A two-stage distributionally robust optimization approach for integrated energy systems: coordinated configuration of CCUS and hydrogen

Energy 347 (2026) 140331

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

A two-stage distributionally robust optimization approach for integrated energy systems: coordinated configuration of CCUS and hydrogen utilization

Zhewei Wang [a], Jie Yuan [a], Yang Li [b], Leiqi Zhang [d,*], Changjun Xie [a,**], Wenchao Zhu [c], Yang Yang [a], Han Wang [a]

[a] School of Automation, Wuhan University of Technology, Wuhan, 430070, China
[b] Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, 41258, Sweden
[c] School of Automotive Engineering, Wuhan University of Technology, Wuhan, 430070, China
[d] State Grid Zhejiang Electric Power Research Institute, Hangzhou, 310014, China

ARTICLE INFO

Keywords:
Integrated energy system
Carbon capture utilization and storage
Hydrogen energy comprehensive utilization
Optimal configuration
Distributionally robust optimization

ABSTRACT

Facing global energy shortages and carbon emission reduction requirements, the optimal configuration of integrated energy systems (IESs) has become a critical research direction. This study proposes a two-stage distributionally robust optimization (DRO) method for IES configuration that considers carbon capture, utilization, and storage as well as comprehensive hydrogen energy utilization. First, an IES architecture incorporating a hydrogen-carbon coordinated power system (HCPS) module is constructed. Specifically, this module achieves carbon-hydrogen coupling by directing captured CO_2 and green hydrogen into a methane reactor for synthetic natural gas production, establishing a 'capture-conversion-reutilization' closed loop. Crucially, a two-stage DRO model based on the Wasserstein distance is developed to coordinate capacity configuration and operational scheduling under renewable scenarios generated via improved K-means clustering. The model is then transformed into a mixed-integer linear programming problem and solved using the Big-M method and strong duality theory. Finally, the proposed method is validated through a case study using data from an IES in the coastal region of Zhejiang, China. The results show that, driven by this coupling mechanism, the IES integrated with HCPS reduces wind and photovoltaic curtailment costs by 82.8%, carbon emission costs by 43%, and the total daily cost by 2.3%, achieving synergistic improvement in both economic efficiency and low-carbon performance. Compared to the classical robust optimization model, the data-driven DRO model based on the Wasserstein ambiguity set achieves a lower total cost. When the sample size increases to 2000, the computation time is only 17.53s, significantly enhancing computational efficiency.

Nomenclature

Superscript

e	electrical energy
h	thermal energy
g	natural gas
H_2	hydrogen energy
c	cold energy
CO_2	carbon dioxide

Subscripts

ch	charge
dis	discharge
buy	system energy procurement
$grid$	power grid
ce	curtailed energy
sys	system
$load$	system load

Symbols

P	power
C	carbon dioxide flow
$\widehat{C}$	free carbon emission quota
c	cost
η	energy conversion coefficient
N_{day}	system operation days
r	discount rate
μ	lifetime of the devices

(*continued on next column*)

(*continued*)

PV	photovoltaic array
WT	wind turbine
GT	gas turbine
GB	gas boiler
FC	fuel cell
EL	electrolyzer
$CCUS$	carbon capture, utilization and storage

(*continued on next page*)

* Corresponding author.
** Corresponding author.
E-mail addresses: 247694@whut.edu.cn (Z. Wang), 610578606@qq.com (J. Yuan), yangli@ieee.org (Y. Li), y_delta@outlook.com (L. Zhang), jackxie@whut.edu.cn (C. Xie), zhuwenchao@whut.edu.cn (W. Zhu), whutyangyang@whut.edu.cn (Y. Yang), 335284@whut.edu.cn (H. Wang).

https://doi.org/10.1016/j.energy.2026.140331
Received 30 October 2025; Received in revised form 6 January 2026; Accepted 3 February 2026
Available online 8 February 2026

(*continued*)

MR	methane reactor	*TC*	total cost
EC	electric chiller	*EOC*	energy ordering cost
AC	absorption chiller	*EMC*	system maintenance cost
TR	thermal recovery auxiliary equipment	*ECC*	wind and solar curtailment cost
HES	hydrogen energy storage	*AIC*	average investment cost
BES	battery energy storage	*CEC*	carbon emission cost

1. Introduction

Against the backdrop of rising global energy demand and accelerating structural transition, energy systems are confronted with the dual challenges of energy shortages and the large-scale integration of renewable energy. The heavy reliance on traditional fossil fuels exacerbates the imbalance between supply and demand. Specifically, in 2024, fossil fuels accounted for approximately 80.1% of global primary energy consumption, triggering environmental issues such as global warming [1]. While the installed capacity of renewable energy sources like wind and solar power has surged, their inherent intermittency pose significant challenges to grid stability [2–4]. The "dual carbon" goals (carbon peak and carbon neutrality) are driving an urgent and imperative transition of the energy system towards cleaner, low-carbon, and higher-efficiency paradigms. Integrated energy systems enhance energy efficiency and promote renewable energy integration through multi-energy complementarity and cascade utilization. Consequently, IESs have emerged as a core pathway to address energy security and environmental constraints and facilitate the low-carbon transition, and they hold a pivotal position in the modern energy landscape. Carbon capture, utilization, and storage technology(CCUS) enables deep emission reductions and can generate revenue through carbon utilization. Meanwhile, hybrid energy storage systems, which incorporate comprehensive hydrogen utilization, leverage hydrogen's high energy density and zero-carbon characteristics to effectively mitigate fluctuations from wind and solar power [5,6]. The integration of these technologies with IES not only strengthens the low-carbon attributes but also enhances economic efficiency and flexibility through multi-energy flow synergy, possessing significant research value for promoting the sustainability, reliability, and economic viability of IES.

Research on the optimal configuration of IES has become an active area of research, with many researchers worldwide focusing on optimization objectives, influencing factors, and system structures. In terms of optimization objectives, early studies primarily prioritized economic viability. For instance, Zheng et al. [7] established a foundational framework to minimize total system costs. As environmental regulations tightened, Wang et al. [8] extended these economic models by incorporating carbon trading mechanisms to address emission constraints. However, considering that cost and emissions are not the only concerns, Han et al. [9] further emphasized system reliability, proposing redundancy strategies to enhance resilience against failures. Regarding influencing factors, existing studies have increasingly incorporated uncertainties, including fluctuations in renewable energy output [10], sudden load changes [11], and adjustments in electricity pricing policies [12]. To mitigate the impact of these uncertainties on configuration results, methods such as scenario clustering and dynamic programming have been widely employed. Concerning system structure, the design scale has expanded from single-region energy systems to multi-park interconnected systems, with frameworks encompassing the collaborative configuration of various equipment such as new energy generation, energy conversion, and storage devices [13,14]. Some studies have also explored optimization of novel structures like combined cooling, heating, and power (CCHP) [15] and shared hybrid energy storage [16]. While recent studies have explored the combination of CCUS and P2G, most treat them as separate rigid modules or focus on unidirectional energy conversion. However, the deep coupling mechanism involving the internal carbon recycling loop—where captured CO_2 is continuously converted back into fuel (methane) via green hydrogen—has not been adequately modeled under high uncertainty.

The application of CCUS technology in energy systems has evolved from standalone carbon capture plants toward the synergistic operation of multiple equipment. This process, which involves post-combustion capture, compression, and storage to reduce carbon emissions, followed by resource utilization that generates additional revenue, significantly enhances the system's low-carbon benefits [17–19]. Liang et al. [20] and Shi et al. [21] successfully utilized deep reinforcement learning and multi-timescale strategies to optimize carbon capture within electrical systems. In parallel, comprehensive hydrogen utilization focuses on the 'production-storage-utilization' chain. Xu et al. [22] and Liu et al. [23] demonstrated that shared storage and seasonal shifting can effectively mitigate renewable fluctuations. However, existing studies often simplify the technological coupling between CCUS and hydrogen systems, and the robustness of the synergistic operation of these two systems under uncertainties such as renewable energy fluctuations remains unexplored.

Traditional robust optimization (RO) typically assumes worst-case scenarios for decision-making [24,25]. While effective for parameter uncertainty, this approach often yields overly conservative solutions that compromise system flexibility and economic benefits. To mitigate this, distributionally robust optimization has gained prominence by introducing probabilistic distribution information into the model, thereby allowing for more informed decisions [26]. This field is rapidly evolving; notably, Zeng et al. [27] recently offered a primal perspective on two-stage DRO, significantly enhancing its algorithmic interpretability and computational tractability, which solidifies the theoretical foundation for applying such frameworks to complex energy systems. However, the efficacy of DRO is largely contingent upon the construction of the ambiguity set. Recent literature broadly categorizes these into moment-based, discrepancy-based, and distance-based sets. Moment-based approaches, while computationally tractable, rely on low-order statistics and often fail to capture the asymmetric and multi-modal characteristics of renewable generation, resulting in ambiguity sets that still contain implausible extreme distributions [28,29]. Similarly, discrepancy-based methods like Kullback-Leibler divergence require continuity [30,31]. They struggle to assign probabilities to unobserved scenarios and often incur high computational complexity when handling continuous data. In contrast, the Wasserstein distance-based approach demonstrates superior adaptability for IES applications [32]. Its primary advantages include asymptotic consistency—guaranteeing convergence to the true distribution with sufficient data—and the ability to account for unobserved but physically possible scenarios by allowing probability shifts around empirical samples [33,34]. This capability is indispensable for systems characterized by strong renewable volatility. In terms of application, DRO has been widely employed in energy system optimization [35,36]. Recent high-impact studies, such as Fan et al. [37], have further demonstrated the model's effectiveness in coordinating flexible resources under spatiotemporal uncertainty. Despite these advances, existing frameworks have not comprehensively addressed the specific "capture-conversion-reutilization" material coupling between hydrogen energy storage and CCUS. This deep coupling is highly sensitive to fluctuations, necessitating a tailored Wasserstein-based DRO framework to optimally balance economic efficiency with the continuity of the internal carbon recycling loop.

To address these challenges, this paper proposes a novel optimal configuration method for IES that incorporates CCUS and comprehensive hydrogen energy utilization. A two-stage DRO model based on the Wasserstein metric is established to optimally balance overall system economy against uncertainties in renewable energy output. The model achieves synergistic optimization of system configuration and scheduling, where the lower-level day-ahead operational dispatch problem supports the upper-level capacity configuration decisions, thereby enhancing overall system performance. The main contributions of this

work are summarized as follows.

(1) A novel IES architecture incorporating an internal carbon recycling loop: We construct a hydrogen-carbon coordinated power system that establishes a "capture-conversion-reutilization" closed-loop. This mechanism physically couples carbon capture with green hydrogen production, transforming carbon emissions from a penalty factor into a schedulable resource.
(2) A tailored two-stage DRO framework for coordinating capacity configuration and operational optimization: To address the fragility of the hydrogen-carbon coupling, we propose a Wasserstein-metric-based DRO model. In this framework, the first stage determines optimal capacity investments, while the second stage optimizes operational scheduling. This structure inherently configures a "safety margin" to ensure material flow continuity under worst-case renewable uncertainties.
(3) An efficient solution strategy addressing the economic-environmental trade-off: We employ the Big-M method and strong duality theory to transform the nonlinear model into a tractable mixed-integer linear programming (MILP) problem. This efficient framework not only ensures computational convergence but also quantitatively balances decarbonization performance with operational costs, rendering the operational objectives feasible in terms of both economic and environmental benefits.

The remainder of this paper is structured as follows. Section 2 elaborates on the overall structure of the proposed IES considering CCUS and hydrogen utilization and presents the modeling of its core components. Section 3 introduces the optimization objective for system configuration and the operational constraints of the equipment. Section 4 establishes the two-stage distributionally robust optimization model based on the Wasserstein distance and details its transformation and solution. Section 5 presents and analyzes the simulation and optimization results. Finally, Section 6 concludes the paper.

2. System description

2.1. System composition

The IES encompasses twelve distinct equipment types. Through electrical buses and natural gas pipelines, the system achieves multi-energy complementarity and a reliable energy supply. It is designed to meet diverse load demands, including electrical, heating, cooling, and hydrogen fuel loads, as depicted in Fig. 1. The electrical power subsystem integrates the wind turbine (WT), photovoltaic array (PV), fuel cell (FC), and battery energy storage (BES). This subsystem supplies electricity to the electrolyzer (EL), CCUS, electric chiller (EC), and electrical load. The hydrogen energy subsystem integrates the EL and hydrogen energy storage (HES). The hydrogen produced supplies the FC and methane reactor (MR), and also meets external hydrogen loads (e.g., for fuel cell vehicles). It is noteworthy that a significant amount of waste heat is generated during the operation of both the FC and EL, accounting for 45% to 60% of their energy input. Consequently, a thermal recovery auxiliary equipment (TR) is incorporated into the thermal subsystem. Together with the gas turbine (GT) and a gas boiler (GB), the TR supplies energy to the absorption chiller (AC) and the thermal load. The natural gas network (NG) and the MR work in concert to supply natural gas to the GT and GB within the natural gas subsystem. Furthermore, the system incorporates CCUS to capture CO_2, primarily from the natural gas subsystem, and direct it to the MR, thereby establishing an internal carbon cycle mechanism that effectively reduces the system's overall carbon emissions.

2.2. Model of the PV and WT

The WT and PV constitute the primary renewable energy sources in the proposed IES. The output power of the WT is determined by the wind speed, while that of the PV is a function of solar irradiance. The models for the WT and PV power generation are presented below.

$$P^{e}_{WT,t}=\begin{cases}0 & v_t \le v_{ci} \text{ or } v_{co} < v_t \\ \dfrac{v_t^3 - v_{ci}^3}{v_r^3 - v_{ci}^3} P_r & v_{ci} < v_t < v_r \\ P_r & v_r \le v_t \le v_{co}\end{cases} \tag{1}$$

$$P^{e}_{PV,t}=\eta^{e}_{PV} S I_{PV,t} \tag{2}$$

where P_r is the rated power of WT. v_r, v_{ci} and v_{co} are the rated, cut-in and cut-out wind speed of WT, respectively. S and $I_{PV,t}$ are the total area of the PV array and solar irradiance, respectively.

2.3. Model of the HCPS module

To enhance hydrogen utilization efficiency, the system integrates the P2G, FC, and CCUS technologies, forming a synergistic module termed

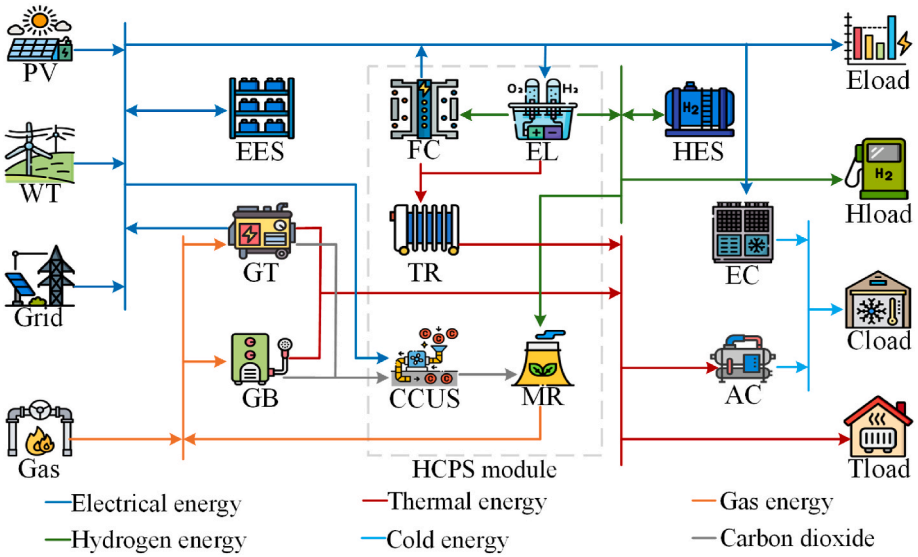

Fig. 1. Structure diagram of the IES.

HCPS, as shown in Fig. 2. To substantively reflect the deep coupling between carbon capture and green hydrogen production, this study establishes a physicochemical linkage model based on mass conservation and reaction stoichiometry, rather than simple energy conversion efficiencies.

The integration of the FC and CCUS fundamentally transforms the system's energy and material flows. Compared to the traditional P2G process, which merely converts electricity unidirectionally into natural gas, the FC enables a bidirectional energy flow by reconverting hydrogen back into electricity, thereby improving hydrogen utilization efficiency and avoiding the energy waste associated with unidirectional conversion. Furthermore, waste heat recovery devices are added to the auxiliary equipment supporting the operation of the FC and EL, reclaiming dissipated heat and further elevating the overall energy utilization efficiency.

$$P^{H_2}_{EL,t} = \eta^{H_2}_{EL} P^{e}_{EL}, t \tag{3}$$

$$P^{h}_{EL,t} = \eta^{h}_{EL} P^{e}_{EL,t} \tag{4}$$

$$P^{e}_{FC,t} = \eta^{e}_{FC} P^{H_2}_{FC,} t \tag{5}$$

$$P^{h}_{FC,t} = \eta^{h}_{FC} P^{H_2}_{FC,} t \tag{6}$$

$$0 \le P^{e}_{EL,t} \le u_{EL} P^{e}_{EL,\max} \tag{7}$$

$$0 \le P^{H_2}_{FC,t} \le u_{FC} P^{H_2}_{FC,\max} \tag{8}$$

$$0 \le u_{EL} + u_{FC} \le 1 \tag{9}$$

The CCUS captures CO_2 generated within the system and supplies it to the MR or conducts carbon sequestration, establishing a carbon recycling mechanism. This process not only reduces carbon emissions but also provides feedstock for methane production, further diversifying the system's energy conversion pathways, which enhances its operational flexibility, stability, and overall sustainability. Unlike traditional models that treat carbon emissions solely as a penalty factor, we introduce the mass flow variable $C^{CO_2}_{t}$ to track the carbon lifecycle. The amount of CO_2 captured by the CCUS unit is physically constrained by the operation of gas-fired units. The operation of the MR is governed by the Sabatier reaction ($CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O$). This implies a strict stoichiometric ratio between hydrogen consumption and carbon utilization. These mathematically enforce the interdependence of the CCUS and P2G, ensuring that the system's ability to accommodate renewable energy via HCPS is physically coupled with its power generation dispatch. The operational constraints of the CCUS and MR are given as follows.

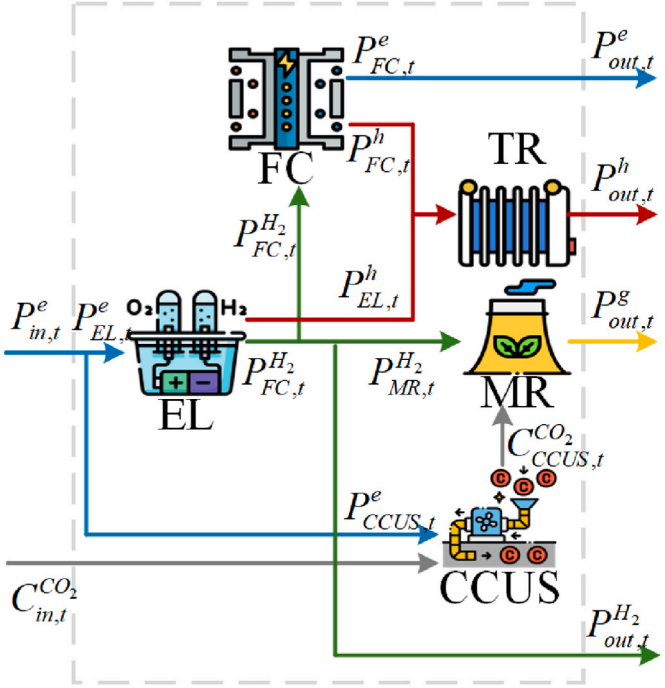

Fig. 2. The synergistic HCPS module.

$$C^{CO_2}_{CCUS,t} = \eta^{CO_2}_{CCUS} P^{e}_{CCUS,} t \tag{10}$$

$$P^{g}_{MR}, t = \eta^{g}_{MR} P^{H_2}_{MR,t} \tag{11}$$

$$C^{CO_2}_{MR,t} = \eta^{CO_2}_{MR} P^{H_2}_{MR,t} \tag{12}$$

$$P^{e}_{CCUS,\min} \le P^{e}_{CCUS,t} \le P^{e}_{CCUS,\max} \tag{13}$$

$$C^{CO_2}_{CCUS,t} \le \eta_{cap} \cdot \left(\eta^{CO_2}_{GT} P^{g}_{GT,t} + \eta^{CO_2}_{GB} P^{g}_{GB,t} \right) \tag{14}$$

$$C^{CO_2}_{CCUS,t} = C^{CO_2}_{MR,t} + C^{CO_2}_{storage,t} \tag{15}$$

$$P^{H_2}_{MR,\min} \le P^{H_2}_{MR,t} \le P^{H_2}_{MR,\max} \tag{16}$$

2.4. Model of the GB and GT

The GB and GT are key components for heat generation within the system. Both devices generate energy through the combustion of natural gas. The fundamental distinction lies in their primary operational outputs: the GB is designed primarily for thermal energy production, whereas the GT is designed for electricity generation, with heat recovered as a useful by-product. The CO_2 emissions from both devices are captured and recycled by the CCUS for further utilization. The operation of both equipment types is required to satisfy the following constraints:

$$P^{h}_{GB,t} = \eta^{h}_{GB} P^{g}_{GB,t} \tag{17}$$

$$P^{e}_{GT,t} = \eta^{e}_{GT} P^{g}_{GT,t} \tag{18}$$

$$P^{h}_{GT,t} = \eta^{h}_{GT} P^{g}_{GT,t} \tag{19}$$

$$P^{g}_{GB,\min} \le P^{g}_{GB,t} \le P^{g}_{GB,\max} \tag{20}$$

$$P^{g}_{GT,\min} \le P^{g}_{GT,t} \le P^{g}_{GT,\max} \tag{21}$$

2.5. Model of the AC and EC

The system's cooling equipment primarily comprises the EC and AC. The AC utilizes waste heat recovered from the GT, GB, EL, and FC to drive its cooling process, which enhances multi-energy complementarity. The EC, in contrast, operates via electric refrigeration to supply cooling loads as required, thereby helping to maintain the overall energy balance of the system. The operation of both chillers is subject to the following constraints:

$$P^{c}_{EC,t} = \eta^{c}_{EC} P^{e}_{EC,t} \tag{22}$$

$$P^{c}_{AC,t} = \eta^{c}_{AC} P^{h}_{AC,t} \tag{23}$$

$$0 \le P^{e}_{EC,t} \le P^{e}_{EC,\max} \tag{24}$$

$$0 \le P^{h}_{AC,t} \le P^{h}_{AC,\max} \tag{25}$$

2.6. Model of the HES and BES

The coordinated operation of the BES and HES allows different components within the IES to compensate for their individual limitations, thereby enhancing the utilization of renewable generation and improving overall system efficiency. Despite their different physical principles, both BES and HES can be described by generic energy storage models based on energy balance and operational state transitions, as formulated below:

$$\begin{cases} SOC_{X,t} = E_{X,t}/E_{X,\max} \\ E_{X,t} = E_{X,t-1} + \eta_{X,ch}P_{X,ch,t} - P_{X,dis,t}/\eta_{X,dis} \\ P_{X,ch,\min} \le P_{X,ch,t} \le u_{X,ch}P_{X,ch,\max} \\ P_{X,dis,\min} \le P_{X,dis,t} \le u_{X,dis}P_{X,dis,\max} \\ 0 \le u_{X,ch} + u_{X,dis} \le 1 \\ E_{X,\min} \le E_{X,t} \le E_{X,\max} \\ E_{X,T} = E_{X,0} \end{cases} \tag{26}$$

2.7. Carbon trading model

Carbon emissions within the IES primarily stem from two sources: indirect emissions from purchased electricity and direct emissions from the combustion of natural gas in the GT and GB. The total system carbon emissions $C^{CO_2}_{sys,t}$ are calculated as follows:

$$C^{CO_2}_{sys,t} = C^{CO_2}_{GT,t} + C^{CO_2}_{GB,t} + C^{CO_2}_{grid,t} - C^{CO_2}_{CCUS,t} \tag{27}$$

$$C^{CO_2}_{GT,t} = \eta^{CO_2}_{GT}P^{g}_{GT,t} \tag{28}$$

$$C^{CO_2}_{GB,t} = \eta^{CO_2}_{GB}P^{g}_{GB,t} \tag{29}$$

$$C^{CO_2}_{grid,t} = \eta^{CO_2}_{grid}P^{e}_{grid,t} \tag{30}$$

Following the common practice in the Chinese carbon market, the initial emission allowance $\widehat{C}^{CO_2}_{sys,t}$ is allocated through a benchmark approach [38,39]. The allocated allowance is determined based on the system's energy output and a predefined carbon emission benchmark, as modeled by the following equation:

$$\widehat{C}^{CO_2}_{sys,t} = \widehat{C}^{CO_2}_{GT,t} + \widehat{C}^{CO_2}_{GB,t} + \widehat{C}^{CO_2}_{grid,t} \tag{31}$$

$$\widehat{C}^{CO_2}_{GT,t} = \widehat{\eta}^{CO_2}_{GT}P^{g}_{GT,t} \tag{32}$$

$$\widehat{C}^{CO_2}_{GB,t} = \widehat{\eta}^{CO_2}_{GB}P^{g}_{GB,t} \tag{33}$$

$$\widehat{C}^{CO_2}_{grid,t} = \widehat{\eta}^{CO_2}_{grid}P^{e}_{grid,t} \tag{34}$$

3. Capacity configuration model of the IES

3.1. Objective function

A low-carbon optimal configuration model for the IES to optimize both economic and environmental performance is established. The model formulates the optimization objective and operational constraints while accounting for uncertainties in wind and photovoltaic power output. The total cost (TC) comprises five components: energy ordering cost (EOC), equipment maintenance cost (EMC), energy curtailment cost (ECC), daily average investment cost (AIC), and carbon emission cost (CEC). The objective function minimizes the total cost, which includes the first-stage annualized investment cost and the expected second-stage operational cost under uncertainty. Let $F(\boldsymbol{x},\widetilde{\boldsymbol{e}})$ represent the minimum operational cost of the second-stage problem, which depends on the first-stage investment decisions $\boldsymbol{x}$ and the realization of uncertainty $\widetilde{\boldsymbol{e}}$. The distribution of $\widetilde{\boldsymbol{e}}$ belongs to a Wasserstein ambiguity set to model the worst-case scenario. Thus, the overall objective function is formulated as follows:

$$\min TC = \min\left\{EOC + EMC + ECC + AIC + CEC + \sup_{\tilde{P}\in\Phi}E_{\tilde{P}}[F(\boldsymbol{x},\widetilde{\boldsymbol{e}})]\right\} \tag{35}$$

$$EOC = \sum_{t=1}^{T}\left[c^{e}_{grid}P^{e}_{grid} + c^{g}_{buy}P^{g}_{buy}\right]\Delta t \tag{36}$$

$$EMC = \sum_{y=1}^{Y}\sum_{t=1}^{T}c_{EMC,y}P_{y}\Delta t \tag{37}$$

$$ECC = \sum_{t=1}^{T}\left[c^{e}_{pv,ce}P^{e}_{pv,ce} + c^{e}_{wt,ce}P^{e}_{wt,ce}\right]\Delta t \tag{38}$$

$$AIC = {}^{1}/_{N_{\text{day}}}\sum_{z=1}^{Z}\frac{r(1+r)^{\mu_z}}{(1+r)^{\mu_z}-1}c_zW_z \tag{39}$$

$$CEC = \sum_{t=1}^{T}c^{CO_2}_{sys}\left[C^{CO_2}_{sys,t} - \widehat{C}^{CO_2}_{sys,t}\right]\Delta t \tag{40}$$

3.2. Constraints

Beyond the component-specific operational constraints presented in Section 2, the system must satisfy a set of energy balance constraints. These include the balance constraints for power, thermal energy, natural gas, hydrogen, and cooling energy. These constraints are formulated as follows:

$$\begin{cases} P^{e}_{grid,t} + P^{e}_{PV,t} + P^{e}_{WT,t} + P^{e}_{FC,t} + P^{e}_{GT,t} + P^{e}_{BES,dis,t} = P^{e}_{EL,t} + P^{e}_{EC,t} + P^{e}_{CCUS,t} + P^{e}_{BES,ch,t} + P^{e}_{load,t} \\ P^{h}_{TR,t} + P^{h}_{GB,t} + P^{h}_{GT,t} = P^{h}_{AC,t} + P^{h}_{load,t} \\ P^{g}_{buy,t} + P^{g}_{MR,t} = P^{g}_{GT,t} + P^{g}_{GB,t} \\ P^{H_2}_{EL,t} + P^{H_2}_{HES,dis,t} = P^{H_2}_{FC,t} + P^{H_2}_{MR,t} + P^{H_2}_{HES,ch,t} + P^{H_2}_{load,t} \\ P^{c}_{AC,t} + P^{c}_{EC,t} = P^{c}_{load,t} \end{cases} \tag{41}$$

4. Construction of wind-solar uncertainty set and DRO model solution

4.1. Modeling of wind and solar uncertainty

Wind and photovoltaic power generation are the primary energy sources for the IES. Wind power uncertainty originates from stochastic fluctuations in wind speed and direction, which directly cause variability in the power output. Similarly, photovoltaic output uncertainty is primarily driven by solar irradiance, which is subject to variations due to factors including cloud cover, seasonal and diurnal cycles, and geographical location. The inherent unpredictability of the WT and PV power output poses a major challenge to the stable and reliable operation of the IES. To address these challenges, this paper employs a two-stage DRO model, which utilizes an ambiguity set constructed via the Wasserstein distance to characterize the uncertainties associated with wind and photovoltaic power generation.

The direct use of raw historical data to construct the ambiguity set would incur a substantial computational burden, thereby reducing the model's solution efficiency. Consequently, an improved K-means clustering method is adopted to generate typical scenarios s for the WT and PV output $\widehat{\boldsymbol{e}}$, along with their corresponding probabilities p_s. This set of scenarios and probabilities constitutes the empirical distribution $\widehat{P}$ within the ambiguity set, achieving significant data reduction while preserving representative scenarios. This method features two core improvements: the selection of initial cluster centers and the convergence criterion. Unlike the traditional K-means algorithm that randomly selects initial centers, this method employs a progressive distance optimization strategy. After randomly selecting the first center K_1, each subsequent center K_n is chosen as the data point that maximizes the sum of distances to all previously selected centers, resulting in a more dispersed and representative initial distribution. For the convergence criterion, the algorithm terminates when the change in the clustering error between consecutive iterations falls below a specified threshold H. The clustering error is defined as the sum of squared distances between each data point and its assigned cluster center. This method, which replaces traditional ones based on center stability or a fixed iteration count, collectively reduces the algorithm's sensitivity to initialization, mitigates the risk of local optima, and enhances the stability and consistency of the clustering results. Using the aforementioned clustering method, typical scenarios for the WT and PV historical data are generated separately. Assuming k_1 and k_2 are the number of clusters for the wind speed and solar irradiance (which are subsequently transformed into power generation outputs via mathematical models of WT and PV), the total number of typical scenarios S is given by $S = k_1 \times k_2$.

In terms of probabilistic metrics, Wasserstein balls provide meaningful, data-driven confidence guarantees: their radius can be calibrated based on sample size and confidence level to yield rigorous out-of-sample performance bounds, while accommodating both discrete and continuous distributions—making them suitable for real-world data-generating processes. In contrast, KL divergence cannot include continuous true distributions when centered at an empirical distribution and thus fails as a valid confidence set, while moment-based ambiguity sets, though computationally simpler, struggle to reliably estimate high-order moments from finite samples and lack comparable finite-sample performance guarantees [40].

The ambiguity set, constructed using the Wasserstein distance, is defined as the set of all probability distributions whose Wasserstein distance to the empirical distribution $\widehat{P}$ is within a constant threshold θ. This set is designed to contain a family of plausible true distributions, including the worst-case distribution that maximizes the expected cost. The integral form of the Wasserstein probability distance is defined as follow [41].

$$W(\widehat{P},\widetilde{P}) = \left[\inf_{Q\in\mathscr{C}(\widehat{P},\widetilde{P})}\int \rho(\widehat{\boldsymbol{e}},\widetilde{\boldsymbol{e}})Q(\mathrm{d}\widehat{\boldsymbol{e}},\mathrm{d}\widetilde{\boldsymbol{e}})\right] \tag{42}$$

Furthermore, the Wasserstein distance can be expressed in an expectation form as:

$$W(\widehat{P},\widetilde{P}) = \left\{\inf_{Q\in\mathscr{C}(\widehat{P},\widetilde{P})} \mathrm{E}_Q[\rho(\widehat{\boldsymbol{e}},\widetilde{\boldsymbol{e}})]\right\} \tag{43}$$

where $\inf(\cdot)$ is the infimum function. $\mathscr{C}(\widehat{P},\widetilde{P})$ is the set of all joint distributions with marginals $\widehat{P}$ and $\widetilde{P}$. $\mathrm{E}_Q(\cdot)$ is the expectation under the joint distribution Q. $\rho(\cdot)$ is the ground distance of the Wasserstein metric and often chosen as a norm. $\widehat{\boldsymbol{e}} = \{\widehat{e}_1, \widehat{e}_2, ..., \widehat{e}_t, ..., \widehat{e}_T\}$ is a vector of the empirical distribution over the scheduling horizon T. $\widetilde{\boldsymbol{e}} = \{\widetilde{e}_1, \widetilde{e}_2, ..., \widetilde{e}_t, ..., \widetilde{e}_T\}$ is the corresponding random vector from the true distribution.

Then, the Wasserstein ambiguity set is then defined as a "ball" of distributions centered at the empirical distribution $\widehat{P}$ with radius θ.

$$\Omega : \{\widetilde{P} \in R(\Xi) | W(\widehat{P},\widetilde{P}) \le \theta(N)\} \tag{44}$$

where $R(\Xi)$ is the set of all probability distributions supported on Ξ. The radius $\theta(N)$ is typically a function of the sample size N and a confidence level, satisfying $\lim_{N\to+\infty}\theta(N) = 0$. This ensures that the empirical distribution converges to the true distribution as more data becomes available, guaranteeing the asymptotic consistency of this data-driven approach.

For computational tractability, the theoretically infinite-dimensional ambiguity set Ω is often approximated by a finite one based on the S generated typical scenarios. This reformulation leverages a scenario-wise conditional distribution and support set to capture the uncertainty structure. The detailed mathematical formulation of this data-driven ambiguity set $\mathscr{F}_W(\theta)$ is given by Eq. (45), where the key idea is to constrain the expected ground distance for each scenario. The probabilities p_s define the center of the Wasserstein ball. They represent the weights of the typical scenarios derived from the improved K-means clustering. The radius θ appears on the right-hand side of the inequality constraint. This constraint explicitly enforces that the total Wasserstein distance between the true distribution and the empirical distribution must not exceed the radius θ.

$$\mathscr{F}_W(\theta) = \left\{ \widetilde{P} \in \Phi\left(\mathbf{R}^{T+1}\times[S]\right) \middle| \begin{array}{l} [(\widetilde{\boldsymbol{e}}_{WT},\widetilde{\nu}_{WT}),\widetilde{s}] \sim \widetilde{P} \\ [(\widetilde{\boldsymbol{e}}_{PV},\widetilde{\nu}_{PV}),\widetilde{s}] \sim \widetilde{P} \\ \mathrm{E}_{\widetilde{P}}[\widetilde{\nu}_{WT}|\widetilde{s}\in[S]] \le \theta \\ \mathrm{E}_{\widetilde{P}}[\widetilde{\nu}_{PV}|\widetilde{s}\in[S]] \le \theta \\ \widetilde{P}[(\widetilde{\boldsymbol{e}}_{WT},\widetilde{\nu}_{WT}) \in \Xi_s|\widetilde{s}=s] = 1, \forall s\in[S] \\ \widetilde{P}[(\widetilde{\boldsymbol{e}}_{PV},\widetilde{\nu}_{PV}) \in \Xi_s|\widetilde{s}=s] = 1, \forall s\in[S] \\ \widetilde{P}[\widetilde{s}=s] = p_s, \forall s\in[S] \end{array} \right\} \tag{45}$$

Where $\widetilde{\boldsymbol{e}}_{WT}$ and $\widetilde{\boldsymbol{e}}_{PV}$ are the true values of the WT and PV output. $\widetilde{\nu}_{WT}$ and $\widetilde{\nu}_{PV}$ are auxiliary random variables. Φ represents the set of true distributions $\widetilde{P}$. $\widetilde{s}$ is the scenario index variable. Ξ_s is the support set for scenario s is defined as:

$$\Xi_s = \left\{(\widetilde{\boldsymbol{e}},\widetilde{\nu}) \in \mathbf{R}^{T+1} | \underline{\widehat{\boldsymbol{e}}} \le \widetilde{\boldsymbol{e}} \le \overline{\widehat{\boldsymbol{e}}}, \rho(\widehat{\boldsymbol{e}}_s,\widetilde{\boldsymbol{e}}) \le \widetilde{\nu}\right\}, s\in[S] \tag{46}$$

Where $\underline{\widehat{\boldsymbol{e}}}$ and $\overline{\widehat{\boldsymbol{e}}}$ are the vectors of minimum and maximum values of the empirical values, respectively. $\widehat{\boldsymbol{e}}_s$ represents the s-th scenario of the empirical value.

The ambiguity set, governed by the radius θ, encapsulates a family of distributions that are plausible given the empirical data. The corresponding distributionally robust optimization model is formulated to be immunized against the worst-case distribution within this set, thereby ensuring system reliability under uncertainty. The radius θ directly

controls the conservativeness of the model. By tuning θ, the decision-maker can constrain the size of the ambiguity set, thereby balancing the robustness of the solution against its economic performance. Consequently, the Wasserstein ambiguity set provides a principled and data-driven framework to effectively capture uncertainty while maintaining computational tractability, even in systems with multiple interdependent uncertain sources, including the IES with wind and photovoltaic power generation.

4.2. Transformation and solution of the two-stage DRO model

In the proposed capacity configuration model, the maximum allowable power in the energy storage operational constraints (Eq. (26)) depends on the capacity configuration. Therefore, it functions as a continuous decision variable rather than a fixed parameter. Consequently, the operational constraint involves a nonlinear term. Taking the discharging state as an instance, this nonlinearity arises from the product of the binary status variable $u_{X,dis}$ and the continuous capacity variable $P_{X,dis,max}$. To linearize this bilinear term, we introduce an auxiliary continuous variable $\widetilde{P}_{X,dis}$ and employ the Big-M method. The original nonlinear constraint is reformulated into the following set of mixed-integer linear inequalities:

$$\begin{cases} \widetilde{P}_{X,dis} = u_{X,dis} P_{X,dis,\max} \\ \widetilde{P}_{X,dis} \le P_{X,dis,\max} \\ \widetilde{P}_{X,dis} \le M u_{X,dis} \\ \widetilde{P}_{X,dis} \ge P_{X,dis,\max} - M\left(1 - u_{X,dis}\right) \end{cases} \tag{47}$$

where M is a sufficiently large constant (set to 10^6 in this study). The charging process of the energy storage systems, as well as the operational constraints of the FC and EL (Eqs. (7) and (8)), involve identical bilinear non-linear terms (products of binary operational status variables and continuous capacity decision variables). These terms are linearized using the same Big-M formulation strategy as derived above and are therefore not repeated here for brevity.

The overall two-stage DRO model for IES configuration can be summarized in the following compact form:

$$\begin{cases} \min\limits_{\boldsymbol{x}} \boldsymbol{c}^T\boldsymbol{x} + \sup\limits_{\tilde{P}\in\mathscr{F}_W} E_{\tilde{P}}[F(\boldsymbol{x},\widetilde{\boldsymbol{e}})] \\ s.t. \quad \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b} \end{cases} \tag{48}$$

While the Wasserstein ambiguity set effectively captures the spatiotemporal correlations of uncertainties, it renders the resulting DRO model computationally intractable for direct solution. To address this, we incorporate a linear affine policy. This policy explicitly parameterizes the second-stage decisions $\boldsymbol{y}$ as affine functions of the uncertainties. This re-parameterization preserves the correlation structure and enables a tractable reformulation.

$$\boldsymbol{y}(\boldsymbol{x},\widetilde{\boldsymbol{e}}) = \boldsymbol{x} + \boldsymbol{\alpha} \odot \widetilde{\boldsymbol{e}}_{\boldsymbol{WT}} + \boldsymbol{\beta} \odot \widetilde{\boldsymbol{e}}_{\boldsymbol{PV}} \tag{49}$$

Where $\boldsymbol{y}$ is the second-stage decision variables. $\boldsymbol{x}$ is the first-stage decision vector. $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are the affine adjustable parameters, respectively.

Consequently, the second-stage objective function and its corresponding constraints can be expressed as:

$$\begin{cases} F(\boldsymbol{x},\widetilde{\boldsymbol{e}}) = \min\limits_{\boldsymbol{x}} \boldsymbol{d}^T\boldsymbol{y}(\boldsymbol{x},\widetilde{\boldsymbol{e}}) \\ s.t. \quad \boldsymbol{E}\boldsymbol{y} \le \boldsymbol{q}(\widetilde{\boldsymbol{e}}) \end{cases} \tag{50}$$

To solve this model, it is reformulated based on strong duality theory. According to this theory, the worst-case expectation $\sup\limits_{\tilde{P}\in\Phi} E_{\tilde{P}}[F(\boldsymbol{x},\widetilde{\boldsymbol{e}})]$ in the model's objective function can be expressed as [42]:

$$\sup_{\tilde{P}\in\Phi} E_{\tilde{P}}[F(\boldsymbol{x},\widetilde{\boldsymbol{e}})] = \inf_{\gamma\ge 0}\left\{\gamma\theta + \sum_{s=1}^{S} \sup_{\tilde{\boldsymbol{e}}\in\Xi} p_s[F(\boldsymbol{x},\widetilde{\boldsymbol{e}}) - \gamma\|\widetilde{\boldsymbol{e}} - \widehat{\boldsymbol{e}}_s\|]\right\} \tag{51}$$

where, γ is the dual variable related to $\inf\limits_{Q\in\mathscr{C}(\widehat{P},\widetilde{P})}\{\mathrm{E}_Q[\rho(\widehat{\boldsymbol{e}},\widetilde{\boldsymbol{e}})]\} \le \theta$ in Φ.

Substituting Eq. (51) into Eq. (48), the configuration optimization model can be equivalently written as:

$$\begin{cases} \min\limits_{\boldsymbol{x}} \boldsymbol{c}^T\boldsymbol{x} + \gamma\theta + \sum\limits_{s=1}^{S} \sup\limits_{\tilde{\boldsymbol{e}}\in\Xi} p_s[F(\boldsymbol{x},\widetilde{\boldsymbol{e}}) - \gamma\|\widetilde{\boldsymbol{e}} - \widehat{\boldsymbol{e}}_s\|] \\ s.t. \quad \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b} \\ \boldsymbol{E}\boldsymbol{y}(\boldsymbol{x},\widetilde{\boldsymbol{e}}) \le \boldsymbol{q}(\widetilde{\boldsymbol{e}}), \forall \widetilde{\boldsymbol{e}} \in \Xi \end{cases} \tag{52}$$

The objective function still contains a min-max problem, making it difficult to solve directly. Therefore, an auxiliary variable ξ_s for each scenario s is introduced, transforming the above expression into the following form [43]:

$$\begin{cases} \min\limits_{\boldsymbol{x}} \boldsymbol{c}^T\boldsymbol{x} + \gamma\theta + \sum\limits_{s=1}^{S} p_s\xi_s \\ s.t. \quad \sup\limits_{\tilde{\boldsymbol{e}}\in\Xi}[F(\boldsymbol{x},\widetilde{\boldsymbol{e}}) - \gamma\|\widetilde{\boldsymbol{e}} - \widehat{\boldsymbol{e}}_s\|] \le \xi_s, \forall s \in [S] \\ \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b} \\ \boldsymbol{E}\boldsymbol{y}(\boldsymbol{x},\widetilde{\boldsymbol{e}}) \le \boldsymbol{q}(\widetilde{\boldsymbol{e}}), \forall \widetilde{\boldsymbol{e}} \in \Xi \end{cases} \tag{53}$$

Because $F(\boldsymbol{x},\widetilde{\boldsymbol{e}})$ is a convex function with respect to $\widetilde{\boldsymbol{e}}$, the optimal solution of $\sup\limits_{\tilde{\boldsymbol{e}}\in\Xi}[F(\boldsymbol{x},\widetilde{\boldsymbol{e}}) - \gamma\|\widetilde{\boldsymbol{e}} - \widehat{\boldsymbol{e}}_s\|]$ is necessarily attained at one of the boundary values (either the lower or upper limit) of the random variable's domain [44]. Furthermore, the model can be represented as:

$$\begin{cases} \min\limits_{\boldsymbol{x}} \boldsymbol{c}^T\boldsymbol{x} + \gamma\theta + \sum\limits_{s=1}^{S} p_s\xi_s \\ s.t. \quad \gamma \ge 0 \\ \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b} \\ \boldsymbol{d}^T\boldsymbol{y}(\boldsymbol{x},\widehat{\boldsymbol{e}}_s) \le \xi_s, \forall s \in [S] \\ \boldsymbol{d}^T\boldsymbol{y}(\boldsymbol{x},\overline{\widetilde{\boldsymbol{e}}}) - \gamma\mathbf{1}^T\rho(\overline{\widetilde{\boldsymbol{e}}},\widehat{\boldsymbol{e}}_s) \le \xi_s, \forall s \in [S] \\ \boldsymbol{d}^T\boldsymbol{y}(\boldsymbol{x},\underline{\widehat{\boldsymbol{e}}}) + \gamma\mathbf{1}^T\rho(\underline{\widehat{\boldsymbol{e}}},\widehat{\boldsymbol{e}}_s) \le \xi_s, \forall s \in [S] \\ \boldsymbol{E}\boldsymbol{y}(\boldsymbol{x},\overline{\widetilde{\boldsymbol{e}}}) - \boldsymbol{q}(\overline{\widetilde{\boldsymbol{e}}}) \le 0 \\ \boldsymbol{E}\boldsymbol{y}(\boldsymbol{x},\underline{\widehat{\boldsymbol{e}}}) - \boldsymbol{q}(\underline{\widehat{\boldsymbol{e}}}) \le 0 \end{cases} \tag{54}$$

The formulation involves continuous and discrete decision variables, with both the objective function and constraints being linear functions of these variables. Thus, the original two-stage distributionally robust optimization model is transformed into a MILP model, which can be solved directly by commercial solvers.

In summary, the proposed system configuration optimization framework comprises three key steps. First, scenario reduction techniques are applied to historical WT and PV data to generate a finite set of typical scenarios forming the empirical distribution, which is used to construct a data-driven ambiguity set via the Wasserstein metric. Second, a two-stage DRO framework is constructed. The first stage then determines the optimal capacity investment and a nominal operational schedule, based on forecast data to maximize economic efficiency. In the second stage, the adjustment actions are optimized against the worst-case distribution within the ambiguity set, ensuring system feasibility and cost-effectiveness even under extreme renewable generation scenarios. The linear affine policy is central to this, enabling flexible device operation that responds to uncertainty realizations, contingent on the first-stage capacity decisions. Finally, the entire model is rendered computationally tractable through a series of transformations: the Big-M method linearizes bilinear terms, strong duality theory reformulates the worst-case expectation, and the affine policy parameterizes the adjustment decisions, leading to a final MILP problem solvable by commercial optimizers. The entire modeling and solution process is summarized in Fig. 3.

5. Case study and result analysis

This section begins with the data description and comparative case

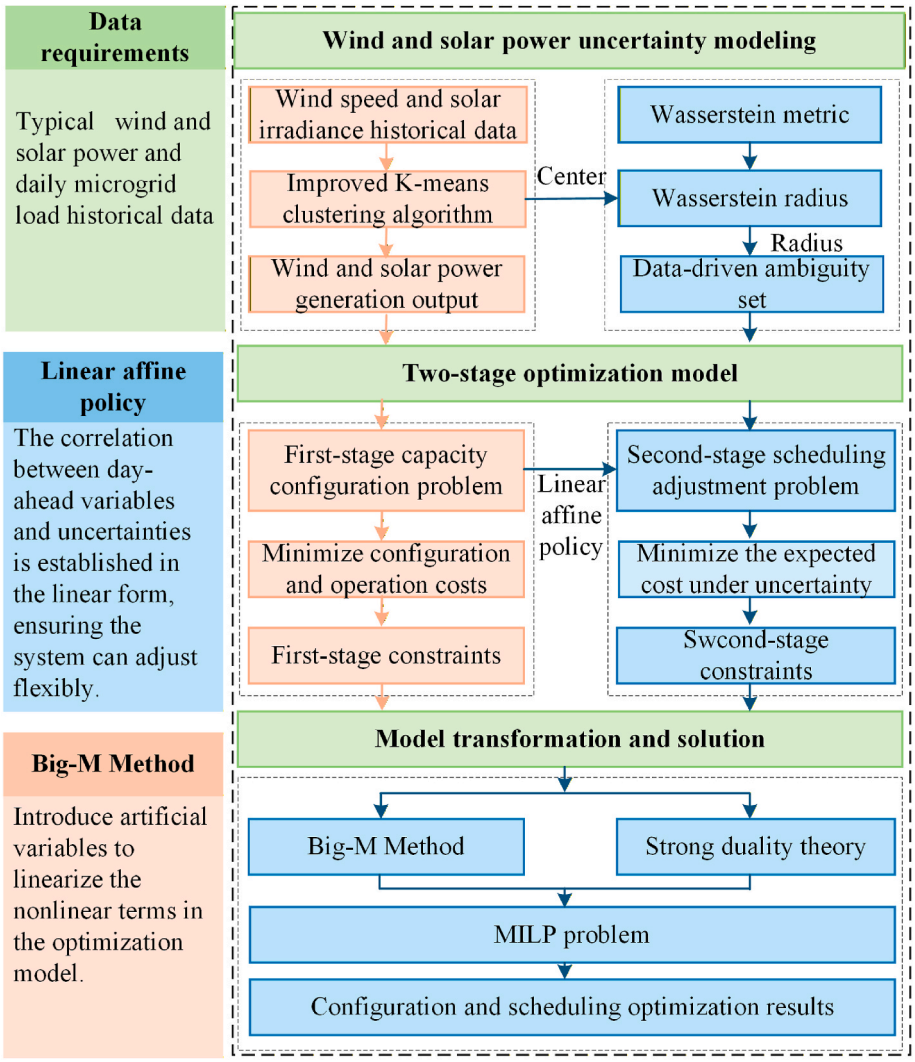

Fig. 3. The framework of modeling and solving procedure.

study to evaluate the technical and economic benefits of integrating CCUS and hydrogen utilization technologies within the IES optimal configuration framework. Subsequently, we analyze the multi-energy complementary scheduling strategies among different energy flows under fluctuating renewable generation, highlighting the system's operational flexibility. Furthermore, a sensitivity analysis is performed on key algorithmic parameters, particularly the radius θ of the Wasserstein ball, to assess the trade-off between robustness and economics in the proposed DRO approach. Finally, the proposed Wasserstein-metric-based DRO model is benchmarked against traditional stochastic programming and robust optimization methods to demonstrate its superior balance of solution robustness and computational efficiency for IES applications.

All simulations and optimization models are implemented in MATLAB R2022b and solved using the Gurobi 10.0.1 solver via the RSOME modeling toolbox [45]. The computations were performed on a computer equipped with an Intel Core i5-12500H CPU (2.50 GHz) and 16 GB of RAM.

Table 1
Technical parameters of the equipment.

Parameters	Value	Parameters	Value
$\eta_{EL}^{H_2}, \eta_{EL}^{h}$	0.7/0.27	η_{EC}^{c}	4
$\eta_{MR}^{g}, \eta_{MR}^{CO_2}$	0.8/0.165	η_{AC}^{c}	1.3
$\eta_{FC}^{e}, \eta_{FC}^{h}$	0.7/0.27	$\eta_{BES,ch}, \eta_{BES,dis}$	0.95/0.95
$\eta_{CCUS}^{CO_2}$	3.8	$E_{BES,\min}, E_{BES,\max}$	0.1/0.9
$\eta_{GB}^{h}, \eta_{GB}^{CO_2}$	0.9/230	$\eta_{HES,ch}, \eta_{HES,dis}$	0.9/0.9
$\eta_{GT}^{e}, \eta_{GT}^{h}, \eta_{GT}^{CO_2}$	0.35/0.5/230	$E_{HES,\min}, E_{HES,\max}$	0.1/0.9

5.1. Data description

A model-based case study is conducted to evaluate the performance of the proposed IES optimal configuration method integrating CCUS and hydrogen utilization. The model encompasses multiple energy flows

Table 2
Economic parameters of the equipment.

Equipment	Installation cost(CNY/kW)	Maintenance cost(CNY/kWh)
PV	4200	0.024
WT	5200	0.017
EL	7000	0.012
MR	1000	0.02
FC	3500	0.012
CCUS	1200	0.01
GB	2300	0.016
GT	2300	0.029
EC	2370	0.018
AC	1600	0.016
BES	2000	0.01
HES	2500	0.01

(electricity, heat, natural gas, hydrogen, and cooling) with distinct units. To ensure consistency and accuracy in the optimization, all relevant parameters were converted to a unified energy basis. The technical parameters of the energy conversion devices, compiled from publicly available project data and equipment specifications, are summarized in Table 1.

A time-of-use electricity price mechanism is implemented for the IES. The prices are 0.48 CNY/kWh (off-peak: 00:00–07:00, 23:00–24:00), 0.88 CNY/kWh (mid-peak: 08:00–11:00, 15:00–18:00), and 1.1 CNY/kWh (on-peak: 12:00–14:00, 19:00–22:00). The natural gas price is 3.89 CNY/m^3, with a calorific value equivalent to approximately 0.35 CNY/kWh. The project economic assumptions include a system lifetime of 20 years and a discount rate of 8%. Economic parameters for equipments are listed in Table 2. Note that the installation costs for renewable energy generation (i.e., PV and WT) are parameterized based on comprehensive EPC prices for coastal projects in Zhejiang (ca. 2023). These values incorporate necessary premiums for structural hardening (e.g., typhoon-resistant mounting) and anti-corrosion treatments required to ensure system reliability in harsh coastal environments [46,47]. For the CCUS system, the levelized cost of carbon capture is estimated to be approximately 400 CNY/ton, which covers investment, energy consumption, and solvent degradation [48].

5.2. Construction of historical uncertain scenarios

This study employs one year of hourly wind speed and solar irradiance data from a coastal demonstration project in Zhejiang, China, as shown in Fig. 4. This dataset is used to derive the empirical distribution for the Wasserstein ambiguity set.

First, we evaluated several commonly used clustering algorithms (using the same original dataset), introducing four performance metrics—relative error (RE), root mean square deviation (RMSD), silhouette coefficient (SC), and computation time—to assess the feasibility of our proposed method for large-scale engineering applications. The results are presented in Table 3.

The experimental results show that K-Shape achieves the best fitting accuracy (in terms of RE and RMSD), but its computational cost is nearly four times that of our proposed method. In contrast, standard K-means is the fastest (approximately 2 s) but yields the poorest accuracy and clustering quality (SC). The improved K-means algorithm proposed in this paper achieves the best overall trade-off: it attains the highest clustering quality (SC), delivers fitting accuracy nearly on par with K-Shape (e.g., the RE for WT differs by only about 0.8%), and maintains computational efficiency (approximately 2.5 s) comparable to that of standard K-means. Consequently, for IES scenario reduction tasks—-where a balance among accuracy, clustering reliability, and computational efficiency is essential—the proposed improved K-means algorithm demonstrates superior engineering applicability.

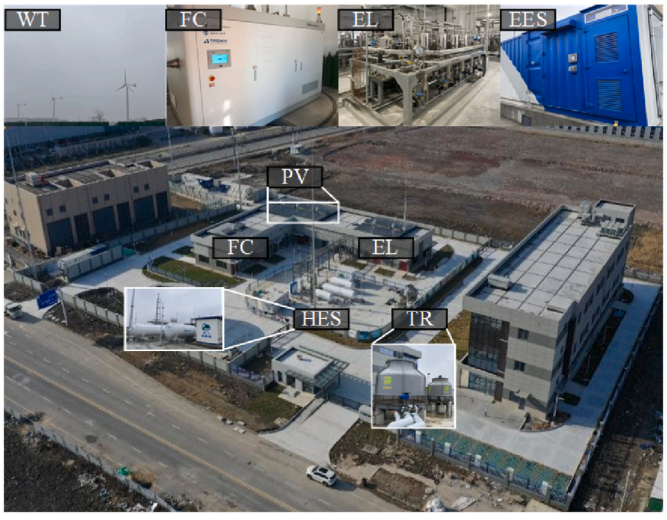

Fig. 4. Actual IES case from a demonstration project in the coastal area of Zhejiang, China.

Based on the improved K-means clustering method, the historical data of wind speed and solar irradiance variations are clustered and reduced, generating 4 scenarios each for WT and PV power generation. Consequently, the ambiguity set contains a total of 16 combined scenarios. The historical maximum and minimum values are selected as boundaries to represent the uncertainty envelope of the historical renewable energy data. The generated wind and solar scenarios and their corresponding probabilities are shown in Figs. 5 and 6, respectively. The wind power scenarios exhibit significant variability and intermittency, which is characteristic of coastal regions. The four scenarios capture distinct patterns, reflecting the unpredictable nature of the wind resource. In contrast, the solar irradiance intensity clearly follow the expected diurnal pattern. While all scenarios share a common shape—rising through the morning, peaking at midday, and declining in the afternoon—they differ in the ramp rates and peak magnitudes, effectively representing a range of typical solar generation days. Furthermore, the scenario probabilities generated by the clustering algorithm align with its effectiveness. Higher-probability scenarios lie closer to the cluster centers, representing the most prototypical patterns of renewable generation. Conversely, lower-probability scenarios are more dispersed, capturing rare but plausible outlier conditions, such as those during extreme weather events. These results collectively demonstrate that the improved K-means method effectively generates a representative set of scenarios for both wind and solar power. The resulting scenarios and their probabilities realistically capture the key characteristics and uncertainties of renewable generation, thereby providing a high-quality, data-driven foundation for the subsequent Wasserstein-based distributionally robust optimization.

5.3. Analysis of configuration optimization and economic performance

This subsection analyzes the configuration results and economic benefits of the proposed IES under different technology integration cases, based on the established WT and PV scenarios.

5.3.1. Case settings

To validate the benefits of the proposed framework, four cases are defined.

Case 1. A basic IES without the P2G, CCUS, and FC.

Case 2. The IES with hydrogen utilization (including the EL and FC).

Case 3. The IES with the HCPS module, operating under a partial carbon recycling strategy.

Case 4. The IES with the HCPS module, operating under a full carbon recycling strategy.

Case 1 represents a typical Integrated Energy System structure and serves as the benchmark for the configuration optimization results. Case 2 introduces a hydrogen energy loop by integrating an electrolyzer and a fuel cell. Cases 3 and 4 both incorporate the proposed HCPS module. The critical distinction is their carbon management strategy: Case 3 employs partial carbon capture and recycling, balancing economic and environmental objectives, whereas Case 4 enforces complete carbon recycling to achieve a zero-carbon operational target.

5.3.2. Capacity configuration analysis

The optimal equipment capacities for all cases are shown in Fig. 7. Case 1, lacking hydrogen utilization and carbon management, is not equipped with the FC, MR, and CCUS system. Due to constrained power and heat generation options and limited hydrogen absorption capacity,

Table 3
The performance of different clustering algorithms.

Method	K-means		K-medoids		K-shape		Proposed	
	WT	PV	WT	PV	WT	PV	WT	PV
RE (%)	15.33	11.51	12.61	10.82	9.45	7.53	10.20	9.19
RMSD	0.121	0.045	0.087	0.033	0.069	0.032	0.083	0.035
SC	0.45	0.54	0.52	0.58	0.58	0.61	0.62	0.65
Computation time (s)	2.1	1.8	7.3	6.8	11.4	10.2	2.5	2.6

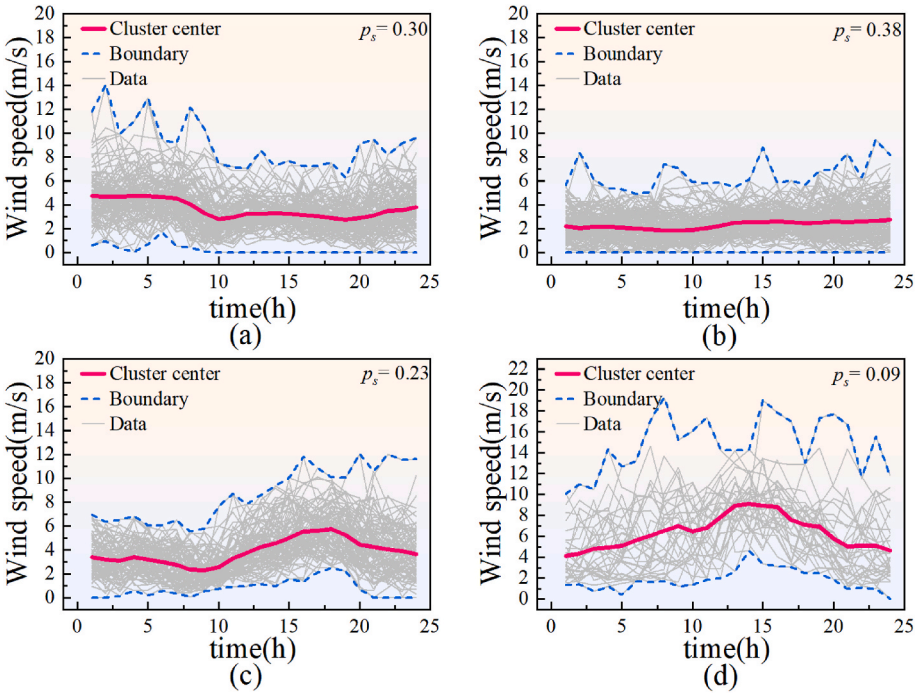

Fig. 5. Wind speed scenario construction and corresponding probabilities.

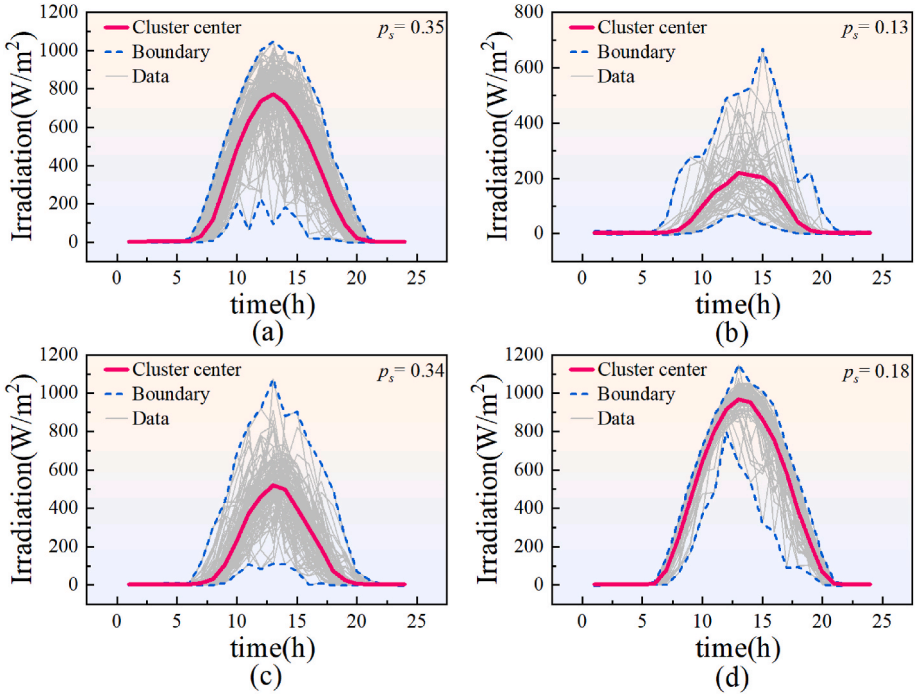

Fig. 6. Solar irradiance scenario construction and corresponding probabilities.

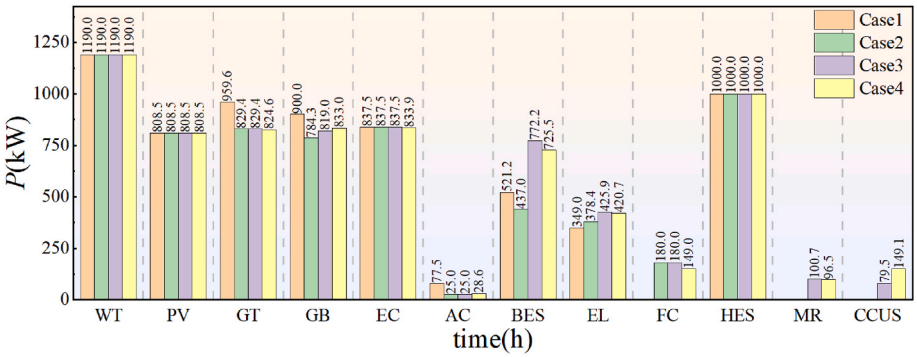

Fig. 7. The IES capacity configuration results.

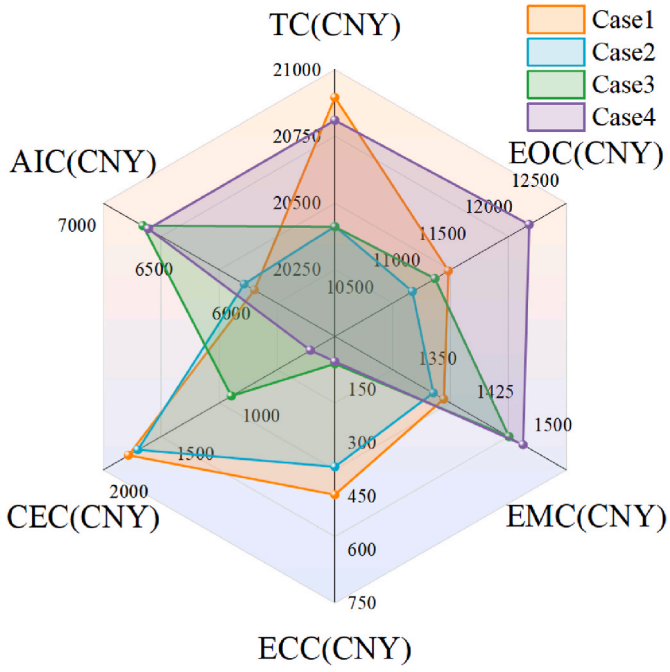

Fig. 8. System configuration optimization results under different cases.

Table 4
Cost components of system configuration optimization under different cases.

Case	TC(CNY)	AIC (CNY)	EOC (CNY)	EMC (CNY)	CEC (CNY)	ECC (CNY)
1	20895.67	6020.63	11227.37	1391.09	1811.14	445.45
2	20412.19	6088.75	10838.91	1377.46	1740.54	366.53
3	20410.95	6743.17	11081.92	1475.95	1033.08	76.83
4	20809.55	6710.18	12099.78	1493.85	435.34	70.41

the solution necessitates larger capacities for the GT, GB, and AC, while the EL capacity remains relatively small. The introduction of hydrogen energy in Case 2 fundamental alters the capacity configuration logic through an energy substitution mechanism. The FC provides an alternative pathway for power and heat supply, effectively substituting part of the traditional fossil-fuel-based generation. Consequently, compared to Case 1, the configured capacities of the GT and GB in Case 2 decrease by 13.6% and 12.9%, respectively. The integration of the HCPS module in Case 3 further highlights the trade-off between equipment investment and system flexibility. While the inclusion of the CCUS and MR enables carbon recycling, it introduces complex energy coupling that requires robust buffering. As a result, the capacity of the BES in Case 3 increases significantly by 76.7% compared to Case 2. This substantial increase indicates that the HCPS-integrated system prioritizes higher investment in flexibility resources (like BES) to accommodate the volatility of renewable energy and the operational constraints of carbon capture, thereby achieving the optimal balance between decarbonization and system stability. Case 4, which mandates full carbon recycling, requires an 87.5% larger CCUS system compared to Case 3. This comes at the expense of reduced capacities for the BES, EL, and FC, reflecting the stringent investment trade-offs under a zero-carbon constraint.

Fig. 8 and Table 4 present the optimized results for various cost components under the different cases. Although Case 1 incurs the lowest investment cost due to its minimal equipment portfolio, its limited flexibility leads to poor renewable energy integration, resulting in the highest wind and solar curtailment cost. The integration of the FC in Case 2 enhances the utilization efficiency of hydrogen and waste heat, thereby reducing the energy purchasing cost, curtailment cost, and daily operating cost. Nevertheless, the absence of carbon management in Cases 1 and 2 results in substantially higher carbon emission costs. In contrast, Cases 3 and 4 demonstrate remarkable improvements in energy efficiency and environmental performance. Compared to Case 1, their wind and solar curtailment costs decrease by 82.8% and 84.2%, respectively, while carbon emission costs drop by 43% and 76%. However, these environmental gains come with higher total costs due to the significant investment and operation expenses of CCUS-related equipment. A comparison between Cases 3 and 4 reveals a critical trade-off. Case 4 prioritizes environmental performance, achieving zero direct carbon emissions but with a higher total cost. Case 3, by tolerating a small amount of emissions, strikes a superior balance, reducing the total daily cost by 2.3% compared to the Case 1 while still delivering significant carbon reductions. This demonstrates the economic-environmental co-benefits of the proposed low-carbon IES configuration with HCPS module.

5.4. Scheduling results analysis

Fig. 9(a)–(e) depict the optimized dispatch schedules for multiple energy carriers in case 3. The electrical power balance shows that the WT and PV are the primary energy sources, supplemented by the FC and GT as auxiliary generation units, collectively supplying power to the EC, CCUS system, and electrical load. During the peak electricity consumption period from 18:00 to 21:00, electricity is purchased from the main grid to ensure supply reliability. Conversely, during periods of excess generation, power is primarily directed to the EL for hydrogen production, thereby enhancing the overall energy efficiency. Regarding the hydrogen balance, the hydrogen load primarily considers daytime demands such as fuel cell vehicles. The EL produces hydrogen to meet the hydrogen load from 06:00 to 20:00. Excess hydrogen produced during other periods is directed to the FC and MR. The HES ensures the hydrogen balance through flexible hydrogen charging and discharging operations. The demand for natural gas by the GB and GT is relatively high and is largely met through external purchases. However, the

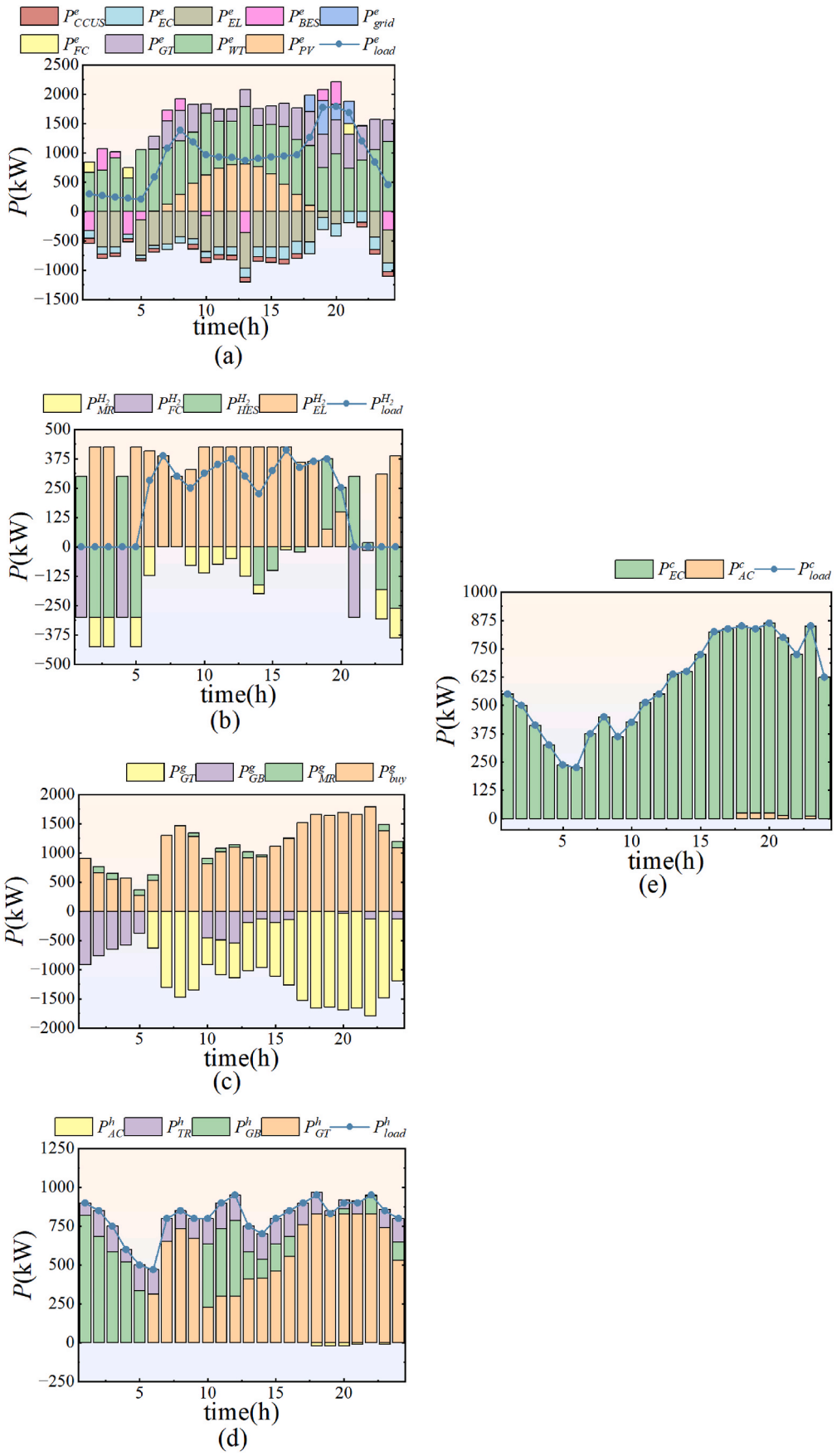

Fig. 9. Optimal dispatch results for (a) electrical power, (b) hydrogen, (c) natural gas, (d) thermal energy, (e) cooling energy.

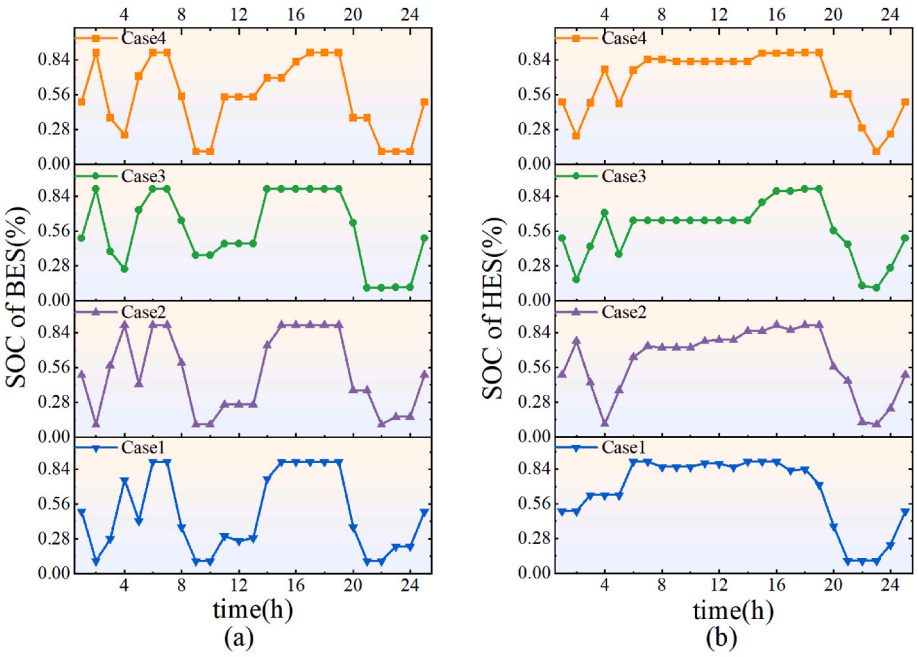

Fig. 10. Hybrid hydrogen-electric energy storage dispatch optimization results, (a) SOC of BES, (b) SOC of HES.

synergistic operation of the CCUS and MR enables internal production of a portion of the natural gas via carbon recycling. This establishes a closed-loop "capture–conversion–reutilization" cycle for the carbon emissions from the GT and GB, significantly advancing system decarbonization. The thermal load is primarily met by the GT and GB, with waste heat recovery from the FC and EL serving as a supplementary source. For cooling, the EC is dispatched first due to its higher economic efficiency (fewer energy conversion steps), while the AC supplies the remainder. The dispatch strategies for the battery and hydrogen storage systems differ fundamentally. The BES, leveraging its rapid response capability, undergoes frequent charge/discharge cycles to maintain real-time power balance. In contrast, the HES capitalizes on its long-duration storage capability, operating with fewer start-stop cycles over longer periods, which is characteristic of the EL and FC.

The state of charge (SOC) profiles for both storage systems across the four cases are compared in Fig. 10 to further elucidate the role of hybrid hydrogen-electric energy storage. A horizontal comparison of the dispatch results from Case 1 to Case 4 reveals a consistent charging and discharging strategy for the BES across the cases. During the night (00:00–05:00), when renewable generation and system load are low, the BES charges and discharges to maintain system power balance. In the early morning (06:00–09:00), as load demand increases, the BES discharges rapidly to meet the demand. Increased renewable generation during the midday to afternoon period allows the BES to charge. Subsequently, during the evening peak (19:00–22:00), when various load demands surge sharply, the BES discharges again to supply power to the system. Notably, Cases 3 and 4, with their larger configured BES capacities, demonstrate superior regulation capability due to the enhanced flexibility afforded by the HCPS module. The configured capacity of the HES is identical across all four cases, and each case can release hydrogen to supply the system during the evening peak period. During nighttime hours, Cases 2 to 4 can frequently charge and discharge hydrogen to support the system, whereas Case 1, lacking hydrogen conversion equipment, participates minimally in dispatch, demonstrating insufficient flexibility. A critical difference emerges during daytime hours. The HES in Cases 1 and 2 remains nearly full, indicating no available margin for additional hydrogen storage. In contrast, Cases 3 and 4 maintain a substantial buffer, allowing for further charging. This indicates that the HCPS in Cases 3 and 4, by consuming hydrogen for carbon recycling, creates higher hydrogen energy throughput and thus achieves superior overall energy utilization efficiency. Therefore, the coordinated operation of the hybrid hydrogen-electric energy storage successfully leverages the complementary characteristics of both technologies. Crucially, the scheduling results highlight the unique hydrogen-carbon coupling characteristics of the proposed HCPS module. Unlike decoupled systems where power shortages merely lead to load shedding or purchasing electricity, a shortage of green hydrogen in this system directly interrupts the methanation process (MR), breaking the internal carbon recycling loop. The substantial buffer observed in the HES of Case 3 and 4 is not merely for energy arbitrage, but serves as a vital 'material buffer'

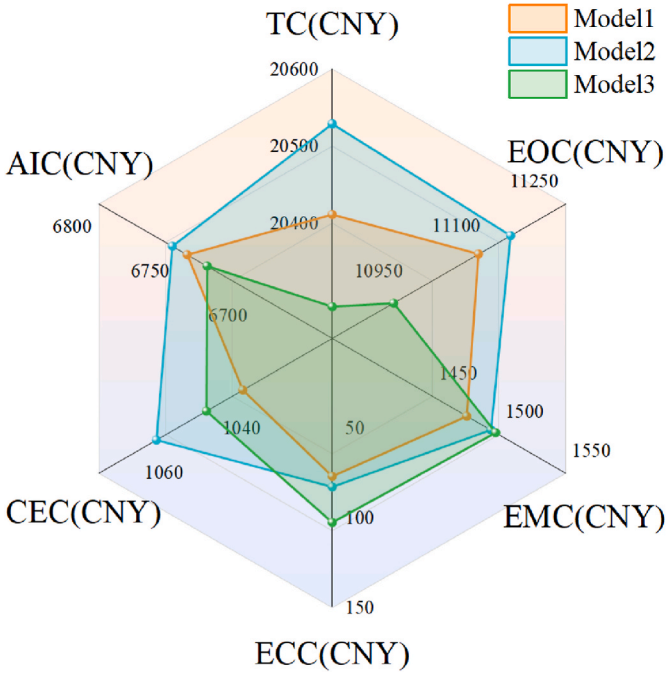

Fig. 11. System configuration optimization costs for different algorithmic models.

to ensure the continuity of the chemical carbon fixation process against renewable volatility.

5.5. Methods comparison

5.5.1. Algorithmic performance comparison

To demonstrate the superiority of the proposed Wasserstein-based DRO model over RO and stochastic optimization (SO) and to systematically validate its effectiveness, a comprehensive comparative analysis is conducted from three perspectives—economic performance, operational robustness, and computational efficiency—through the establishment and evaluation of three distinct models for the IES.

Model 1. The proposed method using the Wasserstein ambiguity set.

Model 2. A RO method employing a budget uncertainty set [49].

Model 3. A SO method that uses clustered historical scenarios as the true distribution [50].

Simulation results indicate significant differences in the total cost and cost structure among the different optimization methods, as shown in Fig. 11. Although Model 3 achieves the lowest total cost, this result is critically dependent on the assumption of accurate probability distributions, leading to insufficient reliability in practical uncertain environments. In contrast, the proposed Model 1 incurs a total cost only 0.59% above Model 3, while outperforming Model 2 with a total cost 0.57% lower. More importantly, Model 1 demonstrates superior performance in key operational and environmental metrics. Its wind and solar curtailment cost is 7.1% and 25.1% lower than that of Model 2 and Model 3, respectively, demonstrating its superior capability to handle renewable volatility and enhance utilization. Simultaneously, Model 1 attains the lowest carbon emission cost, 2.1% below that of Model 2, underscoring its better environmental performance.

This comparison reveals the specific application value of the DRO framework for the rigid hydrogen-carbon coupling constraint, where the stability of the downstream carbon cycle relies heavily on the upstream renewable supply. Model 3 (SO), by prioritizing average scenarios, fails to account for tail risks; under extreme low-wind conditions, it lacks sufficient hydrogen reserves to sustain the methanation reactor, leading to the interruption of the carbon recycling loop and a spike in emissions. Conversely, Model 2 (RO) adopts an overly conservative boundary approach, resulting in economically prohibitive capacity investments and excessive grid reliance. Model 1 (DRO) effectively resolves this dilemma by identifying the worst-case probability distribution that specifically threatens the material flow continuity. By incorporating this distributional information, it inherently configures a precise 'safety margin' for the electrolyzer and hydrogen storage—neither as dangerously thin as Model 3 nor as wasteful as Model 2. Therefore, the DRO serves not merely as a cost-optimization tool, but as a necessary reliability guarantee for maintaining the continuous operation of the internal carbon recycling loop under uncertainty.

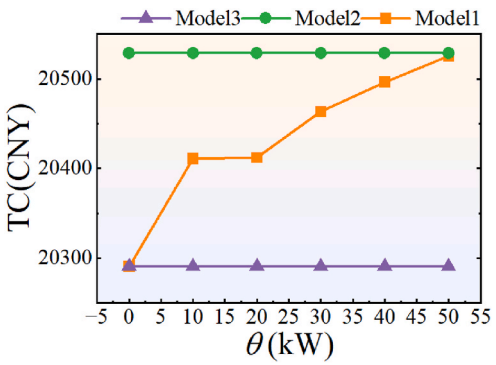

Fig. 12. Influence of parameter θ on system optimization costs.

5.5.2. Influence of Wasserstein radius

A sensitivity analysis is conducted to examine the impact of the Wasserstein ball radius θ on the total system cost in the proposed DRO model (Model 1). Based on our engineering experience, we selected a series of values of θ. The calculation results are shown in Fig. 12 and Table 5. When $\theta = 0$, the ambiguity set contains only the empirical distribution causing the DRO model to degenerate into a SO model equivalent to Model 3. This represents an overly optimistic strategy that minimizes cost under the nominal distribution but offers no robustness guarantee. As the θ increases, the optimal cost of Model 1 increases monotonically and approaches that of the classical RO (Model 2), reflecting an increasingly conservative and risk-averse decision stance. Therefore, the radius θ provides a crucial tuning parameter for decision-makers to effectively navigate the trade-off between economic performance (low θ) and operational robustness (high θ).

Due to the model structure and optimization mechanism, increasing θ implies that the decision-maker considers a broader range of more adverse probability distributions; consequently, during the optimization process, the system tends to select technology portfolios and capacity configurations that remain feasible across a wider spectrum of uncertainty. Essentially, θ directly influences the planning and operation outcomes of the IES by adjusting the decision-maker's perceived boundary of distributional uncertainty, thereby providing a qualitative basis for tuning decisions—from risk tolerance to investment orientation.

5.5.3. Influence of sample sizes

Fig. 13 demonstrates the effect of historical data volume on the optimization results and computation time for Model 1 and Model 3. As the sample size increases, the total cost for both models gradually declines, indicating that a larger number of samples enhances decision-making precision. This trend highlights the data-driven nature of both DRO and SO methods, as a larger dataset better approximates the true uncertainty distribution. Across different sample sizes, the total cost of Model 1 is consistently slightly higher than that of Model 3, though the difference is marginal and stabilizes as the sample size increases. Furthermore, the declining trend of Model 1 aligns closely with Model 3, demonstrating its comparable data-driven effectiveness. Notably, Model 1 exhibits a decisive advantage in computational efficiency. As the sample size grows from 100 to 2000, the computation time for Model 1 increases only modestly from 11.57 s to 17.53 s (a 1.5-fold increase). In contrast, the solving time for Model 3 surges from 14.73 s to 175.65 s (a 12-fold increase). This efficiency is attributed to the model transformation techniques employed to reduce the solving complexity of the DRO model. It confirms that the proposed method retains the data-driven performance while offering greater robustness and far superior computational efficiency for large-scale problems.

6. Conclusion

This study proposes an IES model integrating a synergistic HCPS module, establishing a closed-loop "carbon capture-hydrogen energy conversion-multi-energy complementarity" system. This framework overcomes the limitations of traditional IES designs, which often focus solely on single energy equipment or independent low-carbon modules. Simulation results demonstrate that the proposed IES achieves an 82.7% reduction in renewable curtailment cost, a 43% reduction in carbon emission cost, and a 2% decrease in total daily cost compared to the baseline IES configuration. These results highlight the significant advantages of the proposed modeling framework in simultaneously enhancing the low-carbon performance and renewable energy accommodation capacity of the system. To address the challenges posed by the randomness and volatility of wind and photovoltaic power output on IES stability, an improved K-means clustering algorithm was developed for scenario reduction. Furthermore, a two-stage distributionally robust optimization framework based on the Wasserstein ambiguity set was

Table 5
Total system cost (CNY) under different values of parameter θ(kW).

Model \ θ	0	10	20	30	40	50
1	20290.95	20410.95	20412.15	20463.1	20496.25	20524.95
2	20528.75	20528.75	20528.75	20528.75	20528.75	20528.75
3	20290.95	20290.95	20290.95	20290.95	20290.95	20290.95

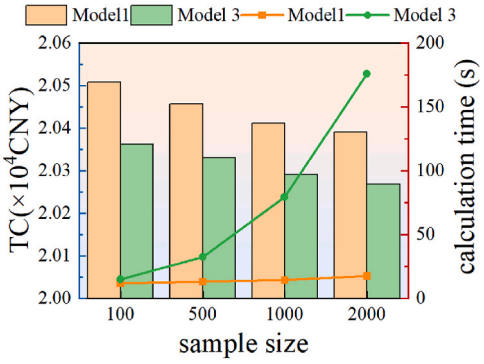

Fig. 13. System optimization results and computation time under different sample sizes.

proposed, effectively achieving the collaborative optimization of IES configuration and operational scheduling. Critically, the DRO framework ensures the robustness of the material flow within the hydrogen-carbon coupling process, preventing the interruption of the carbon recycling loop caused by renewable fluctuations. The results indicate that compared to the traditional robust optimization algorithm, the proposed method reduces the total cost by 0.57%. Moreover, when the sample size increases to 2000, the solution time is only 17.53 s, which demonstrates superior scalability and computational efficiency, making it particularly suitable for optimization problems under large-scale uncertainty conditions.

Concurrently, this study is subject to certain limitations regarding its scalability. While the proposed framework is validated using a single-region case, extending it to multi-region or transnational energy grids presents significant challenges. Specifically, the expansion to a multi-node network would precipitate an exponential growth in binary variables and constraints within the two-stage DRO framework, potentially leading to a 'curse of dimensionality.' Moreover, such an extension requires rigorous modeling of physical network constraints, including tie-line transmission capacities and power flows, alongside the management of policy heterogeneity regarding carbon pricing and market mechanisms across different jurisdictions. To overcome these barriers, future research will investigate advanced distributed algorithms, such as the alternating direction method of multipliers (ADMM) or Benders decomposition, to decouple large-scale optimization problems and alleviate the 'curse of dimensionality.' This will be complemented by refined network modeling that explicitly accounts for inter-regional transmission limits, alongside the application of game-theoretic approaches to simulate stakeholder interactions and benefit allocation under heterogeneous policy environments.

CRediT authorship contribution statement

Zhewei Wang: Writing – review & editing, Writing – original draft, Software. **Jie Yuan:** Software. **Yang Li:** Supervision. **Leiqi Zhang:** Project administration. **Changjun Xie:** Investigation, Funding acquisition. **Wenchao Zhu:** Methodology. **Yang Yang:** Methodology. **Han Wang:** Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This paper is supported by the National Natural Science Foundation of China (NSFC, No. U24B20103) and the China Scholarship Council (No. 202406950113).

Data availability

Data will be made available on request.

References

[1] GlobalEnergyReview2025.pdf. International Energy Agency; 2025.
[2] Hu F, Zhou D, Zhu Q, Wang Q. How dynamic renewable portfolio standards affect trading behavior of power generators? Considering green certificate and reward/penalty mechanism. Appl Energy 2024;375:124114.
[3] Das P, Kayal P. An advantageous charging/discharging scheduling of electric vehicles in a PV energy enhanced power distribution grid. Green Energy Intell Transp 2024;3:100170.
[4] Thulasingam M, Periyanayagam ADVR. Radial distribution systems performance enhancement through RE (renewable energy) integration and comprehensive contingency ranking analysis. Green Energy Intell Transp 2025;4:100245.
[5] Lu T, Yi X, Li J, Wu S. Collaborative planning of integrated hydrogen energy chain multi-energy systems: a review. Appl Energy 2025;393.
[6] Xuan A, Shen X, Luo Y. Bi-level integrated electricity and natural gas system retrofit planning model considering carbon capture, utilization and storage. Appl Energy 2025;385:125476.
[7] Zheng B, Wu X. Integrated capacity configuration and control optimization of off-grid multiple energy system for transient performance improvement. Appl Energy 2022;311:118638.
[8] Wang L, Shi Z, Dai W, Zhu L, Wang X, Cong H, et al. Two-stage stochastic planning for integrated energy systems accounting for carbon trading price uncertainty. Int J Elec Power 2022;143:108452.
[9] Han Z, Zhang J, Zhang H. Optimization design research of combined cooling, heating, and power system based on novel redundancy strategies. Energy 2025;316:134489.
[10] Chen Y, Xu J, Wang J, Lund PD, Wang D. Configuration optimization and selection of a photovoltaic-gas integrated energy system considering renewable energy penetration in power grid. Energy Convers Manag 2022;254:115260.
[11] Yang M, Liu Y. A two-stage robust configuration optimization framework for integrated energy system considering multiple uncertainties. Sustain Cities Soc 2024;101:105120.

[12] Shang J, Gao J, Jiang X, Liu M, Liu D. Optimal configuration of hybrid energy systems considering power to hydrogen and electricity-price prediction: a two-stage multi-objective bi-level framework. Energy 2023;263:126023.
[13] Zhang Z, Jiang P, Liu Z, Fu L, Wang P. Capacity optimal configuration and collaborative planning of multi-region integrated energy system. Energy 2023;278:127970.
[14] Sun X, Chen J, Yang S, Pan X, Gu X, Guo J. Optimal scheduling of multi-regional integrated energy systems with shared energy storage under dynamic rental prices mechanism. J Energy Storage 2025;106:114902.
[15] Zeng R, Ni Z, Tang X, Zhang X, Li H, Zhang G. Optimization of PVT- AFC -CCHP system configuration under multi-climate zones. Int J Hydrogen Energy 2025;102:1210–30.
[16] Pu Y, Li Q, Huo S, Breaz E, Chen W, Gao F. Optimal configuration for shared electric-hydrogen energy storage for multiple integrated energy systems with Mobile hydrogen transportation. Renew Energy 2024;237:121828.
[17] Ni Q, Xu C, Cai Z, Zhou Y, Lai LL. Optimal dispatch of integrated energy system with CCUS-P2G coupling and hydrogen-doped gas equipment based on ladder-type carbon trading mechanism. Energy 2025;336:138397.
[18] Sheng X, Lin S, Liang W, Xie H, Liu M. Optimal long-term planning of CCUS and carbon trading mechanism in offshore-onshore integrated energy system. Appl Energy 2025;379:124983.
[19] Liu W. Towards power-to-X and carbon capture technologies for decarbonized green cities: a multi-layer cost-effective energy scheduling approach in integrated energy systems for targeting carbon footprint reduction. Energy 2025;335:137436.
[20] Liang T, Chai L, Tan J, Jing Y, Lv L. Dynamic optimization of an integrated energy system with carbon capture and power-to-gas interconnection: a deep reinforcement learning-based scheduling strategy. Appl Energy 2024;367:123390.
[21] Shi Z, Yang Y, Xu Q, Wu C, Hua K. A low-carbon economic dispatch for integrated energy systems with CCUS considering multi-time-scale allocation of carbon allowance. Appl Energy 2023;351:121841.
[22] Xu F, Li X, Jin C. Optimal capacity configuration and dynamic pricing strategy of a shared hybrid hydrogen energy storage system for integrated energy system alliance: a bi-level programming approach. Int J Hydrogen Energy 2024;69:331–46.
[23] Liu X, Zu L, Wei Z, Wang Y, Pan Z, Xiao G, et al. Two-layer optimal scheduling of integrated electric-hydrogen energy system with seasonal energy storage. Int J Hydrogen Energy 2024;82:1131–45.
[24] Mulvey JM, Vanderbei R. Robust optimization of large-scale systems: an emerging new technology. Oper Res: J Oper Res Soc Am 1995;43.
[25] Bertsimas D, Sim M. The price of robustness. Oper Res 2004;52:35–53.
[26] Steve Zymler, Daniel Kuhn, Ber Rustem. Distributionally robust joint chance constraints with second-order moment information. Math Program 2011;137:167–98.
[27] Lu Z, Zeng B. Two-stage distributionally robust optimization: intuitive understanding and algorithm development from the primal perspective. 2024.
[28] Wang Y, Song M, Jia M, Li B, Fei H, Zhang Y, et al. Multi-objective distributionally robust optimization for hydrogen-involved total renewable energy CCHP planning under source-load uncertainties. Appl Energy 2023;342:121212.
[29] Han F, Zeng J, Lin J, Gao C. Multi-stage distributionally robust optimization for hybrid energy storage in regional integrated energy system considering robustness and nonanticipativity. Energy 2023;277.
[30] Wang Q, Xiao Y, Tan H, Mohamed MA. Day-ahead scheduling of rural integrated energy systems based on distributionally robust optimization theory. Appl Therm Eng 2024;246.
[31] He C, Hu Z, Liu X, Nan L, Liu T. Coordinated expansion planning of integrated electricity and natural gas distribution systems for enhancing both inherent and dynamic resilience: a distributionally robust approach considering decision-dependent uncertainty. Energy 2025;333.
[32] Chen Z, Kuhn D, Wiesemann W. Technical Note—data-driven chance constrained programs over wasserstein balls. Oper Res 2024;72.
[33] Ma M, Long Z, Liu X, Lee KY. Distributionally robust optimization of electric–thermal–hydrogen integrated energy system considering source–load uncertainty. Energy 2025;316.
[34] Ma K, Feng Y, Yang J, Cai Y, Yang B, Guan X. Two-stage distributionally robust optimization for port integrated energy system with berth allocation under uncertainties. Energy 2025;322.
[35] Kong F, Mi J, Wang Y. A two-stage distributionally robust optimization model for optimizing water-hydrogen complementary operation under multiple uncertainties. J Clean Prod 2022.
[36] Wang Y, Yang Y, Fei H, Song M, Jia M. Wasserstein and multivariate linear affine based distributionally robust optimization for CCHP-P2G scheduling considering multiple uncertainties. Appl Energy 2022;306.
[37] Fan W, Fan Y, Liu P, Wang Y, Tong F, Yi B, et al. Distributionally robust optimization scheduling model for electric power and computing power coordination considering spatiotemporal response. Appl Energy 2025;402:126895.
[38] Zhen J, Liu X, Wu C, Ji L, Huang G. Operation optimization and performance evaluation of photovoltaic-wind-hydrogen-based integrated energy system under carbon trading mechanism and uncertainty for urban communities. J Clean Prod 2024;476.
[39] Fan W, Ju L, Tan Z, Li X, Zhang A, Li X, et al. Two-stage distributionally robust optimization model of integrated energy system group considering energy sharing and carbon transfer. Appl Energy 2023;331.
[40] Mohajerin Esfahani P, Kuhn D. Data-driven distributionally robust optimization using the wasserstein metric: performance guarantees and tractable reformulations. Math Program 2017;171:115–66.
[41] Dou X, Anitescu M. Distributionally robust optimization with correlated data from vector autoregressive processes. Oper Res Lett 2019;47:294–9.
[42] Wu Q, Li C. Modeling and operation optimization of hydrogen-based integrated energy system with refined power-to-gas and carbon-capture-storage technologies under carbon trading. Energy 2023;270:126832.
[43] Huang H, Liang R, Lv C, Lu M, Gong D, Yin S. Two-stage robust stochastic scheduling for energy recovery in coal mine integrated energy system. Appl Energy 2021;290:116759.
[44] Zhao C, Guan Y. Data-driven risk-averse stochastic optimization with wasserstein metric. Oper Res Lett 2018;46:262–7.
[45] Chen Z, Sim M, Xiong P. Robust stochastic optimization made easy with RSOME. Manag Sci 2020;66:3329–39.
[46] China Photovoltaic Industry Development Roadmap. 2022-2023 edition ed. Beijing: China photovoltaic industry Association(CPIA). 2023.
[47] Djalab A, Djalab Z, El Hammoumi A, Marco Tina G, Motahhir S, Laouid AA. A comprehensive review of floating photovoltaic systems: tech advances, marine environmental influences on offshore PV systems, and economic feasibility analysis. Sol Energy 2024;277.
[48] Liu G, Cai B, Li Q, Zhang X, Ouyang T. China's pathways of CO2 capture, utilization and storage under carbon neutrality vision 2060. Carbon Manag 2022;13:435–49.
[49] Tostado-Véliz M, Rezaee Jordehi A, Fernández-Lobato L, Jurado F. Robust energy management in isolated microgrids with hydrogen storage and demand response. Appl Energy 2023;345:121319.
[50] Mei F, Zhang J, Lu J, Lu J, Jiang Y, Gu J, et al. Stochastic optimal operation model for a distributed integrated energy system based on multiple-scenario simulations. Energy 2021;219:119629.