field study should replace $\mathcal{F}(s_t)$ with an unbalanced AC optimal-power-flow or hardware-in-the-loop safety layer and retrain the policy against that layer.

The learning results depend on the stated attack library and observation model. Novel exploits, coordinated deception of several data sources, compromised auditors, and sustained distribution shifts can exceed the training envelope. The risk multipliers limit outcomes under sampled episodes; they do not provide a formal worst-case guarantee. Confidence intervals quantify seed variability, not model uncertainty about a real VPP. The 240- and 300-bus rows test computational size only and should not be read as validation on the WECC or IEEE systems.

The financial model is also deliberately compact. The administration term aggregates issuance and audit cost, but the model does not separately estimate legal fees, placement cost, illiquidity, secondary-market spread, sponsor or counterparty default, taxation, rating-agency treatment, investor capital charges, jurisdiction-specific insurance regulation, or collateral basis risk. Including these items would normally raise the required coupon, reduce affordable principal, or lower the rebate available for hardening. Default and close-out rules would add new state variables and could make some currently feasible contracts unacceptable. The cash-flow table is therefore an internally consistent screen, not evidence of market acceptance [18].

8.2. Deployment Architecture and Safeguards

A practical implementation should keep the learned policy outside the protection layer. The policy proposes bounded setpoints and allocations. Existing protection, inverter controls, and the deterministic feasibility projection retain final authority. Table 12 gives a staged architecture that connects the model to utility systems without allowing the bond logic to bypass engineering controls.

Table 12. Engineering deployment interfaces, timing, and fallback controls.

Layer	Interface	Operational Role	Failure Response
Field and protection	Inverters, relays, meters, battery management, local controllers	Enforce device ratings, protection settings, SOC, and local voltage/frequency behavior.	Reject unsafe setpoints; retain autonomous protection and local droop.
Substation/edge	IEC 61850 [53,54] logical nodes, MMS/GOOSE mapping, protocol gateways	Time-align telemetry, authenticate commands, run the feasibility projection, and cache the last safe schedule.	Isolate a failed gateway; continue the cached schedule; switch to local fallback.
SCADA/EMS/DERMS	Existing state estimator, dispatch, outage management, alarm and historian services	Supply approved measurements and receive projected active-power, reserve, isolation, and recovery actions.	Freeze financial actions; return control to operator-approved dispatch.
Digital twin	Segregated cloud or control-center service	Run forecasting, policy evaluation, stress tests, trigger reconstruction, and audit replay.	No direct protection authority; loss of service cannot disable field controls.
Evidence/oracle	Signed meter records, identity logs, market records, synchronized clocks	Reconstruct the trigger and preserve an auditable event trail for the SPV and regulator.	Use quorum and reconciliation rules; suspend settlement if provenance fails.

Table 12. *Cont.*

Layer	Interface	Operational Role	Failure Response
Human supervision	Operator console, incident commander, SPV and regulatory approval workflow	Approve renewal actions, monitor explanations and constraint margins, and intervene during novel events.	Manual override, emergency operating plan, and post-event review.
Resilience	Redundant edge hardware, diverse communications, offline backups, tested recovery images	Remove single points of failure and preserve minimum service during cloud or link loss.	Fail over by tested priority; shed noncritical load before critical service.
Governance	Role-based access, signed models, change control, independent audit, applicable cyber rules	Separate policy development, deployment, trigger certification, and settlement authority.	Revoke credentials, roll back the model, and block renewal or payment.

The operating path can be staged. Offline replay and digital-twin testing come first. A shadow mode then compares policy proposals with operator decisions without issuing commands. Limited dispatch follows on a small resource subset under conservative projection limits. Wider deployment requires latency tests, red-team exercises, failover drills, operator training, and jurisdiction-specific review. The communication-delay sensitivity shows that performance begins to deteriorate before 250 ms; a site should therefore measure end-to-end telemetry, inference, projection, approval, and actuation delay rather than relying on compute time alone.

9. Conclusions

This study connects VPP cyber defense, network-feasible operation, restoration, and resilience finance in one finite-horizon stochastic game. The bond provides restricted preventive capital and cumulative, non-repeating collateral releases. The controller coordinates role-specific actions while hard engineering limits and four episode-level risk budgets remain explicit.

Across VPP-39 and VPP-118, the full policy achieves normalized social costs of 0.691 and 0.704 and weighted basis risks of 0.065 and 0.071. Recovery takes 8.4 and 11.8 h, compared with 10.2 and 14.3 h for the closest rule-based benchmark. The ablations identify the source of these gains. The coupon–control link raises verified hardening and reduces tail loss, while the feasibility projection controls unsafe actions. The basis-risk penalty and separate agents are also necessary for an admissible solution. These results support the three hypotheses within the disclosed synthetic environment.

The evidence does not establish field performance or market-clearing bond terms. An unbalanced AC network, protection dynamics, new attacks, imperfect audit, issuance frictions, default, taxation, and regulation may narrow the feasible design space. The next engineering step is a utility-specific digital twin with signed SCADA evidence, an AC or hardware-in-the-loop safety layer, shadow-mode operation, and independent financial calibration. The archived source, data, equations, seed-level outputs, and table and figure scripts provide a reproducible starting point for that test.

Author Contributions: Conceptualization, T.Z., B.Y., J.Y., H.T. and A.P.Z.; methodology, T.Z., B.Y. and A.P.Z.; software, formal analysis, investigation, and data curation, T.Z.; validation, T.Z., J.Y. and H.T.; writing—original draft preparation, T.Z.; writing—review and editing, B.Y., J.Y., H.T. and A.P.Z.; supervision, B.Y., H.T. and A.P.Z.; project administration and funding acquisition, B.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Young Talent Program of State Grid Energy Research Institute Co., Ltd., grant number SGN202114010.

Data Availability Statement: The synthetic scenario generator, aggregate experimental results, and table-reconstruction script are included in the accompanying archive. No confidential utility, customer, or insurance-claim data were used.

Conflicts of Interest: Authors Biao Yang and Jingru Yu were employed by State Grid Energy Research Institute Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. This study received funding from the Young Talent Program of State Grid Energy Research Institute Co., Ltd., grant number SGNY202114010. The funder provided financial support for the research. Apart from the individual scholarly contributions of Biao Yang and Jingru Yu reported in the Author Contributions statement, the funder had no involvement in the study design; collection, analysis, or interpretation of data; preparation of the manuscript; or the decision to submit the article for publication.

Appendix A. Literature Comparison

Table A1 summarizes the streams used to position the model. The table keeps the detailed comparison outside the main narrative while showing which elements each study covers.

Table A1. Representative literature and the specific boundary addressed in this study.

Study	Primary System	Main Method	Useful Coverage	Element Outside Scope
Liu et al. [13]	Power-system cyber risk	EVT catastrophe bond	Tail attachment and securitization	Operational hardening and recovery control
Braun et al. [14]	Cyber insurance-linked securities	Risk-transfer analysis	Investor capacity and model uncertainty	VPP network operation
Motlagh et al. [15]	Disaster resilience finance	Review	Mitigation-linked capital	Adaptive cyber operation
Singh et al. [11]	Virtual power plant	Unsupervised autoencoder	Network-attack detection	Financing and dispatch
Diaba et al. [12]	Future energy systems	Cyber-physical review	Attack and countermeasure taxonomy	Contract control
Seghior et al. [22]	Photovoltaic systems	AI and time-frequency diagnosis	Component fault detection	Portfolio cyber finance
Ou and Chen [23]	Photovoltaic forecasting	Adaptive hybrid learning	Volatility-aware forecasts	Attack and collateral states
Ou et al. [24]	Microgrid and microalgae system	Integrated feasibility analysis	Net-zero multi-energy operation	Cyber-risk transfer
Ou et al. [25]	Intelligent VPP	Green power cloud	Cloud-enabled VPP architecture	Financial incentives and basis risk
Chen et al. [26]	Virtual power plant	Revenue and communication co-optimization	Forecasting and time-segmented access	Adaptive attacks and bond settlement
Eling and Woods [18]	Cyber risk and insurance	Gap analysis	Insurance protection and evidence gaps	Grid-constrained control
Gela et al. [27]	Multiple VPPs	Stochastic Stackelberg method	Joint wind, load, and price uncertainty	Cyber resilience finance
Kong et al. [28]	VPP in spot markets	Coordinated decisions	Multi-market participation	Adaptive cyber recovery
Yan et al. [30]	Distribution-network VPP	Bi-level pricing and mixed game	P2P sharing and strategic operation	Collateralized cyber tail risk
This study	Cyber-physical VPP	Constrained multi-agent control	Dispatch, attacks, verified hardening, recovery, cumulative settlement, cash flow, and basis risk	Field and market validation

References

1. Pudjianto, D.; Ramsay, C.; Strbac, G. Virtual power plant and system integration of distributed energy resources. *IET Renew. Power Gener.* **2007**, *1*, 10–16. [\[CrossRef\]](#)
2. Raeispour, M.; Yan, S.; Meegahapola, L.G.; Yu, X. Cyber-physical security of virtual power plants: A survey. *IEEE Open J. Ind. Electron. Soc.* **2026**, *7*, 416–441. [\[CrossRef\]](#)
3. Weng, Y.; Yang, K.; Liu, Z.; He, W.; Tang, X. Lgvlm-miot: A lightweight generative visual-language model for multilingual IoT applications. *IEEE Internet Things J.* **2025**, *12*, 13311–13326. [\[CrossRef\]](#)
4. Liu, Y.; Ning, P.; Reiter, M.K. False data injection attacks against state estimation in electric power grids. In Proceedings of the 16th ACM Conference on Computer and Communications Security, Chicago, IL, USA, 9–13 November 2009; pp. 21–32. [\[CrossRef\]](#)
5. Wang, J.; Deng, J.; Rashwan, T.L.; Qu, G.; Ren, S.; Liu, D. Oxygen-lean thermokinetics and early-warning of Laojunmiao non-caking coal in Xinjiang coalfield. *Int. J. Min. Sci. Technol.* **2026**, *in press*.
6. Li, P.; Yan, X.; Qian, T.; Shen, Y.; Alhazmi, M. Biodegradable transient data centers: A multi-objective optimization framework for sustainable computing in extreme environments. *IEEE Trans. Ind. Appl.* **2026**, *in press*.
7. Lai, C.H.; Wu, T.E.; Wang, C.C. Enhancing information security in smart manufacturing through least significant bit steganography in engineering drawings. *J. Comput. Inf. Sci. Eng.* **2025**, *25*, 091006. [\[CrossRef\]](#)
8. Zhang, Z.; Xu, C.; Ma, C. Efficient multi-designated receiver authenticated broadcast encryption for group messaging. *IEEE Trans. Netw.* **2026**, *34*, 3449–3459. [\[CrossRef\]](#)
9. Kong, L.; Jiang, J.; Xu, W.; Wu, L. Backdoor rationalization. In Proceedings of the AAAI Conference on Artificial Intelligence, Singapore, 20–27 January 2026; Volume 40, pp. 31447–31455.
10. Gao, M.; Zhou, S.; Gu, W.; Fan, J.; Guan, A.; Zhu, H.; Wei, L.; Hu, Z. Enhancing distribution system state estimation under limited measurements: Leveraging large language model and multimodal information. *CSEE J. Power Energy Syst.* **2025**, *12*, 622–631. [\[CrossRef\]](#)
11. Singh, K.N.; Goswami, A.K.; Chudhury, N.B.D.; Shuaibu, H.A.; Ustun, T.S. Enhancing cybersecurity in virtual power plants by detecting network based cyber attacks using an unsupervised autoencoder approach. *Sci. Rep.* **2025**, *15*, 32374. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Diaba, S.Y.; Shafie-khah, M.; Elmusrati, M. Cyber-physical attack and the future energy systems: A review. *Energy Rep.* **2024**, *12*, 2914–2932. [\[CrossRef\]](#)
13. Liu, Z.; Wei, W.; Wang, L. An extreme value theory-based catastrophe bond design for cyber risk management of power systems. *IEEE Trans. Smart Grid* **2022**, *13*, 1516–1528. [\[CrossRef\]](#)
14. Braun, A.; Eling, M.; Jaenicke, C. Cyber insurance-linked securities. *ASTIN Bull.* **2023**, *53*, 684–705. [\[CrossRef\]](#)
15. Motlagh, F.; Hamideh, S.; Gallagher, M.; Yan, G.; van de Lindt, J.W. Bonds for disaster resilience: A review of literature and practice. *Int. J. Disaster Risk Reduct.* **2024**, *104*, 104318. [\[CrossRef\]](#)
16. Pagano, A.; Romagnoli, F.; Vannucci, E. Quantitative and financial aspects of resilience bonds in recursive insurance contracts. *Environ. Clim. Technol.* **2020**, *24*, 387–402. [\[CrossRef\]](#)
17. Guo, Y.; Chen, R.; Allaoui, H.; Choudhary, A.; Li, W. A dual attention graph neural network framework for sustainable supply chain optimization under energy performance contracting. *Transp. Res. Part E Logist. Transp. Rev.* **2026**, *208*, 104702. [\[CrossRef\]](#)
18. Eling, M.; Woods, D.W. Cyber risk and cyber insurance: Closing the gaps. *Geneva Pap. Risk Insur. Issues Pract.* **2026**, advance online publication. [\[CrossRef\]](#)
19. Guan, A.; Zhou, S.; Gu, W.; Liu, H.; Gao, M.; Zhao, B.; Lv, H.; Chen, J. Region-driven two-layer refined scheduling for multi-stack-integrated alkaline electrolyzer in wind-hydrogen system. *IEEE Trans. Smart Grid* **2026**, *17*, 3845–3857. [\[CrossRef\]](#)
20. Zhang, Z.; Xu, C.; Jiang, C. Privacy-preserving electricity trading matching based on blockchain in smart grid. *Expert Syst. Appl.* **2025**, *303*, 130682. [\[CrossRef\]](#)
21. Wang, L.; Yuan, M.; Song, Y.; Zhao, N.; Zhang, Z.; Chen, Q. Research on complex mountain wind power spectrum modeling based on frequency coupling mechanism. *Renew. Energy* **2026**, *273*, 126079. [\[CrossRef\]](#)
22. Seghiour, A.; Bendjedou, Y.; Mostefaoui, I.M.; Chouder, A.; Alharbi, H.; Bin Humayd, A.S.; Wondimeneh, A.W.; Babqi, A. Fault detection and diagnosis in photovoltaic systems using artificial intelligence and time-frequency analysis. *Sci. Rep.* **2026**, *16*, 10056. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Ou, T.-C.; Chen, Y.-T. Volatility-aware adaptive hybrid deep learning for short-term photovoltaic power forecasting. *Eng. Appl. Artif. Intell.* **2026**, 115319. [\[CrossRef\]](#)
24. Ou, T.-C.; Luo, S.-Y.; Chen, Y.-T. Performance and feasibility of integrated microgrid and microalgae hybrid systems for net-zero energy solutions. *Energy Convers. Manag.* **2026**, *356*, 121340. [\[CrossRef\]](#)
25. Ou, T.-C.; Tieng, H.; Tsai, T.-H.; Li, Y.-Y.; Hung, M.-H.; Cheng, F.-T. Design of green power clouds for intelligent virtual power plants. *IEEE Trans. Autom. Sci. Eng.* **2025**, *22*, 18051–18062. [\[CrossRef\]](#)
26. Chen, Z.; Chen, Y.; Tang, Y. Co-optimization of revenue and communication for virtual power plants via renewable energy forecasting and time-segmented access. *Appl. Sci.* **2026**, *16*, 8682. [\[CrossRef\]](#)
27. Gela, T.T.; Mohandes, M.; Al-Shaikhi, A.; Al-Ismail, F.S.; Gebriel, A.M.; Zalhaf, A.S. Metamodel-assisted stochastic Stackelberg coordination of multiple VPPs under joint wind/load/price uncertainty. *Sustain. Energy Grids Netw.* **2026**, *47*, 102491. [\[CrossRef\]](#)

28. Kong, X.; Chen, Y.; Zhang, X.; Liu, G.; Zhang, P.; Zhang, D. Coordinated decision-making method for VPP participating in multiple power spot markets. *Energy* **2026**, *362*, 142183. [\[CrossRef\]](#)
29. Wang, X.; Zhang, R.; Wu, L.Y. Performance evaluation of PBFT-based blockchain via a stage-informed queueing model. *J. King Saud Univ. Comput. Inf. Sci.* **2026**, *38*, 507. [\[CrossRef\]](#)
30. Yan, X.; Gao, C.; Liu, Z.; Song, M.; Francois, B. Bi-level dynamic pricing and P2P energy sharing in a distribution network VPP: Mixed game. *Sustain. Energy Grids Netw.* **2026**, *47*, 102468. [\[CrossRef\]](#)
31. Rockafellar, R.T.; Uryasev, S. Optimization of conditional value-at-risk. *J. Risk* **2000**, *2*, 21–41. [\[CrossRef\]](#)
32. Altman, E. *Constrained Markov Decision Processes*; Chapman and Hall/CRC: Boca Raton, FL, USA, 1999.
33. Li, P.; Hu, Z.; Shen, Y.; Cheng, X.; Alhazmi, M. Short-term electricity load forecasting based on large language models and weighted external factor optimization. *Sustain. Energy Technol. Assess.* **2025**, *82*, 104449. [\[CrossRef\]](#)
34. Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O. Proximal policy optimization algorithms. *arXiv* **2017**, arXiv:1707.06347.
35. Achiam, J.; Held, D.; Tamar, A.; Abbeel, P. Constrained policy optimization. In Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017; pp. 22–31.
36. Soltan, S.; Mittal, P.; Poor, H.V. BlackIoT: IoT botnet of high wattage devices can disrupt the power grid. In Proceedings of the 27th USENIX Security Symposium, Baltimore, MD, USA, 15–17 August 2018; pp. 15–32.
37. Böhme, R.; Schwartz, G. Modeling cyber-insurance: Towards a unifying framework. In Proceedings of the Workshop on the Economics of Information Security, Cambridge, MA, USA, 7–8 June 2010.
38. Li, P.; Gu, C.; Cheng, X.; Li, J.; Alhazmi, M. Integrated energy-water systems for community-level flexibility: A hybrid deep Q-network and multi-objective optimization framework. *Energy Rep.* **2025**, *13*, 4813–4826. [\[CrossRef\]](#)
39. Marotta, A.; Martinelli, F.; Nanni, S.; Orlando, A.; Yautsiukhin, A. Cyber-insurance survey. *Comput. Sci. Rev.* **2017**, *24*, 35–61. [\[CrossRef\]](#)
40. Cremer, F.; Sheehan, B.; Mullins, M.; Fortmann, M.; Materne, S.; Murphy, F. Enhancing cyber insurance strategies: Exploring reinsurance and alternative risk transfer approaches. *J. Cybersecur.* **2024**, *10*, tyae027. [\[CrossRef\]](#)
41. Shapley, L.S. Stochastic games. *Proc. Natl. Acad. Sci. USA* **1953**, *39*, 1095–1100. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Littman, M.L. Markov games as a framework for multi-agent reinforcement learning. In Proceedings of the Eleventh International Conference on Machine Learning, New Brunswick, NJ, USA, 10–13 July 1994; pp. 157–163.
43. Yu, C.; Velu, A.; Vinitzky, E.; Gao, J.; Wang, Y.; Bayen, A.; Wu, Y. The surprising effectiveness of PPO in cooperative multi-agent games. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 24611–24624. [\[CrossRef\]](#)
44. Biagioni, D.; Zhang, X.; Wald, D.; Vaidhyanathan, D.; Chintala, R.; King, J.; Zamzam, A.S. PowerGridworld: A framework for multi-agent reinforcement learning in power systems. In Proceedings of the Thirteenth ACM International Conference on Future Energy Systems, Virtual Event, 28 June–1 July 2022; pp. 565–570. [\[CrossRef\]](#)
45. Li, P.; Shen, Y.; Shang, Y.; Alhazmi, M. Innovative distribution network design using GAN-based distributionally robust optimization for DG planning. *IET Gener. Transm. Distrib.* **2025**, *19*, e13350. [\[CrossRef\]](#)
46. Vu, L.; Vu, T.; Vu, T.L.; Srivastava, A. Multi-agent deep reinforcement learning for distributed load restoration. *arXiv* **2023**, arXiv:2306.14018.
47. Zimmerman, R.D.; Murillo-Sánchez, C.E.; Thomas, R.J. MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Trans. Power Syst.* **2011**, *26*, 12–19. [\[CrossRef\]](#)
48. Farivar, M.; Low, S.H. Branch flow model: Relaxations and convexification—Part I. *IEEE Trans. Power Syst.* **2013**, *28*, 2554–2564. [\[CrossRef\]](#)
49. Kundur, P. *Power System Stability and Control*; McGraw-Hill: New York, NY, USA, 1994.
50. Baran, M.E.; Wu, F.F. Network reconfiguration in distribution systems for loss reduction and load balancing. *IEEE Trans. Power Deliv.* **1989**, *4*, 1401–1407. [\[CrossRef\]](#)
51. Panteli, M.; Mancarella, P. Influence of extreme weather and climate change on the resilience of power systems: Impacts and possible mitigation strategies. *Electr. Power Syst. Res.* **2015**, *127*, 259–270. [\[CrossRef\]](#)
52. Wang, Y.; Chen, C.; Wang, J.; Baldick, R. Research on resilience of power systems under natural disasters—A review. *IEEE Trans. Power Syst.* **2016**, *31*, 1604–1613. [\[CrossRef\]](#)
53. IEC 61850-7-4:2010+AMD1:2020 CSV; Communication Networks and Systems for Power Utility Automation—Part 7-4: Basic Communication Structure—Compatible Logical Node Classes and Data Object Classes. International Electrotechnical Commission: Geneva, Switzerland, 2020.
54. IEC 61850-8-1:2011+AMD1:2020 CSV; Communication Networks and Systems for Power Utility Automation—Part 8-1: Specific Communication Service Mapping (SCSM)—Mappings to MMS (ISO 9506-1 and ISO 9506-2) and to ISO/IEC 8802-3. International Electrotechnical Commission: Geneva, Switzerland, 2020.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.