processes

MDPI

Article

Data-Driven Risk-Aware Approximate Dynamic Programming Algorithm for Resilient Power System Operation Under High Renewable Uncertainty

Zike Guo [1], Peng Yang [1], Xue Du [1], Wanmei Zhao [1], Jiehua Lu [1], Siliang Liu [2,3,*] and Yingqi Yi [2,3]

[1] Foshan Power Supply Bureau of Guangdong Power Grid Co., Ltd., Foshan 528000, China; guozike@fs.gd.csg.cn (Z.G.); yangpeng@fs.gd.csg.cn (P.Y.); duxue@fs.gd.csg.cn (X.D.); zhaowanmei@fs.gd.csg.cn (W.Z.); lujiehua@fs.gd.csg.cn (J.L.)
[2] Guangzhou Power Electrical Technology Co., Ltd., Guangzhou 510700, China; yi_yingqi@foxmail.com
[3] School of Electric Power, South China University of Technology, Guangzhou 510641, China
* Correspondence: lsliang_ps@sina.com

Abstract

The accelerating integration of renewable energy sources into modern power grids has created unprecedented operational challenges, with significant system cost volatility under extreme uncertainty events. To address this challenge, this paper presents a risk-aware stochastic approximate dynamic programming (SADP) algorithm based on machine learning and parallel computing architectures. The algorithm learns optimal coordination strategies for source-grid-load-storage resources while explicitly quantifying and mitigating tail risk events that conventional approaches overlook. First, a risk-averse stochastic optimization model is constructed, which captures the complex interdependencies between renewable generation uncertainty, demand variability, and flexible resource coordination through second-order cone programming formulations. This model integrates the GlueVaR (Glued Value-at-Risk) metric, enabling simultaneous optimization across multiple risk horizons with adjustable conservatism parameters. Second, to solve the established model efficiently, an SADP algorithm based on risk-averse approximate value functions (RAVFs) is proposed, in which the training process of the RAVFs employs machine learning principles to directly encode risk preferences into operational decisions. By integrating GlueVaR into offline training across 5000 probabilistically weighted scenarios, the algorithm discovers emergent coordination patterns between distributed resources, which are rarely identified by human operators. Third, a large-scale parallel computing architecture is implemented for the SADP algorithm. This architecture decomposes the multi-period optimization problem into single-period coordinated sub-problems. During offline training, parallel computing of a series of single-period sub-problems can be performed across all probabilistic scenarios, significantly reducing training time. Extensive validation on both the modified IEEE 33-bus and 69-bus systems with integrated wind turbines, photovoltaic plants, energy storage systems, and demand response capabilities demonstrates remarkable performance improvements. Convergence analysis reveals that the AVFs stabilize within 30 training iterations, achieving sub-160 s solution times in online application even for complex networks with heterogeneous resources. By enabling real-time risk-aware decision-making under severe uncertainty, the proposed method provides grid operators with actionable strategies that balance economic efficiency and operational resilience.

Keywords: new power system; data-driven; source-load uncertainty; risk-aware training; stochastic dynamic programming

Check for updates

Academic Editors: Lefeng Cheng, Xiaoshun Zhang and Huaizhi Wang

Received: 3 November 2025
Revised: 11 December 2025
Accepted: 14 April 2026
Published: 5 July 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

Recently, the renewable energy sources, particularly wind and photovoltaic (PV) systems, have seen significant growth and are increasingly connected to the power systems [1–3]. However, as the proportion of wind power and PV access continues to grow, the forecast for renewable energy output is becoming increasingly uncertain, which presents great challenges to the secure and cost-effective operation of power systems [4]. Addressing the negative impacts of uncertainties in renewable energy output and load, while minimizing the risk of elevated operation costs, is essential for ensuring the secure and cost-effective operation of the power systems.

Much literature has studied the problem of optimal dispatch of power systems under uncertainty. Plytaria et al. [5] proposed an energy dispatch model of a microgrid with solar and stationary battery systems, based on the stochastic optimization (SO) method. Zhang et al. [6] developed a multi-stage SO model, offering greater flexibility compared to the traditional two-stage approach, which can better handle the uncertainties associated with renewable energy outputs. Based on the quasi-Monte Carlo simulation technique to generate wind farm output scenarios, Chen et al. [7] proposed an SO model for economic dispatch that incorporates the trading of flexible ramping services. Nguyen et al. [8] used a robust optimization (RO) model for energy management of energy storage systems (ESSs) and renewable energies in DC microgrids, and a convex reformulation to deal with the nonlinear non-convex equations. Chen and Wei [9] proposed a novel two-stage RO model for power systems to reduce operational risks based on the decision-related uncertainty sets.

The aforementioned optimal dispatch methods considering uncertainties can be classified into two categories: RO and SO methods. However, the RO method ignores the random variables' probability distribution, potentially leading to overly conservative decisions [10]. And the SO method considers the probability distribution information and minimizes the expected operation costs but ignores the extreme risk scenarios in the tails of the probability distributions. The extreme risk scenarios often have a low probability of occurrence, but if they do occur, they will result in high operating costs for the grid [11]. To mitigate the shortcomings of the risk-neutral SO decisions described above, risk metrics have been introduced into the SO method with the aim of achieving more economic and secure decisions by avoiding probabilistic tail risks in advance.

Daneshvar et al. [12] considered the risk under uncertainties during system operation by means of a robust function based on information gap decision-making and obtained the risk-averse strategies. Li et al. [13] and Farzan et al. [14] proposed a risk-averse SO method for multi-energy grids based on a scenario approach, and introduced a conditional value at risk (CVaR) metric for quantifying tail risks within a specific probability distribution, thereby facilitating risk-averse decisions. Recently, Belles-Sampera et al. [15] proposed a new risk metric in finance, named GlueVaR (Glued Value-at-Risk), which has more flexible parameter combinations than CVaR, and is able to take into account multiple risk needs of decision makers [16]. Furthermore, the stochastic approximate dynamic programming (SADP) algorithm has been widely used in SO problems to address the problem of low solution efficiency, which can transform a complex multi-temporal model into a series of single-period models in the form of recursive solution [17]. Lin et al. [18] developed an SO model for an islanded microgrid considering PV ancillary services and solved the model by the SADP algorithm. Lin et al. [19] used an improved SADP algorithm with the segmented linear approximate value functions (AVFs) to solve the SO model of power system economic dispatch considering pumped storage plants. Das et al. [20] proposed an SADP algorithm based on strategy function accelerated solution and applied it to the SO dispatch problem. Therefore, this paper focuses on how to design an improved SADP algorithm that can efficiently formulate risk-averse strategies based on the new risk metric.

The major contributions can be summarized as follows:

(1) A risk-averse stochastic optimization model is constructed, which introduces the advanced financial risk measure GlueVaR into power system dispatch. This model employs second-order cone programming formulations to capture complex interdependencies and enables simultaneous optimization across multiple risk horizons with adjustable conservatism parameters, thereby explicitly quantifying and mitigating tail risks often overlooked by conventional methods.

(2) An SADP algorithm based on risk-averse approximate value functions (RAVFs) is proposed to solve the established model efficiently, in which the training process of the RAVFs employs machine learning principles to directly encode risk preferences into operational decisions by integrating GlueVaR metric.

(3) A parallel computing architecture is embedded within the SADP algorithm framework, decomposing the multi-period optimization problem into coordinated single-period sub-problems. This design enables the parallel computation of all probabilistic scenarios during offline training, drastically reducing the training time and ensuring solution efficiency for online risk-aware decision-making.

The rest of this paper is structured as follows. Section 2 introduces the risk-averse SO model of new power system. Section 3 introduces the improved SADP algorithm. Section 4 presents the case studies in the power system. Section 5 presents the detailed discussion and further study. Section 6 presents the conclusions.

2. Risk-Averse Stochastic Optimization Model for the New Power Systems

This paper enhances the model by incorporating the risk metric, creating a risk-averse SO model that balances economic efficiency with risk management. The detailed model is introduced as follows.

2.1. Objective Function

The traditional SO model [6] is risk-neutral, focusing solely on minimizing the expected operating cost. Meanwhile, the proposed risk-averse SO model minimizes a weighted risk metric cost. By explicitly embedding a quantifiable risk measure into the objective, our model facilitates a tunable trade-off between economic efficiency and operational resilience under extreme events. This weighted risk metric cost comprises a weighted sum of two components: the expected value and the risk metric value of the system's cost under stochastic variables, as delineated in (1) and (2).

$$\min \sum_{t=1}^{T} R_t^{\text{risk}} \tag{1}$$

$$R_t^{\text{risk}} = (1-\lambda) \cdot E[C_t] + \lambda \cdot \text{Glue}[C_t] \tag{2}$$

where R_t^{risk} is the weighted risk metric cost at period t; $\text{Glue}[C_t]$ is the risk cost calculated by the risk metric; λ is the weighted coefficient; C_t is the system cost at period t. The system cost C_t contains electricity purchase cost, active network loss cost, curtailment penalties for wind and solar power, ESS operation cost, gas turbine (GT) power generation cost, and demand response load dispatch cost, as (3):

$$C_t = \Bigg[c_{\text{buy},t} P_{\text{buy},t} + c_{\text{loss},t} \sum_{ij \in \psi_l} r_{ij} \hat{I}_{ij,t} + \sum_{g}^{G_1} c_{g,t}^{\text{GT}} P_{g,t}^{\text{GT}} + \sum_{w=1}^{N_W} c_{w,t}^{\text{Pen}} \left(P_{w,t}^{\text{W,max}} - P_{w,t}^{\text{W}} \right) + \sum_{p=1}^{N_P} c_{p,t}^{\text{Pen}} \left(P_{p,t}^{\text{PV,max}} - P_{p,t}^{\text{PV}} \right) + \sum_{b=1}^{B} c_{b,t}^{\text{ESS}} \left(P_{b,t}^{\text{ch}} + P_{b,t}^{\text{dis}} \right) + \sum_{t=1}^{T} \sum_{i=1}^{N_L} c_{j,t}^{\text{L}} P_{j,t}^{\text{DR}} \Bigg] \cdot \Delta T \tag{3}$$

where $c_{\text{buy},t}/c_{\text{loss},t}/c_{w,t}^{\text{Pen}}/c_{p,t}^{\text{Pen}}/c_{b,t}^{\text{ESS}}/c_{j,t}^{\text{L}}$ is the cost coefficient for the electricity purchase price/active power loss/wind power/solar power curtailment/ESS/demand response; $P_{\text{buy},t}$ is the purchased power at period t; $\hat{I}_{ij,t}$ is the current squared value flowing through line ij at period t; is the output power of g-th GT at period t; G_1 is the total number of GTs; r_{ij} is the resistance of line ij; $P_{w,t}^{\text{W,max}}/P_{w,t}^{\text{W}}$ denotes the maximum/actual active power of wind turbine (WT) w at period t; $P_{p,t}^{\text{PV,max}}/P_{p,t}^{\text{PV}}$ denotes the maximum/actual active power of PV plant p at period t; $P_{b,t}^{\text{ch}}/P_{b,t}^{\text{dis}}$ is the charging/discharging power of the b-th ESS; B/N_{L} is the total number of ESSs/demand response buses; $P_{j,t}^{\text{DR}}$ is the load regulation of the demand response in bus j at period t; ΔT is the dispatch interval.

2.2. Network Power Flow Constraint

To guarantee the operational safety of the system in the face of the volatility of renewable energy sources such as wind farms and photovoltaic (PV) plants, it is crucial to ensure bus voltage and line current within safe limits. The branch-flow model precisely captures current distribution and system security constraints, convertible to second-order cone constraints as illustrated [18]:

$$\sum_{k \in \delta(j)} P_{jk,t} - \sum_{i \in \pi(j)} (P_{ij,t} - r_{ij} \hat{I}_{ij,t}) = P_{j,t} \tag{4}$$

$$\sum_{k \in \delta(j)} Q_{jk,t} - \sum_{i \in \pi(j)} (Q_{ij,t} - x_{ij} \hat{I}_{ij,t}) = Q_{j,t} \tag{5}$$

$$\hat{U}_{j,t} - \hat{U}_{i,t} + 2(x_{ij} Q_{ij,t} + r_{ij} P_{ij,t}) - (r_{ij}^2 + x_{ij}^2) \hat{I}_{ij,t} = 0 \tag{6}$$

$$\left\| 2P_{ij,t}, 2Q_{ij,t}, \hat{U}_{i,t} - \hat{I}_{ij,t} \right\|_2 \leq \hat{U}_{i,t} + \hat{I}_{ij,t} \tag{7}$$

where $\delta(j)/\pi(j)$ is buses set of the first/last bus in the branch contained j; P_{jk}/Q_{jk} is active/reactive power at the first end of the branch jk; x_{ij} is reactance of branch ij; b_j is grounding conductance of bus j; $\hat{U}_{j,s}$ is voltage magnitude's square value of bus j. When bus j is connected to the GT/WT/PV plant/ESS/load, the injected power can be written as follows:

$$\begin{cases} P_{j,t} = P_{g,t}^{\text{GT}} + P_{w,t}^{\text{W}} + P_{p,t}^{\text{PV}} + P_{b,t}^{\text{dis}} - P_{b,t}^{\text{ch}} - P_{j,t}^{\text{L}} \\ Q_{j,t} = Q_{w,t}^{\text{W}} + Q_{p,t}^{\text{PV}} - Q_{j,t}^{\text{L}} \end{cases} \tag{8}$$

where $Q_{w,t}^{\text{W}}/Q_{p,t}^{\text{PV}}$ is actual reactive power of WT w/PV plant p at period t; $P_{j,t}^{\text{L}}/Q_{j,t}^{\text{L}}$ is actual active/reactive power loads at bus j at period t.

2.3. Generator Unit Operation Constraint

For GTs in the system, the active power is required to satisfy the output constraints as well as the ramp-up and ramp-down constraints during normal operation, as follows:

$$P_{\text{min},g}^{\text{GT}} \leq P_{g,t}^{\text{GT}} \leq P_{\text{max},g}^{\text{GT}} \tag{9}$$

$$\begin{cases} P_{g,t}^{\text{GT}} - P_{g,t-1}^{\text{GT}} \leq R_g^{\text{up}} \\ P_{g,t-1}^{\text{GT}} - P_{g,t}^{\text{GT}} \leq R_g^{\text{down}} \end{cases} \tag{10}$$

where $P^{GT}_{min,g}/P^{GT}_{max,g}$ is minimum/maximum values of g-th GT's active output at period t; R^{up}_g/R^{down}_g is maximum values of g-th GT's ramp-up/ramp-down active output.

For integrated renewable energy sources, the actual active power output of these sources must adhere to the following upper and lower limits:

$$0 \le P^{W}_{w,t} \le P^{W,max}_{w,t} \; 0 \le P^{W}_{w,t} \le P^{W,max}_{w,t} \tag{11}$$

$$P^{W}_{w,t} \cdot \tan \varphi^{W}_{w,min} \le Q^{W}_{w,t} \le P^{W}_{w,t} \cdot \tan \varphi^{W}_{w,max} \tag{12}$$

$$0 \le P^{PV}_{p,t} \le P^{PV,max}_{p,t} \; 0 \le P^{PV}_{p,t} \le P^{PV,max}_{p,t} \tag{13}$$

$$P^{PV}_{p,t} \cdot \tan \varphi^{PV}_{p,min} \le Q^{PV}_{p,t} \le P^{PV}_{p,t} \cdot \tan \varphi^{PV}_{p,max} \tag{14}$$

where $\varphi^{W}_{w,min}/\varphi^{W}_{w,max}$ is minimum/maximum power factor angle of WT w; $\varphi^{PV}_{p,min}\varphi^{PV}_{p,min}/\varphi^{PV}_{p,max}$ is minimum/maximum power factor angle of PV plant p. In the actual operation, $P^{W,max}_{w,t}$ and $P^{PV,max}_{p,t}$ have strong uncertainties due to natural factors such as wind speed and light intensity. Therefore, $P^{W,max}_{w,t}$ and $P^{PV,max}_{p,t}$ are random variables in the established model.

2.4. ESS Operation Constraints

At present, ESSs are increasingly being integrated into power systems as a distributed flexible resource to smooth out the new energy output volatility. In the actual operation, the ESS needs to meet the following operational constraints [19]:

$$R_{b,t} = R_{b,t-1} + P^{ch}_{b,t}\eta_b \Delta T - P^{dis}_{b,t}\Delta T \tag{15}$$

$$R_{b,min} \le R_{b,t} \le R_{b,max} \tag{16}$$

$$0 \le P^{ch}_{b,t} \le P^{ch}_{b,max} \tag{17}$$

$$0 \le P^{dis}_{b,t} \le P^{dis}_{b,max} \tag{18}$$

where $R_{b,t}$ is storage capacity of ESS b at period t; η_b is conversion efficiency; $R_{b,min}/R_{b,max}$ is lower/upper storage capacity; $R_{b,0}$ is the initial storage capacity; $P^{ch}_{b,max}/P^{dis}_{b,max}$ is maximum charging/discharging power.

2.5. Demand Response Constraints

In practice, well-crafted demand response strategies can motivate users to modulate the electricity usage in accordance with the dynamics of grid supply and demand [21–23]. This can be achieved by implementing peak pricing and a tiered incentive system to prompt load reduction, thereby minimizing operational costs. When regulating demand resources, two key approaches should be considered: price-based and incentive-based demand response [24]. Price-based response entails users adjusting their load according to the electricity market's time-of-use pricing, thus managing load fluctuations [25–27]. In contrast, incentive-based response directly compensates users for load reductions during peak times to balance grid supply and demand [28,29]. The formula for calculating the demand response load adjustment $P^{DR}_{j,t}$ is presented below:

$$\begin{cases} P^{DR}_{j,t} = P^{DR,d}_{j,t} - P^{DR,u}_{j,t} + P^{DR,I}_{j,t} \\ P^{L}_{j,t} = P^{L0}_{j,t} - P^{DR}_{j,t} \end{cases} \tag{19}$$

where $P^{DR,u}_{j,t}/P^{DR,d}_{j,t}$ denotes the load increase/decrease due to price-based demand response; $P^{DR,I}_{j,t}$ denotes the load reduction under incentive-based demand response; $P^{L}_{j,t}/P^{L0}_{j,t}$

denotes the active load after/before the demand response. $P_{j,t}^{\mathrm{L0}}$ is also a random variable in the model.

In addition, to ensure the rationality and effectiveness of load regulation, (20) and (21) limit the maximum and minimum adjustment amount of the demand response load:

$$\begin{cases} \sum_{t=1}^{T}\sum_{j=1}^{N_L} P_{j,t}^{\mathrm{DR,u}} = \sum_{t=1}^{T}\sum_{j=1}^{N_L} P_{j,t}^{\mathrm{DR,d}} \\ 0 \le P_{j,t}^{\mathrm{DR,u}} \le \overline{P}_{j,t}^{\mathrm{DR,u}} \\ 0 \le P_{j,t}^{\mathrm{DR,d}} \le \overline{P}_{j,t}^{\mathrm{DR,d}} \end{cases} \tag{20}$$

$$0 \le P_{j,t}^{\mathrm{DR,I}} \le \overline{P}_{j,t}^{\mathrm{DR,I}} \tag{21}$$

where $\overline{P}_{j,t}^{\mathrm{DR,u}}$ and $\overline{P}_{j,t}^{\mathrm{DR,d}}$ are the maximum values of the upward and downward load adjustments for price-based demand response at bus j at period t, respectively; $\overline{P}_{j,t}^{\mathrm{DR,I}}$ is the maximum load reduction under incentive-based demand response.

The established model is a stochastic second-order cone programming model containing complex risk metric operations, which cannot be solved directly by the existing solver, and an improved SADP algorithm based on GlueVaR will be designed in the following to efficiently find the optimal risk-averse strategy.

3. Solution Methodology

3.1. Model Reformulation

Define the state of ESS at period t as $\boldsymbol{R}_t = \{R_{b,t}\}$; the external random variables as $\boldsymbol{W}_t = \{P_{w,t}^{\mathrm{W,max}}, P_{p,t}^{\mathrm{PV,max}}, P_{j,t}^{\mathrm{L0}}\}$; the decision variables as $\boldsymbol{X}_t = \{P_{g,t}^{\mathrm{GT}}, P_{w,t}^{\mathrm{W}}, P_{p,t}^{\mathrm{PV}}, P_{b,t}^{\mathrm{dis}}, P_{b,t}^{\mathrm{ch}}, P_{j,t}^{\mathrm{DR,d}}, P_{j,t}^{\mathrm{DR,u}}, P_{j,t}^{\mathrm{DR,I}}\}$; and introduce $\boldsymbol{S}_t = (\boldsymbol{W}_t, \boldsymbol{R}_t)$ as the pre-decision state and $\mathbf{S}_t^x = (\mathbf{W}_t, \mathbf{R}_t^x)$ as the post-decision state. According to the risk-averse Bellman recursive equation [30], the multi-temporal model can be transformed into a recursive solution of the single-temporal model when the model uses a consistent risk metric to calculate the cost of risk:

$$V_t(\mathbf{S}_t) = \min_{\mathbf{X}_t \in \psi_t} \left(C_t(\mathbf{S}_t, \mathbf{X}_t) + V_t^x(\mathbf{S}_t^x)\right) \tag{22}$$

$$V_t^x(\mathbf{S}_t^x) = R_t^{\mathrm{risk}}(V_{t+1}(\mathbf{S}_{t+1})|(\mathbf{S}_t^x)) \tag{23}$$

where $V_t(\mathbf{S}_t)$ and $V_t^x(\mathbf{S}_t^x)$ are the value functions of $\mathbf{S}_t$ and $\mathbf{S}_t^x$, respectively, and ψ_t is the decision feasible domain. When the analytic expression of $V_t^x(\mathbf{S}_t^x)$ is known, (22) is a deterministic optimization model. However, the expression of the risk-averse cost (23) depends on the information of random variables that are difficult to obtain accurately in practice, which makes the recursive solution difficult. Thus, this paper proposes a risk-averse SADP algorithm based on GlueVaR metric to strive to obtain a computationally efficient near-optimal strategy.

3.2. GlueVaR-Based Risk Metric

Choosing the right risk metrics is critical to accurately measuring the cost of risk. The most popular risk metrics used in power systems are mainly VaR and CVaR, which are calculated as (24) and (25):

$$\mathrm{VaR}_\alpha(Y) = \inf_u \{P(Y \le u) \ge \alpha\} \tag{24}$$

$$\mathrm{CVaR}_\alpha(Y) = \inf_u \{u + 1/(1-\alpha)E[(Y-u)^+]\} \tag{25}$$

where α is the confidence level; u is the VaR value; $P(Y \leq u) \geq \alpha$ means the probability of $Y \leq u$ is at least α; and $(Y-u)^+$ denotes that the $Y-u$ is taken is equal to 0, except when $Y \geq u$.

VaR and CVaR's conservatism hinges on the given confidence level, often set high to minimize the risk costs of system operation. However, decisions at higher confidence levels can impact system operation economics, requiring decision-makers to balance risk avoidance with economic considerations. GlueVaR, a flexible risk metric, has been increasingly used in finance. It offers adjustable parameters to suit diverse risk needs, encompassing VaR and CVaR. The detailed definitions of GlueVaR can be referred in [15], by introducing into the distortion function, the GlueVaR risk metric $\mathrm{Glue}(Y)$ can be expressed in the form of the linear combinations of CVaR and VaR as follows:

$$\begin{cases} \mathrm{Glue}(Y) = k_1 \mathrm{CVaR}_\beta(Y) + k_2 \mathrm{CVaR}_\alpha(Y) + k_3 \mathrm{VaR}_\alpha(Y) \\ k_3 = 1 - (k_1 + k_2), k_1 \geq 0, k_2 \geq 0 \end{cases} \tag{26}$$

While the VaR and CVaR methods can only metric at a single given confidence level, GlueVaR is able to metric at two given confidence levels α and β simultaneously, and by adjusting the parameters it can metric any cost of risk between VaR_α and CVaR_β. The GlueVaR metric introduces two key parameters, k_1 and k_2, which are not merely abstract coefficients but direct levers for specifying operational risk preferences. Their practical interpretation is pivotal for real-world applications:

(1) The parameter k_1 governs aversion to extreme "disaster" scenarios. It applies a risk weight to the very far tail of the loss distribution (associated with a high-confidence VaR). A high k_1 value directs the optimization to prioritize hedging against low-probability, high-impact events (e.g., a compound failure during a once-in-a-decade storm).

(2) The parameter k_2 governs aversion to severe "stress" scenarios. It weighs the conditional expectation of losses beyond a specified VaR threshold (aligned with CVaR). A high k_2 value focuses the strategy on mitigating more frequent, severe stress periods (e.g., a week of consecutively low renewable generation and high demand).

This structure allows system operators to move beyond a single risk measure. For practical tuning, the following guidelines are proposed:

(1) Set (k_1, k_2) = (0, 1) to emulate a pure CVaR-like policy focused on "stress" management; set (1, 0) to emulate a pure high-confidence VaR-like policy focused on "disaster" prevention.

(2) Adjust k_1 and k_2 based on the specific reliability mandates, observed volatility of renewable assets, and the perceived trade-off between preparing for extreme tail events versus managing severe but more common conditions. For instance, a balanced profile like (0.5, 0.5) explicitly tells the algorithm to weigh both concerns equally.

Therefore, the GlueVaR metric provides flexible adjustment of risk preferences, enabling a tailored balance among diverse decision-making needs in practical applications.

3.3. Steps of Improved SADP Algorithm Based on GlueVaR

The core of the proposed algorithm lies in designing and training a set of explicit, parameterized risk-averse AVFs to encapsulate economic risks and overcome the computational "curse of dimensionality". To preserve the physical interpretability of the power system dispatch strategy, piecewise linear functions are employed to parameterize the value functions. This formulation not only effectively captures the marginal effects of the value functions as the system state changes but also ensures that its parameters directly correspond to the risk-adjusted marginal value of dispatchable resources under specific system conditions.

To determine the slopes of segmented linear functions, a large number of sample sets need to be generated during the offline training of the AVFs. Then, the proposed algorithm calculates the risk-averse cost of each sample set by combining with GlueVaR, and updates the slope of the segmented function with the information of the risk of each sample set. After obtaining the converged AVFs, they can be applied online in conjunction with the operation state of the power system. The proposed algorithm can be divided into the following steps as follows:

3.3.1. Initialization

The Monte Carlo approach is applied in generating a large number of error data of random variables as the training sample set $\mathbf{\Phi}$, and the sample set $\mathbf{\Phi}$ is randomly and equally divided into M batches of small sample sets $\mathbf{\Phi}_m$; set the initial slopes $v = 0$, $m = 1$.

3.3.2. Training Approximation Value Functions

(i) For each error scenario φ of a small sample set $\mathbf{\Phi}_m$, solve the corresponding (22) based on segmented linear functions and obtain the optimal decision $\boldsymbol{X}_{\varphi,t}$ and the objective function $C_{\varphi,t}$, and update the state $\mathbf{R}^{\mathrm{x}}_{\varphi,t}$. In order to improve the computational efficiency, each error scenario can be computed in parallel for the same sample set.

(ii) In order to make the AVFs capture the decision-maker's risk appetite, the training objective is not to minimize the traditional expected cost, but to minimize the risk-adjusted future cost distribution measured by GlueVaR. Thus, calculate the cost of risk metrics based on the following (27):

$$\begin{cases} \hat{R}^{\mathrm{risk}}_{t,m} = \frac{(1-\lambda)}{|\Phi_m|} \sum\limits_{\varphi \in \Phi_m} R^{\mathrm{risk}}_{t,m}(\varphi) + \lambda \mathrm{Glue}\left(R^{\mathrm{risk}}_{t,m}\right) \\ R^{\mathrm{risk}}_{t,m}(\varphi) = C_{\varphi,t} + V^{\mathrm{x}}_{\varphi,t} \end{cases} \tag{27}$$

Combining (24)–(26), the risk metric $\mathrm{Glue}\left(R^{\mathrm{risk}}_{t,m}\right)$ can be obtained by solving the following linear optimization problem:

$$\begin{aligned} \mathrm{Glue}\left(R^{\mathrm{risk}}_{t,m}\right) = \min_{\mathbf{u},\mathbf{p}^+} \{ & k_1 u_\beta + (k_2 + k_3) u_\alpha + \\ & \frac{k_1}{(1-\beta)|\Phi_m|} \sum_{\varphi \in \Phi_m} p^+_\beta(\varphi) + \frac{k_2}{(1-\alpha)|\Phi_m|} \sum_{\varphi \in \Phi_m} p^+_\alpha(\varphi) \} \\ \text{s.t. } & R^{\mathrm{risk}}_{t,m}(\varphi) - u_\alpha - p^+_\alpha(\varphi) \le 0 \\ & R^{\mathrm{risk}}_{t,m}(\varphi) - u_\beta - p^+_\beta(\varphi) \le 0 \\ & p^+_\alpha(\varphi) \ge 0, p^+_\beta(\varphi) \ge 0 \end{aligned} \tag{28}$$

where u_α, q^+_α, u_β, q^+_β are auxiliary variables. This objective function indicates that the direction of parameter updates is predominantly determined by the most adverse (i.e., high-cost) batch of scenarios, with the degree of adversity being precisely calibrated by the GlueVaR parameters (k_1, k_2). In this way, the risk measure is directly encoded into the parameters of the AVFs, rather than evaluating the strategy afterward in deep reinforcement learning method.

(iii) Apply a small perturbation ΔR_b to each post-decision state variable at $t-1$ (in which $t = 2, 3, \ldots, T$) and calculate the post-perturbation risk metric cost $\hat{R}^{\mathrm{risk}}_{t,m,b}$ similarly to steps (i), (ii).

(iv) Obtain the observed slope of the post-decision state at the moment $t-1$ (in which $t = 2, 3, \ldots, T$):

$$\hat{v}_{m,b,t-1} = \left(\hat{R}^{\mathrm{risk}}_{t,m} - \hat{R}^{\mathrm{risk}}_{t,m,b}\right) / \Delta R_b \tag{29}$$

where $\hat{\rho}_{t,m,b}$ is the risk metric cost calculated after applying a perturbation to the b-th state variable at time period t in the m-th batch of sample sets; ΔR_b is the amount of the given state perturbation. The provisional slope of the AVFs is updated in the following equation:

$$\overline{v}_{m,b,t-1} = (1 - k_m)v_{m-1,b,t-1} + k_m\hat{v}_{m,b,t-1} \quad (30)$$

where k_m is the slope update step. After obtaining the temporary slope, the successive projection approximation algorithm [31] is used to perform the mapping operation to maintain the convexity of the AVFs and to obtain the updated slope $v_{m,b,t-1}$.

(v) Let $m = m + 1$, if $m < M$, then return to step (i); otherwise, the training process ends and the trained AVFs is output.

3.3.3. Online Applications of Approximate Value Functions

Based on the trained AVFs and the real-time grid operation state data, the established model (22) is solved and output the online optimal dispatch decision ($t = 1, 2, \ldots, T$), and send the decision command to each unit of the grid.

4. Case Studies

4.1. System Parameters

The case studies are conducted on the IEEE 33-bus system, with its topology depicted in Figure 1. The number of GTs, ESSs, PV plants, and WTs in the system are 4, 2, 9, and 4, respectively. Take $R_{b,\min} = 100$ kWh, $R_{b,\max} = 1000$ kWh, $P^{\mathrm{ch}}_{b,\max} = P^{\mathrm{dis}}_{b,\max} = 240$ kW. The forecast data of load, PV, and WT are detailed in Figure 2. The time-of-use electricity prices are detailed in Table 1. Buses 8 and 24 access the incentive-type demand response loads, whose loads and maximum reductions are shown in Figure 3, and buses 18 and 31 access the price-based demand response loads, whose loads and maximum reductions are shown in Figure 4. The uncertainties associated with active load and renewable generation forecast errors are modeled using normal distributions, specifically $N(0, 0.03^2)$ for load and $N(0, 0.15^2)$ for wind/PV output, with output errors truncated at $\pm 10\%$ of capacity. This modeling choice is a common simplification for capturing forecast errors in SO method studies [19], focusing on the mean and variance as the primary moments of uncertainty. Then 5000 error scenarios are extracted as the training sample set using the Monte-Carlo method, and randomly divided equally into 100 batches of small sample sets, each containing 50 error scenarios. All the simulations are made on MATLAB R2019a and each problem is solved by CPLEX solver.

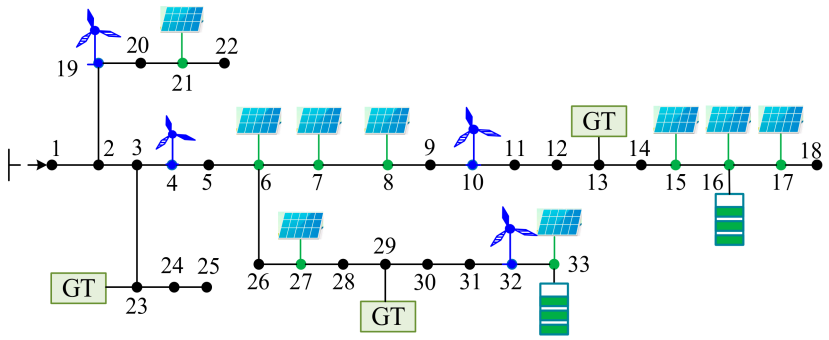


Figure 1. Topology of an IEEE 33-bus system.

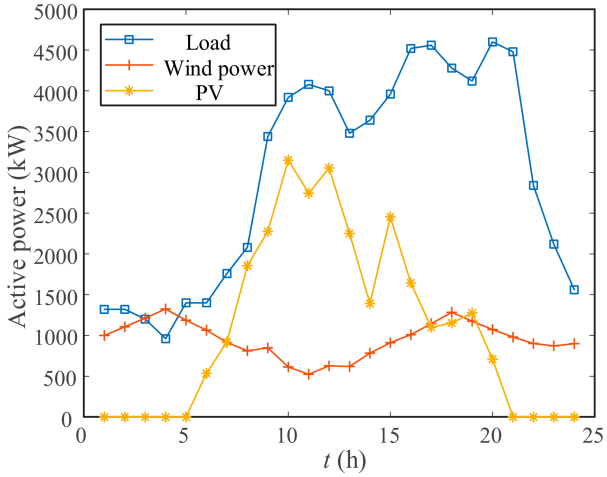


Figure 2. Forecast curve data of Load, Photovoltaic, and Wind Turbine.

Table 1. Time-varying electricity prices used to incentivize economic dispatch and demand response.

Type	Period	Trade Price/(¥/kWh)
Peak	09:00–13:00, 16:00–21:00	1.322
Off-peak	06:00–8:00, 14:00–15:00	0.832
Valley	1:00–5:00, 22:00–24:00	0.369

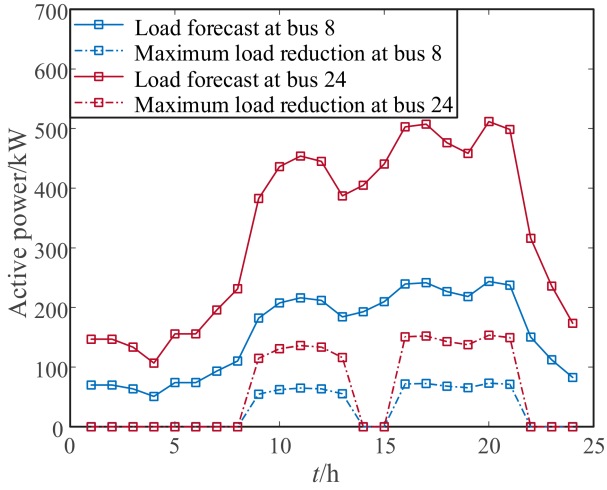


Figure 3. Baseline data profile and available capacity for incentive-based demand reduction.

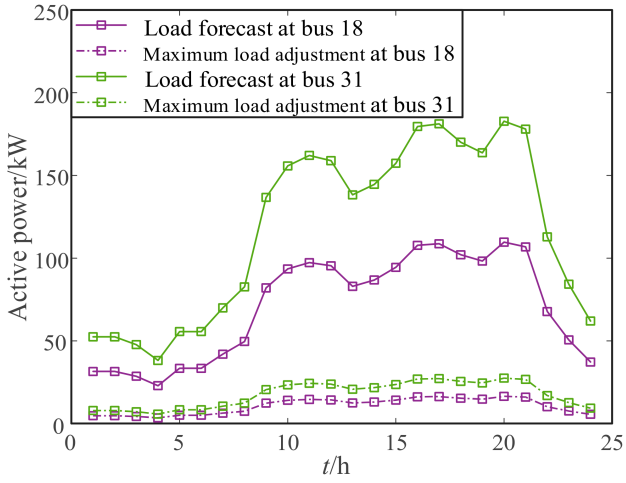


Figure 4. Baseline load data and shiftable capacity for price-based demand response.

4.2. Analysis of Training Results of Improved SADP Algorithm

To assess the convergence and learning performance of the improved SADP algorithm incorporating GlueVaR risk metrics, we use the offline optimal solution as a reference benchmark for comparison, assuming complete knowledge of uncertainties to minimize the total cost. The changes in the average operation cost of the improved risk-averse SADP algorithm, the risk-neutral SADP algorithm ($\lambda = 0$), and the offline optimal solution for each batch of samples during the training process are compared in Figure 5. Meanwhile, the changes in the VaR value at the confidence level of 95% of each algorithm in each batch of samples are compared in Figure 6. The improved SADP algorithm demonstrates good convergence during training. Initially, convergence is challenging due to inaccurate AVFs; however, performance improves as more samples refine the model. After about 30 iterations, the system's operation cost decreases and stabilizes. The improved SADP algorithm's efficiency stems from leveraging the value function's convexity to quickly find the optimal strategy without exhaustively searching the state space. Compared to the risk-neutral SADP algorithm, the risk-averse SADP algorithm in this study shows higher average costs but offers more stability and lower 95%-VaR risk costs in extreme scenarios, suggesting that risk-averse decision of the proposed algorithm leads to more consistent performance under uncertainty.

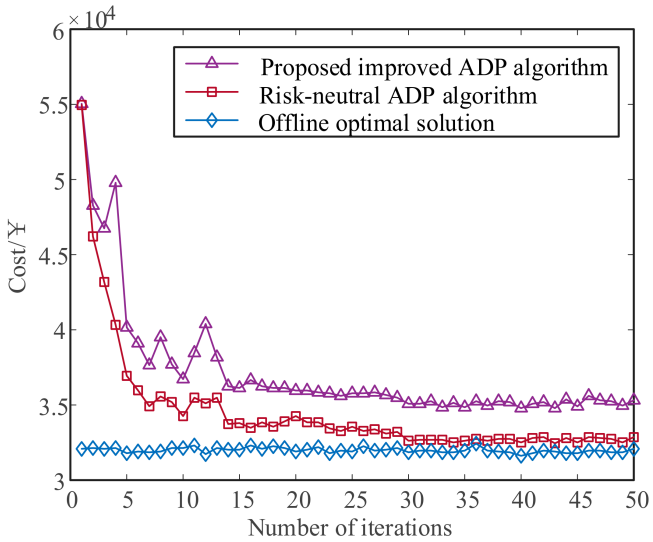


Figure 5. Convergence of average operation cost across training iterations.

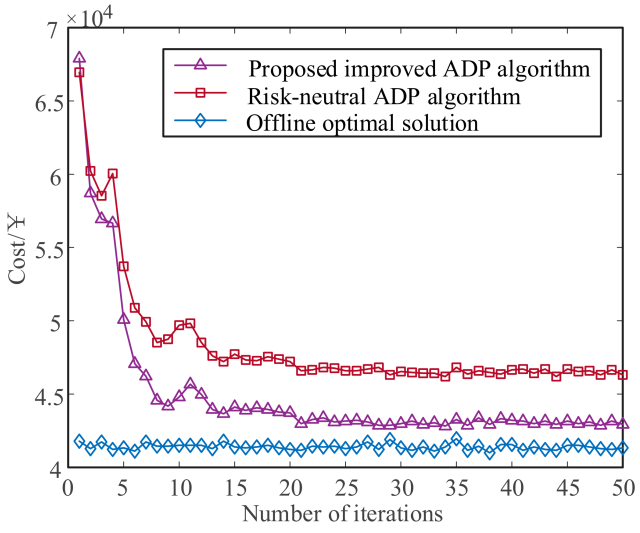


Figure 6. Convergence of the 95% value-at-risk (extreme event cost) across training iterations.

4.3. Analysis of the Algorithm's Solution Results in the Prediction Scenario

Results are presented for a representative prediction scenario. The statistical performance of the proposed policy, including variance analysis across multiple uncertainty realizations, is provided in Section 4.4 (Table 2). The trained AVFs are applied to the decision-making under the predicted scenario of random variables, as illustrated in Figures 7–9. Among them, Figure 7 shows the output of various types of units at each time period under the predicted scenario, and Figures 8 and 9 show the variation curves of the residual power of ESSs and the load demand response, respectively. The GT as a traditional energy source remains almost unchanged, acting as a base energy source. In periods 1 to 13, WTs and PV plants provide a large proportion of the system power, and the excess power is utilized to recharge ESSs, at which point their residual power of ESSs increases. Meanwhile, the system can maintain its power balance without purchasing electricity from the external system. And the load reductions at demand response buses are negative, as shown in Figure 9, indicating that the system power output is abundant and can be increased to run part of the load.

Table 2. Statistical performance comparison under different risk preferences.

(k_1, k_2) Value/Model	Total Average Cost/$	Reduction in 95%-VaR vs. Risk-Neutral/%
Risk-neutral	33,214.31 ± 105.27	0% (Benchmark)
$CVaR_{0.80}$	34,273.89 ± 118.64	4.2% ± 0.6%
$CVaR_{0.95}$	37,186.13 ± 145.92	9.2% ± 0.9%
(0.25, 0.0)	33,521.64 ± 110.85	1.0% ± 0.2%
(0.0, 1.0)	34,273.89 ± 118.64	4.2% ± 0.6%
(0.25, 0.75)	34,750.18 ± 122.56	5.5% ± 0.7%
(0.4, 0.3)	35,121.40 ± 128.40	7.5% ± 0.8%
(0.5, 0.5)	35,888.41 ± 135.77	8.1% ± 0.8%
(0.75, 0.25)	36,112.97 ± 139.65	8.6% ± 0.9%
(1.0, 0.0)	37,186.86 ± 146.01	9.2% ± 0.9%

Note: Values are presented as mean ± standard deviation. The "Reduction in 95%-VaR" is calculated as the relative percentage decrease compared to the risk-neutral benchmark's VaR (47,213.16 $). Its standard deviation reflects the variability of this improvement across runs.

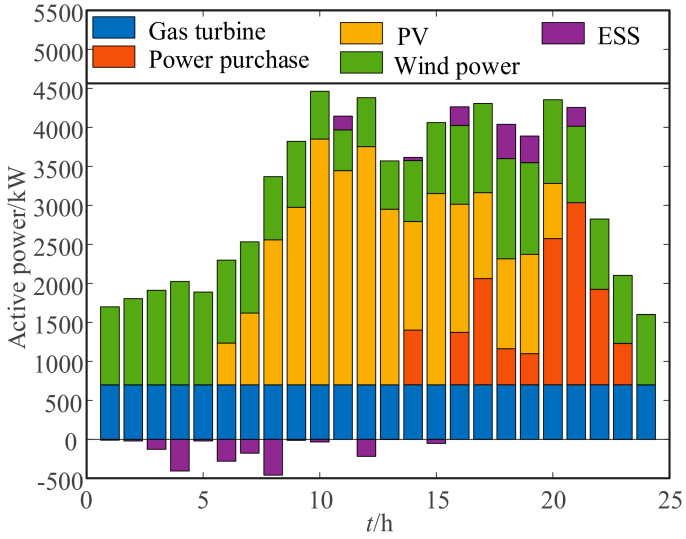


Figure 7. Optimal dispatch schedule: power output of conventional, renewable, and storage units.

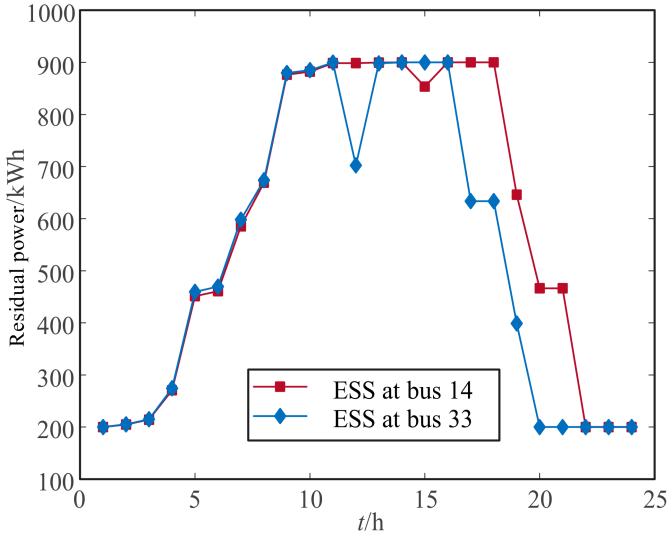


Figure 8. State-of-charge dynamics of the energy storage systems during the dispatch period.

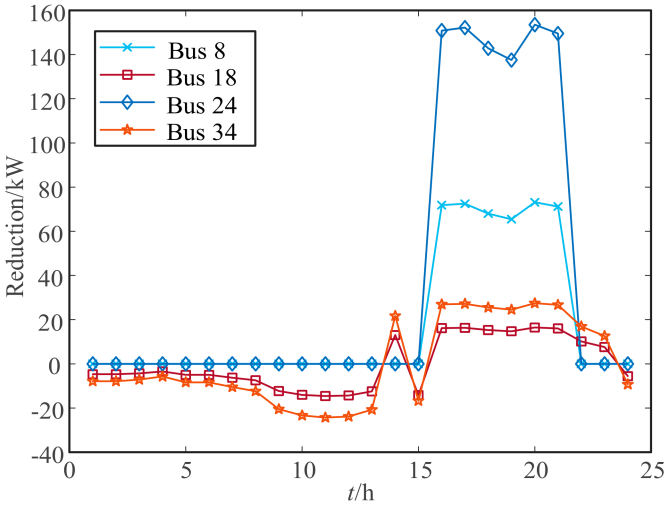


Figure 9. Activated load shifting and curtailment through demand response.

During the periods 14–24, the PV plant's outputs gradually reduced due to insufficient light intensity, and the decision maker began to buy electricity from external systems, the buses of the ESS gradually discharged, and the demand response buses began to reduce the load in response to system demand to meet the system power balance. The whole decision process reflects the flexibility of the source-grid-load-storage coordination, which can flexibly adapt the decision according to the characteristics of renewable energy output.

4.4. Comparison of Test Results Under Different Risk Preferences

The proposed risk-averse SO model utilizes GlueVaR for risk metric, with decision at a given confidence level hinging on the parameters (k_1, k_2). Ref. [15] establishes that GlueVaR is a coherent risk metric, characterized by essential properties such as sub-additive and positively homogeneous within a specific range of parameter values. By adjusting (k_1, k_2) within the range and training the AVFs for each scenario, the AVFs' performance on an independently sampled set of 1000 groups of scenarios is tested. The results, including the 95%-VaR and average cost, are depicted in Figure 10. It is observed that as the values of (k_1, k_2) approach (1, 0), the degree of risk aversion increases, making the average cost in the test set larger, while the 95%-VaR cost is relatively lower; when (k_1, k_2) takes a value farther away from (1, 0), the lower the degree of decision-making risk aversion, the more

economical the decision, i.e., the lower the average cost in the test set, but the extreme cost becomes higher. Meanwhile, it is observed that selecting more conservative values for risk parameters, such as (1, 0), can mitigate the risk costs associated with extreme scenarios and, to some extent, narrow the fluctuation range of operation costs, although it increases the average cost.

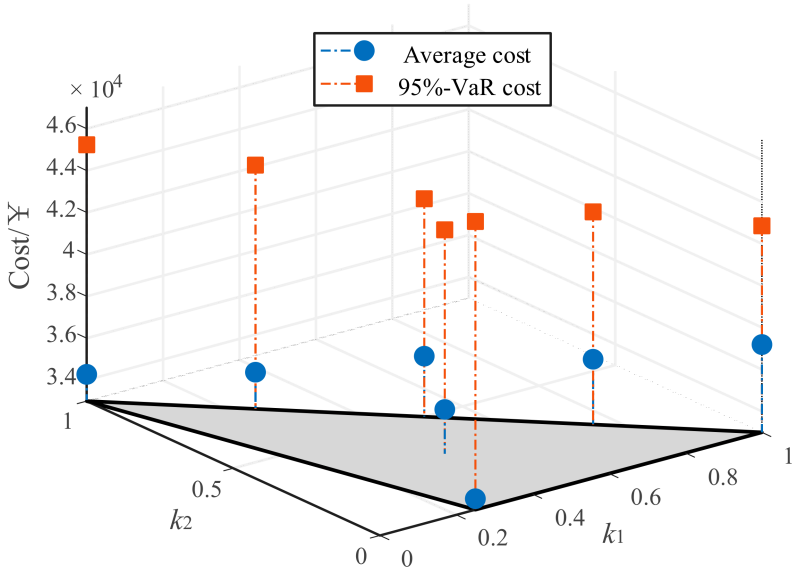


Figure 10. Comparative test results under different GlueVaR risk preference parameters (k_1, k_2) values.

The statistical comparison of test results under different risk preferences is presented in Table 2. Based on 50 independent Monte Carlo runs, the results are expressed as "mean ± standard deviation" demonstrating the statistical stability of the algorithm's performance. It can be observed that the results for $CVaR_{0.80}$ are nearly identical to those of the GlueVaR metric with parameters (k_1, k_2) = (0, 1), while the results for $CVaR_{0.95}$ are very close to those of GlueVaR with (k_1, k_2) = (1, 0). This correspondence occurs because, with these parameter sets, the GlueVaR measure simplifies to $CVaR_{0.80}$ and $CVaR_{0.95}$, respectively, as shown in (26).

More importantly, the parameters (k_1, k_2) serve as practical levers for tuning risk attitudes beyond these fixed corner cases. As outlined in Section 3.2, k_1 primarily governs aversion to extreme, low-probability "disaster" scenarios, whereas k_2 controls aversion to more frequent, severe "stress" scenarios. The performance spectrum in Table 2 directly reflects this operational interpretation. For instance, the configuration (0.5, 0.5), which explicitly weighs both disaster mitigation and stress management equally, achieving a balanced trade-off: it yields a stable reduction of 8.1% (with a standard deviation of ±0.8%) in extreme risk (95%-VaR) at the cost of an 8.1% increase in the average total cost. In contrast, the (0.0, 1.0) configuration provides less extreme risk reduction for a lower cost penalty, while the (1.0, 0.0) configuration achieves the highest risk reduction at the highest economic cost. This parametric flexibility allows the GlueVaR metric to adapt to specific, practical decision-making needs, demonstrating a clear advantage over single risk measures with fixed confidence levels.

4.5. Comparison of Test Results Under Different Methods

To demonstrate the advantages of our proposed method, the comparison with a wider range of state-of-the-art risk-averse methods and AI-based methods in the modified IEEE 33-bus system are shown in Table 3. Five representative methods are as follows:

Table 3. Performance comparison of different methods on the modified IEEE 33-Bus system.

Method	Average Cost ($)	95% VaR ($)	Offline Time (s)	Online Time (s)
Proposed method (GlueVaR-SADP)	35,121	43,687.	<8471	158
Risk-Neutral S-OPF	33,214	47,213	–	2468
IGDT-based Method	38,900	46,100	–	3498
CVaR-SOPF (α = 0.95)	37,886	42,854	–	3141
Robust OPF (Γ = 0.8)	38,350	50,300	–	1142
DRL Method	10,900	48,500	>65,000	241

Note: "–" indicates that the method does not require a traditional offline training phase.

(1) Risk-neutral stochastic optimal power flow (S-OPF) [5]: serves as the economic performance lower bound.

(2) Information-gap decision theory (IGDT) based robust method [12]: the non-probabilistic, extreme uncertainty-averse method.

(3) CVaR based S-OPF [13]: the standard single-metric probabilistic risk-averse method.

(4) Robust OPF (R-OPF) [8]: the conservative worst-case optimization method.

(5) Deep Reinforcement Learning (DRL) [32]: the data-driven AI baseline method.

The proposed GlueVaR-SADP method demonstrates a great balance between risk mitigation, economic efficiency, and computational tractability. In terms of risk management, it achieves a 7.5% reduction in 95% VaR from the risk-neutral benchmark, effectively curbing extreme event costs. While its average operational cost is moderately higher than the risk-neutral lower bound, it remains more economical than all other risk-averse benchmarks, including IGDT, CVaR-SOPF, and R-OPF, highlighting its efficient trade-off. In terms of computational performance, the proposed method requires less than 8500 s for offline training, which is an order of magnitude faster than the DRL baseline (exceeding 65,000 s). This efficiency stems from the parallelizable architecture of the SADP framework. During online application, it delivers a solution in approximately 160 s, striking a favorable balance between the prolonged solving times of scenario-based methods (including S-OPF, CVaR-SOPF) and the simpler but more conservative robust approach. Moreover, unlike the black-box DRL policy, the proposed method provides explicit and interpretable risk quantification, a crucial feature for operational acceptance.

4.6. Test Results in the Modifed IEEE 69-Bus System

To thoroughly evaluate the scalability and generalization capability of the proposed method, comprehensive tests are conducted on a larger and more complex modified IEEE 69-bus distribution system. This system integrates a higher penetration of renewable resources, distributed storage, with its topology depicted in Figure 11. Under the same probabilistic scenario set, the systematic comparison between the proposed GlueVaR-SADP method and the five benchmark methods described in Section 4.5 is presented here.

As shown in Table 4, the proposed method maintains its consistent advantages in the 69-bus system. In terms of risk control, its 95% VaR is significantly lower than those of the risk-neutral, IGDT, and robust optimization benchmarks, and is comparable to the CVaR method, while offering more flexible risk preference tuning. Economically, its average cost remains lower than all other risk-averse benchmarks except the risk-neutral one, demonstrating a good balance between risk mitigation and economic efficiency.

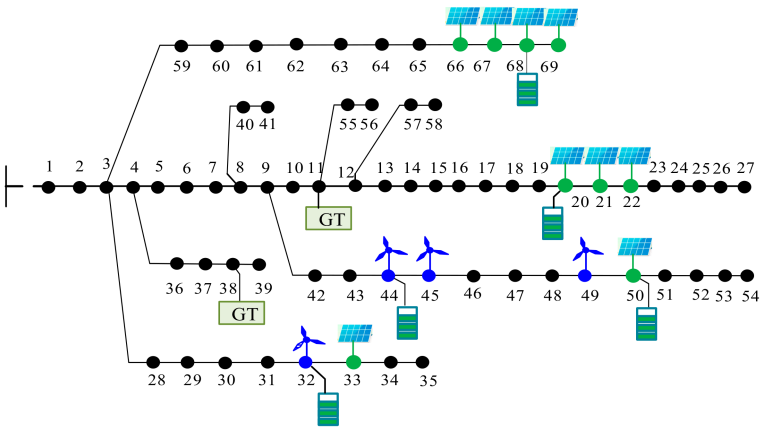


Figure 11. Topology of the IEEE 69-bus system.

Table 4. Performance comparison of different methods on the modified IEEE 69-bus system.

Method	Average Cost ($)	95% VaR ($)	Offline Time (s)	Online Time (s)
Proposed method (GlueVaR-SADP)	72,150	91,105	<16,000	285
Risk-Neutral S-OPF	66,550	103,420	–	6847
IGDT-based Method	78,900	95,215	–	6541
CVaR-SOPF (α = 0.95)	75,880	90,300	–	6970
Robust OPF (Γ = 0.8)	82,450	102,500	–	2210
DRL Method	61,200	104,550	>140,000	654

The comparison of computational performance is particularly notable. The offline training time of the proposed method is far less than that of the DRL method. Its online solution time achieves the best trade-off between accuracy and speed: it is an order of magnitude faster than scenario-based stochastic optimization methods (S-OPF, CVaR-SOPF) while avoiding the excessive conservatism of robust optimization (R-OPF). Compared with the results from the 33-bus system, as the system scale increases, the growth in the online computation time for the proposed method is much lower than the increase in problem complexity, verifying the excellent scalability of its parallel computing architecture.

5. Discussion

This section provides an in-depth interpretation of the core findings, mechanisms, positioning, and significance of the proposed risk-averse SADP algorithm. To ensure a clear presentation, the discussion is organized into the following subsections.

5.1. Key Findings and Core Validation

The empirical evidence from our case studies reveals that the risk-averse SADP algorithm can greatly reduce the operation risk costs under extreme uncertainty scenarios while maintaining computational tractability for real-time grid operations. This achievement directly validates our hypothesis that integrating GlueVaR risk metrics into value function approximation enables more robust decision-making without sacrificing the economic efficiency that system operators require.

The mechanism underlying this improvement warrants careful examination. Traditional stochastic optimization approaches in power systems have struggled with what we might call the "uncertainty paradox"—the need to account for rare but catastrophic events without becoming paralyzed by over-conservative strategies. Our algorithm addresses this through its dual-layer risk assessment architecture. During offline training, the parallel

computation framework processes multiple uncertainty scenarios, but rather than treating each scenario equally, the GlueVaR metric assigns differential weights based on tail risk characteristics. This selective attention to extreme events, combined with the value function's ability to encode risk preferences directly into state transition mappings, creates a decision framework that remains vigilant against high-impact events without sacrificing day-to-day operational efficiency.

5.2. Multi-Scenario Performance Analysis and Scalability

Power system operators face increasingly complex decision environments where renewable penetration levels exceed 40% in many regions. Our framework's ability to coordinate source-grid-load-storage resources under such conditions represents a fundamental shift in operational philosophy. Rather than treating uncertainty as a constraint to be minimized, the algorithm leverages uncertainty information to create adaptive strategies that exploit favorable conditions while protecting against adverse scenarios. This paradigm shift becomes particularly relevant when considering the evolving regulatory landscape, where grid operators face both reliability mandates and renewable integration targets that often seem contradictory.

To comprehensively evaluate the broader applicability of our approach, we conducted extended sensitivity analyses across diverse operational contexts. The following analysis presents a systematic comparison of algorithm performance under varying grid configurations and uncertainty levels. The data presented in Table 5 illuminates several crucial relationships that merit deeper investigation. Most strikingly, the correlation between renewable penetration levels and algorithm performance exhibits non-linear characteristics that challenge conventional assumptions about renewable integration challenges. Systems with higher renewable penetration, particularly the island system scenario with 80% renewables, demonstrate the greatest improvements in both average cost reduction and tail risk mitigation. This counterintuitive result emerges from the algorithm's sophisticated handling of uncertainty—as renewable variability increases, the value of intelligent coordination between storage, demand response, and conventional generation becomes more pronounced, creating larger optimization spaces where our risk-aware approach can identify superior solutions.

Table 5. Performance metrics under diverse operational scenarios and uncertainty levels.

Scenario Configuration	Renewable Penetration (%)	Average Cost Reduction (%)	95%-VaR Improvement (%)	Convergence Speed (Iterations)	Computational Time (Seconds)
Baseline Grid	35	18.4 ± 2.1	27.3 ± 3.2	28 ± 3	145 ± 12
High Renewable	65	24.7 ± 2.8	31.2 ± 3.6	32 ± 4	156 ± 14
Urban Microgrid	45	21.3 ± 2.4	29.8 ± 3.4	26 ± 3	132 ± 11
Industrial Complex	30	15.2 ± 1.8	22.6 ± 2.7	24 ± 2	128 ± 10
Island System	80	28.9 ± 3.1	35.7 ± 4.1	37 ± 5	168 ± 15

The urban microgrid scenario presents particularly interesting dynamics. Despite moderate renewable penetration at 45%, it achieves convergence faster than any other configuration while maintaining strong performance metrics. This efficiency stems from the dense interconnection topology typical of urban grids, where multiple pathways for power flow create redundancy that the algorithm exploits through its network-aware optimization structure. The algorithm recognizes these topological advantages during the training phase, encoding them into value functions that prioritize flexible routing strategies during high-uncertainty periods.

To statistically validate the algorithm's performance, a Monte Carlo analysis comprising 50 independent runs is conducted for the high-renewable penetration scenario in Table 6. The proposed GlueVaR-SADP algorithm achieves a mean reduction of 31.2%

(with a standard deviation of ±1.7%) in extreme-event costs (95%-VaR) compared to the risk-neutral benchmark. Concurrently, it maintains the average operational cost within 7.5% (mean, ±0.9% standard deviation) of the risk-neutral solution.

Table 6. Statistical performance of the proposed GlueVaR-SADP algorithm in the high-renewable penetration scenario (65% renewables).

Metric	Risk-Neutral Stochastic OPF (Benchmark)/$	Proposed GlueVaR-SADP/$	Deviation
95%-VaR	68,950 ± 312	47,300 ± 285	31.2% ± 1.7% Reduction
Average Operational Cost	28,725 ± 120	30,879 ± 118	+7.5% ± 0.9% (vs. Benchmark)

5.3. Limitations and Future Research Directions

This section comprehensively discusses the limitations of the current study and, based on these limitations as well as the research findings, proposes clear directions for future work.

5.3.1. Limitations of the Current Study

Although the proposed framework has achieved positive results, several limitations constrain its generalizability:

(1) Firstly, the assumption of normally distributed forecast errors, adopted for computational convenience, may not accurately capture the heavy-tailed distribution characteristics often present in extreme weather events. Secondly, the current model assumes perfect compliance from demand response resources, which does not align with real-world phenomena such as user behavior degradation and fatigue effects. Furthermore, the treatment of communication delays and measurement errors as negligible may not hold in deployment scenarios with limited infrastructure.

(2) The current framework is primarily designed for operational optimization at the distribution network and microgrid level. It does not consider $N - 1$ contingency constraints or cascading failure scenarios, which are crucial for transmission system security analysis. Additionally, the model does not incorporate large-scale electric vehicle (EV) fleets as spatiotemporally coupled mobile storage resources, representing a gap in modeling storage flexibility.

5.3.2. Future Research Directions

Addressing the above limitations and building upon the technical pathway established in this research, future work will focus on the following extensions:

(1) A primary direction is the integration of electric vehicle (EV) fleets. This involves designing novel state representation methods to characterize their spatiotemporal uncertainty and probabilistic flexibility, thereby enhancing system economics and demand response capabilities. Concurrently, there is a need to develop refined demand response models capable of capturing dynamic user behaviors to improve the practicality of generated strategies.

(2) To address information uncertainty, developing robust versions of the algorithm is necessary to maintain performance under higher-order uncertain information. Furthermore, the intrinsic relationship between algorithm solving efficiency and decision quality should be systematically investigated, striving to improve computational performance while ensuring the interpretability of the physical models.

(3) The current framework, focused on short-term operations, should be extended to long-term investment decision problems such as generation expansion planning, investigating methods for embedding risk preferences under long-term uncertainty. Exploring

the potential for incorporating emerging technologies like hydrogen energy storage and peer-to-peer energy trading into the modeling framework is also promising.

6. Conclusions

This study proposes a risk-averse SO model for new power systems and develops an improved SADP algorithm for efficient solving. The method incorporates the GlueVaR risk metric into the training of risk-aware approximate value functions. Compared with the results of the risk-neutral SADP algorithm, although the decisions of the improved algorithm increase the total average cost under uncertainty, it can significantly mitigate the costs associated with extreme events and narrows the operation cost fluctuation range. In addition, the introduced GlueVaR metric is more flexible than the CVaR metric for decision-making, which not only covers the CVaR metric decision-making at two different confidence levels, but also enables the risk preference to be flexibly adjusted between the two through the appropriate parameter values, making it more suitable for complex and variable practical applications. Moreover, numerical results indicate a strong relationship between conservatism parameters and reserve capacity allocation, offering practical support for risk-adaptive decision-making.

Furthermore, a key advantage of the proposed algorithm is its computational performance and scalability. Comprehensive tests on both IEEE 33-bus and larger 69-bus systems demonstrate that the algorithm enables fast online decision-making suitable for real-time operations. Its offline training is more efficient and stable than data-driven deep reinforcement learning approaches. Crucially, the parallel computing architecture ensures excellent scalability, as the increase in online solution time remains lower than the growth in problem complexity when moving to larger systems. Therefore, this work provides not only a flexible risk-management paradigm but also a computationally efficient and scalable decision-support tool, offering a practical solution for real-time, risk-aware dispatch in power systems with high renewable penetration.

Author Contributions: The authors confirm their contribution to the paper as follows: Conceptualization, Z.G., P.Y., S.L. and Y.Y.; Data curation, P.Y., W.Z. and J.L.; Writing—original draft, Z.G., P.Y., X.D., S.L. and Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Project of the China Southern Power Grid under Grant number GDKJXM20230487 (030600KC23040024).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: All authors express gratitude for the support and cooperation provided by their respective institutions.

Conflicts of Interest: Authors Zike Guo, Peng Yang, Xue Du, Wanmei Zhao and Jiehua Lu were employed by the company Foshan Power Supply Bureau of Guangdong Power Grid Co., Ltd. Authors Siliang Liu and Yingqi Yi were employed by the company Guangzhou Power Electrical Technology 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.

References

1. Cheng, L.; Yu, F.; Huang, P.; Liu, G.; Zhang, M.; Sun, R. Game-theoretic evolution in renewable energy systems: Advancing sustainable energy management and decision optimization in decentralized power markets. *Renew. Sustain. Energy Rev.* **2025**, *217*, 115776. [CrossRef]
2. Fernández Valderrama, D.; Guerrero Alonso, J.I.; León de Mora, C.; Robba, M. Scenario Generation Based on Ant Colony Optimization for Modelling Stochastic Variables in Power Systems. *Energies* **2024**, *17*, 5293. [CrossRef]

3. Cheng, L.; Peng, P.; Huang, P.; Zhang, M.; Meng, X.; Lu, W. Leveraging evolutionary game theory for cleaner production: Strategic insights for sustainable energy markets, electric vehicles, and carbon trading. *J. Clean. Prod.* **2025**, *512*, 145682. [CrossRef]
4. Dong, W.; Zhang, F.; Li, M.; Li, J.; Sun, Y. Imitation Learning Based Real-Time Decision-Making of Microgrid Economic Dispatch Under Multiple Uncertainties. *J. Mod. Power Syst. Clean. Energy* **2024**, *12*, 1183–1193. [CrossRef]
5. Plytaria, K.A.; Steen, D.; Tuan, L.A.; Galanis, N.M.P. Scenario-based Stochastic Optimization for Energy and Flexibility Dispatch of a Microgrid. *IEEE Trans. Smart Grid* **2022**, *13*, 3328–3341. [CrossRef]
6. Zhang, X.; Ding, T.; Mu, C.; Zhang, Y. Dual Stochastic Dual Dynamic Programming for Multi-stage Economic Dispatch with Renewable Energy and Thermal Energy Storage. *IEEE Trans. Power Syst.* **2024**, *39*, 3725–3737. [CrossRef]
7. Chen, H.; Huang, J.; Lin, Z.; Chen, Q. Stochastic Economic Dispatch Based Optimal Market Clearing Strategy Considering Flexible Ramping Products under Wind Power Uncertainties. *CSEE J. Power Energy Syst.* **2024**, *10*, 1525–1535. [CrossRef]
8. Nguyen, T.A.; Crow, M.L. Stochastic Optimization of Renewable-based Microgrid Operation Incorporating Battery Operating Cost. *IEEE Trans. Power Syst.* **2016**, *31*, 2289–2296. [CrossRef]
9. Chen, Y.; Wei, W. Robust Generation Dispatch with Strategic Renewable Power Curtailment and Decision-dependent Uncertainty. *IEEE Trans. Power Syst.* **2023**, *38*, 4640–4654. [CrossRef]
10. Ding, T.; Xu, Y.; Li, Z. Multi-stage Distributionally Robust Stochastic Dual Dynamic Programming to Multi-Period Economic Dispatch with Virtual Energy Storage. *IEEE Trans. Sustain. Energy* **2022**, *13*, 146–158. [CrossRef]
11. Chen, Z.; Zhu, G.; Zhang, Y.; Ji, T.; Liu, Z.; Lin, X.; Cai, Z. Stochastic Dynamic Economic Dispatch of High Wind-Integrated Electricity and Natural Gas System Considering Security Risk Constraints. *CSEE J. Power Energy Syst.* **2019**, *5*, 324–334. [CrossRef]
12. Daneshvar, M.; Eskandari, H.; Sirous, A.B.; Haghifam, M.R.; Abedi, M. A Novel Techno-economic Risk-averse Strategy for Optimal Scheduling of Renewable-based Industrial Microgrid. *Sustain. Cities Soc.* **2021**, *70*, 102879. [CrossRef]
13. Li, Z.M.; Wu, L.; Xu, Y.; Li, Y. Risk-averse Coordinated Operation of a Multi-Energy Microgrid Considering Voltage/Var Control and Thermal Flow: An Adaptive Stochastic Approach. *IEEE Trans. Smart Grid* **2021**, *12*, 3914–3927. [CrossRef]
14. Farzan, F.; Jafari, M.; Masiello, R.; Cherkaoui, R. Toward Optimal Day-ahead Scheduling and Operation Control of Microgrids under Uncertainty. *IEEE Trans. Smart Grid* **2015**, *6*, 499–507. [CrossRef]
15. Belles-Sampera, J.; Guillén, M.; Santolino, M.; Tomas-Rodriguez, A.M. Beyond Value-at-risk: GlueVaR Distortion Risk Metrics. *Risk Anal.* **2014**, *34*, 121–134. [CrossRef] [PubMed]
16. Ma, Z.; Gao, J.; Hu, W.; Zhang, Y. Risk-adjustable Stochastic Schedule Based on Sobol Augmented Latin Hypercube Sampling Considering Correlation of Wind Power Uncertainties. *IET Renew. Power Gener.* **2021**, *15*, 2356–2367. [CrossRef]
17. Zhu, J.; Zhu, W.; Chen, J.; Xu, Y. A DRO-SDDP Decentralized Algorithm for Economic Dispatch of Multi Microgrids with Uncertainties. *IEEE Syst. J.* **2023**, *17*, 6492–6503. [CrossRef]
18. Lin, S.; Wang, Y.; Liu, M.; Wang, J. Stochastic Optimal Dispatch of PV/Wind/Diesel/Battery Microgrids using State-Space Approximate Dynamic Programming. *IET Gener. Transm. Distrib.* **2019**, *13*, 3409–3420. [CrossRef]
19. Lin, S.; Fan, G.; Jian, G.; Li, Y. Stochastic Economic Dispatch of Power System with Multiple Wind Farms and Pumped-Storage Hydro Stations using Approximate Dynamic Programming. *IET Renew. Power Gener.* **2020**, *14*, 2507–2516. [CrossRef]
20. Das, A.; Wu, D.; Ni, Z.; Xu, Y. Approximate Dynamic Programming with Policy-based Exploration for Microgrid Dispatch under Uncertainties. *Int. J. Electr. Power Energy Syst.* **2022**, *142*, 108359. [CrossRef]
21. Thang, V.V.; Ha, T.; Li, Q.; Zhang, Y. Stochastic Optimization in Multi-Energy Hub System Operation Considering Solar Energy Resource and Demand Response. *Int. J. Electr. Power Energy Syst.* **2022**, *141*, 108132. [CrossRef]
22. Cheng, L.; Wei, X.; Li, M.; Tan, C.; Yin, M.; Shen, T.; Zou, T. Integrating evolutionary game-theoretical methods and deep reinforcement learning for adaptive strategy optimization in user-side electricity markets: A comprehensive review. *Mathematics* **2024**, *12*, 3241. [CrossRef]
23. Cheng, L.; Li, M.; Tan, C.; Huang, P.; Zhang, M.; Sun, R. Computational game-theoretic models for adaptive urban energy systems: A comprehensive review of algorithms, strategies, and engineering applications. *Arch. Comput. Methods Eng.* **2026**, *33*, 2037–2114. [CrossRef]
24. Hosseini, S.E.; Najafi, M.; Akhavein, A.; Haghifam, M.R. Day-ahead Scheduling for Economic Dispatch of Combined Heat and Power with Uncertain Demand Response. *IEEE Access* **2022**, *10*, 42441–42458. [CrossRef]
25. Wen, L.; Zhou, K.; Feng, W.; Yang, S. Demand side management in smart grid: A dynamic-price-based demand response model. *IEEE Trans. Eng. Manag.* **2022**, *71*, 1439–1451. [CrossRef]
26. Cheng, L.; Sun, R.; Wang, K.; Yu, F.; Huang, P.; Zhang, M. Advancing sustainable electricity markets: Evolutionary game theory as a framework for complex systems optimization and adaptive policy design. *Complex Intell. Syst.* **2025**, *11*, 320. [CrossRef]
27. Wan, Y.; Qin, J.; Yu, X.; Yang, T.; Kang, Y. Price-based residential demand response management in smart grids: A reinforcement learning-based approach. *IEEE/CAA J. Autom. Sin.* **2021**, *9*, 123–134. [CrossRef]
28. Cheng, L.; Wang, K.; Peng, P.; Zou, T.; Huang, P.; Zhang, M. Multi-agent stackelberg game for joint optimization of electricity spot and deep peak regulation markets: Strategies and implications for system flexibility. *Int. J. Electr. Power Energy Syst.* **2025**, *171*, 111041. [CrossRef]

29. Alquthami, T.; Milyani, A.H.; Awais, M.; Rasheed, M.B. An incentive based dynamic pricing in smart grid: A customer's perspective. *Sustainability* **2021**, *13*, 6066. [CrossRef]
30. Ruszczynski, A. Risk-averse Dynamic Programming for Markov Decision Processes. *Math. Program.* **2010**, *125*, 235–261. [CrossRef]
31. Nascimento, J.; Powell, W.B. An Optimal Approximate Dynamic Programming Algorithm for Concave, Scalar Storage Problems with Vector-valued Controls. *IEEE Trans. Autom. Control* **2013**, *58*, 2995–3010. [CrossRef]
32. Deng, B.; Chen, J.; Ding, Q.; Pan, Z.; Yu, T.; Wang, K. Multi-task Deep Reinforcement Learning Optimal Dispatch Based on Grid Operation Scenario Clustering. *Power Syst. Technol.* **2023**, *47*, 978–990. [CrossRef]

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.