

Optimal operation of distribution networks under asymmetric renewable energy and load demand uncertainties

Zhisheng Xiong^{a,*} , Bo Zeng^b, Peter Palensky^a , Pedro P. Vergara^a 

^a Intelligent Electrical Power Grids (IEPG) group, Delft University of Technology, 2628 CD, Delft, Netherlands

^b Department of Industrial Engineering, University of Pittsburgh, Pittsburgh, PA 15261, USA

ARTICLE INFO

Keywords:

Information gap decision theory

Probability

Renewable energy

Asymmetric uncertainties

Distribution systems

ABSTRACT

To develop an optimal operational scheme for distribution networks capable of addressing asymmetric uncertainties associated with renewable energy and load demands, this paper presents a confidence level-based information gap decision theory (CL-IGDT) framework. Building on IGDT, the proposed framework utilizes the confidence level to capture the asymmetric characteristics of uncertainties and maximize the risk-averse capability of the solution in a probabilistic manner. To facilitate such probabilistic consideration, the imprecise Dirichlet model is employed to construct the ambiguity sets of uncertainties. Consequently, a two-stage robust optimal operation model for distribution networks using CL-IGDT is developed. An iterative method is proposed to solve the model and determine the upper and lower bounds of the objective function. Case study demonstrates that the proposed approach yields a more robust and statistically optimized solution with required accuracy compared to existing methods, contributing to a reduction in first-stage cost by 0.84%, second-stage average cost by 6.7%, and significantly increasing the reliability of the solution by 8%.

1. Introduction

Distribution networks (DNs) are increasingly integrating high shares of renewable energy sources (RESs). The intermittent and volatile nature of RESs, along with the uncertainty in load demands, can affect both the economics and security of the networks [1]. Failure to appropriately address uncertainty may increase operational costs due to actions such as load shedding, RESs curtailment, and the need for excessive reserves. Moreover, potential risks such as voltage violations and line overloading may occur. These significant challenges to the optimal operation of DNAs may undermine the transition to more sustainable energy systems [2]. Accordingly, numerous studies have been conducted to manage the uncertainties associated with RESs and load demands, aiming to

develop robust operational strategies that enhance the cost-effectiveness and reliability of DNAs.

Most research on the optimal operation of DNAs under uncertainty relies on varying levels of information about the random variables: complete, partial, or none at all. Ref. [3] models the wind power loss and line overloading as chance constraints using a truncated normal distribution, establishing a two-stage stochastic programming (SP) model for unit commitment. Ref. [4] presents a two-stage stochastic dynamic economic dispatch model that effectively addresses wind power uncertainty. This model improves cost efficiency and reliability by pre-dispatching generator output to avoid network congestion in the first stage and re-dispatching resources after the realization of the wind scenarios. Ref. [5] incorporates multiple correlations of RESs through

* Corresponding author.

Email address: zhisheng.energy@gmail.com (Z. Xiong).

Nomenclature**Sets**

\mathcal{N}	Set of nodes of the distribution network
\mathcal{L}	Set of lines
\mathcal{G}	Set of nodes where DG units are connected, $\mathcal{G} \subset \mathcal{N}$
\mathcal{B}	Set of nodes where ESS units are connected, $\mathcal{B} \subset \mathcal{N}$
\mathcal{T}	Set of time periods

Indices

ij	Line $ij \in \mathcal{L}$
i, j	Nodes $i, j \in \mathcal{N}$
t	Time period $t \in \mathcal{T}$

Parameters

c_i, \hat{c}_i	First-/second-stage DG operational cost coefficients
d_i, \hat{d}_i	First-/second-stage electricity purchase cost coefficients
$\bar{e}_i, \underline{e}_i$	Maximum/minimum energy capacity of ESSs
$\bar{p}^{B+}, \bar{p}^{B-}$	Maximum charging/discharging power of ESSs
η_B	Charging/discharging efficiency of ESSs

$\bar{p}_i^G, \underline{p}_i^G$	Maximum/minimum active power of DGs
$\bar{\hat{p}}_i^G, \underline{\hat{p}}_i^G$	Maximum/minimum recourse output of DGs
$\bar{q}_i^G, \underline{q}_i^G$	Maximum/minimum reactive power output of DGs
\bar{p}^l	Maximum power flow limit
\bar{v}, \underline{v}	Maximum/minimum limit of $\hat{v}_{i,t}$
r_{ij}, x_{ij}	Resistance and reactance of line ij
RU_i, RD_i	Ramp up/down limits of DGs
$p^{PV, cap}$	Installed capacity of solar panels
Δt	Duration of one time step

Continuous Variables

$e_{i,t}$	Stored energy of ESS
$\hat{p}_{i,t}^G$	Active power output of DGs after recourse action
p_i^S, \hat{p}_i^S	First-/second-stage purchased power from main grid
$p_{ij,t}, \hat{p}_{ij,t}$	First-/second-stage active power flow
$q_{ij,t}, \hat{q}_{ij,t}$	First-/second-stage reactive power flow
$v_{i,t}, \hat{v}_{i,t}$	First-/second-stage squared bus voltage magnitude
$\delta_{i,t}$	Power factor angle

probability distribution functions (PDFs) and scenario analysis into a multi-time scale SP framework. By managing slow- and fast-response resources to coordinate active and reactive power, the economic performance and the secure operation of the system are ensured under RESs uncertainty. SP features the need for PDFs or uncertain scenarios, making its performance closely tied to the estimation of distributional information or the number of scenarios, which can be well guaranteed with an incredibly large amount of historical data.

Robust optimization (RO) uses the boundary information of uncertainties, making it a practical choice for the optimal operation of DNs [6]. However, solutions derived from RO tend to be conservative because they focus solely on the worst-case scenario within a predefined uncertainty set, ignoring any additional information on the uncertainties. To mitigate the conservatism, some researchers have employed partial information to refine the uncertainty set. Ref. [7] searches for the extreme scenarios from the historical data of RESs to redesign the uncertainty sets, yielding a less conservative yet robust solution for determining the tap ratio of transformers and the capacity of switching capacitors. Ref. [8] reconstructs the uncertainty set of wind power based on the historical data of forecasted errors, aiming to reduce conservatism in economic dispatch. Beyond RO, distributionally robust optimization (DRO) incorporates available probability information on uncertainties to minimize operational costs over the worst-case distribution within the ambiguity sets constructed from historical data. Ref. [9] introduces a moment-based DRO to characterize RESs' uncertainty for real-time power dispatch in DNs. Ref. [10] constructs an ambiguity set of wind power using Wasserstein-based DRO, followed by a chance constraint-based DRO model for reactive power dispatch. Ref. [11] introduces the imprecise Dirichlet model (IDM) to construct the ambiguity set of wind power distribution, with an expected risk indicator to identify the worst-case distribution, aiming to balance operational costs and the risk costs in real-time energy dispatch. Ref. [12] proposes a comprehensive norm-based ambiguity set to represent renewable generation in integrated energy system scheduling, achieving better operational economy, lower renewable power curtailment rates, and higher computational efficiency compared to other methods. Collectively, the above studies, including SP, RO, and DRO, share a predefined paradigm: a probability distribution in SP, an uncertainty set in RO, or an ambiguity set in DRO is specified in advance, and the corresponding objective is then optimized under these uncertainty modeling approaches.

It is well-established that increasing the uncertainty set (ambiguity set) enhances robustness at the cost of reduced economic performance.

Several critical challenges thus arise in the optimization process, such as the size of the uncertainty set, the amount of operational cost budget allocated, as well as the trade-off between robustness and economic performance. Many of the aforementioned approaches fail to simultaneously address these challenges effectively. Information gap decision theory (IGDT) provides a robust decision-making framework for handling severe uncertainties without requiring any information or a predefined uncertainty set [13]. Risk-averse-based IGDT seeks to optimize the uncertainty set (representing risk-averse capability) within a preset financial budget, which is particularly useful when there is a clear target. For instance, given an operational budget, an IGDT-based energy management system is proposed for islanded microgrids in [14] to maximize the allowable range of load and RESs fluctuations while managing the system frequency excursion. Ref. [15] applies IGDT to identify robust operational scheme for resiliency-oriented DNs, addressing uncertainties in RESs, load demand, etc., while adhering to a certain budget. Ref. [16] incorporates the penetration level of RESs at different time periods into the uncertainty modeling framework, and proposes a weighted-IGDT to handle wind power uncertainty in microgrids. To account for the electricity price uncertainty and their correlations, the ellipsoid-bound IGDT has been utilized to model the electric vehicle aggregator scheduling in [17].

The envelope-bound uncertainty modeling dominates the current IGDT studies in network operation [14–16], as well as in network planning [18], electricity market offering strategy [19], etc., due to its simplicity and capability to handle forecasted types of uncertainty sources [13]. However, the uncertainty modelling of IGDT does not incorporate any information on uncertain parameters, and the real-valued symmetric uncertainty sets limit the approach to capture the asymmetric characteristics of uncertainties. The asymmetry of uncertainties primarily refers to scenarios that deviate significantly from the expected values yet have a higher probability of occurrence than closer scenarios, as commonly observed in the heavy tails of PDFs. For instance, scenarios with low RESs generation and high load demand, significantly deviating from expected values, are uncommon yet still possible and may be excluded from the uncertainty sets. These scenarios can have substantial impacts when they occur. However, many existing uncertainty modelling approaches fail to account for these significant scenarios probabilistically, or neglect to incorporate the desired operational budget when defining uncertainty sets. While SP considers the probability of scenarios, as previously discussed, its accuracy is highly dependent on the estimation of distributional information and the size of the dataset. A concise

Table 1
Comparison of uncertainty modeling approaches.

Method	Uncertainty Representation	Decision Perspective	Objective	Geometry
SP [3–5]	Probability/scenarios	Given uncertainty model	Min expected cost	Probabilistic region
RO [6–8]	Uncertainty set	Given uncertainty model	Min worst-case cost	Prescribed set
DRO [9–12]	Ambiguity set	Given uncertainty model	Min worst-case expectation	Distribution space
IGDT [14–19]	Uncertainty set	Given performance limit	Max uncertainty level	Symmetric interval expansion
CL-IGDT (this work)	Uncertainty set	Given performance limit	Max confidence level	Asymmetric interval expansion

comparison of these uncertainty modeling approaches is summarized in Table 1.

Given the advantages and disadvantages of the IGDT approach and its widespread utilization of envelope-bound uncertainty modeling, it is essential to seek potential enhancements to improve its applicability in DN operations. Therefore, we propose a novel uncertainty modelling approach that utilizes confidence levels to capture the asymmetric characteristics of uncertain parameters and maximize risk-averse capability in a probabilistic manner within the IGDT framework, while adhering to a desired budget. The IDM is employed to construct the ambiguity sets, leveraging available data and mitigating the dependency on accurate probability distributions. Consequently, the confidence level-based IGDT (CL-IGDT) framework is developed for the two-stage robust optimal operation of DNs. An iterative method is proposed to solve the model and determine the upper and lower bounds of the objective function. Overall, the main contributions of this paper are as follows:

- A CL-IGDT framework is proposed for two-stage optimal operation of distribution networks, extending traditional uncertainty modelling of IGDT by replacing symmetric intervals with confidence level-driven intervals.
- The IDM is utilized to construct the ambiguity sets for RESs and load demands, and integrated into the proposed framework. This method reduces reliance on precise probability distributions and leverages the available data at hand.
- Building on the proposed framework, a two-stage DN optimal operation model is constructed using CL-IGDT. By capturing the asymmetric characteristics of renewable energy and load demand uncertainties, the proposed approach enhances risk-averse capability in a probabilistic manner, leading to improved robustness and economic performance in distribution network operation.
- To solve the integrated model, an iterative solution method is developed to determine the upper and lower bounds of the objective function with guaranteed accuracy.

The remainder of this paper is organized as follows: Section 2 presents the mathematical formulation of the two-stage optimal operation model for distribution networks. Section 3 introduces the proposed CL-IGDT framework and the construction of ambiguity sets using the IDM. Section 4 describes the solution methodology, and Section 5 provides case study results and discussions. Finally, Section 6 concludes the paper.

2. Mathematical formulation

The necessity for a two-stage operation model arises from the challenges of multi-time scale control in DNs [9]. Consequently, this section introduces an optimization model for DNs' optimal operation problem, consisting of an economic dispatch plan and a recourse control stage. In the first stage, the approximate active power plan is scheduled, aiming at economic dispatch under expected uncertain scenarios. In the second stage, after the actual realization of uncertain scenarios, both active and reactive power regulation are implemented to minimize resource costs and ensure system security. The active power plan serves as a baseline decision that guides second-stage adjustments, ensuring the implementations remain within an economically and operationally

feasible range. Two types of distributed energy resources (DERs) are considered dispatchable: distributed generators (DGs) and energy storage systems (ESSs). It is worth clarifying that the proposed model assumes a forward-looking abstraction, i.e., a centralized coordinator responsible for both network security and DERs scheduling. This coordinator could represent a future DSO with expanded mandates, an aggregator, or a microgrid operator operating under network constraints. The general uncertainty modeling methodology is applicable to any entity facing similar operational challenges.

2.1. First-stage economic dispatch plan

In the first stage, decisions regarding the active power of DGs ($p_{i,t}^G$) and the charging/discharging power of ESSs ($p_{i,t}^{B+}/p_{i,t}^{B-}$) are defined based on expected forecasts for renewable energy ($\bar{p}_{i,t}^{PV}$) and load demands ($\bar{p}_{i,t}^L$), to determine the approximate active power output plan.

2.1.1. Objective function

The operational costs (Λ_1) include the cost of active power output from the DGs and the cost of electricity purchased from the upstream main grid.

$$\min \Lambda_1 = \sum_{i \in \mathcal{T}} \left(\sum_{i \in \mathcal{G}} c_i p_{i,t}^G + d_i p_t^S \right) \quad (1)$$

2.1.2. DGs constraints

The DGs' constraints are composed of the active power output limits (2a), and the ramp up/down rate limits (2b)–(2c).

$$p_i^G \leq p_{i,t}^G \leq \bar{p}_i^G, \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \quad (2a)$$

$$p_{i,t}^G - p_{i,t-1}^G \leq RU_i, \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \quad (2b)$$

$$p_{i,t-1}^G - p_{i,t}^G \leq RD_i, \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \quad (2c)$$

2.1.3. ESSs constraints

The ESSs are considered to operate in power control mode, allowing for pre-scheduling of their active power output [20]. Their operation is governed by constraints on stored energy dynamics (3a), storage limits (3b), and charging/discharging limits (3c)–(3d).

$$e_{i,t} = e_{i,t-1} + \eta_B p_{i,t}^{B+} \Delta t - \frac{p_{i,t}^{B-}}{\eta_B} \Delta t, \forall i \in \mathcal{B}, \forall t \in \mathcal{T} \quad (3a)$$

$$e_i \leq e_{i,t} \leq \bar{e}_i, \forall i \in \mathcal{B}, \forall t \in \mathcal{T} \quad (3b)$$

$$0 \leq p_{i,t}^{B+} \leq z_{i,t} \bar{p}^{B+}, \forall i \in \mathcal{B}, \forall t \in \mathcal{T} \quad (3c)$$

$$0 \leq p_{i,t}^{B-} \leq (1 - z_{i,t}) \bar{p}^{B-}, \forall i \in \mathcal{B}, \forall t \in \mathcal{T} \quad (3d)$$

where $z_{i,t}$ is the binary variable ensuring ESSs operate in only one state at a time—charging, discharging, or stand-by.

2.1.4. Power flow constraints

Constraints (4a) model the active power balance equations, and constraints (4b) consider the net active loads ($p_{i,t}$).

$$\sum_{hi \in \mathcal{L}} p_{hi,t} - \sum_{ij \in \mathcal{L}} p_{ij,t} = p_{i,t}, \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \quad (4a)$$

$$p_{i,t} = p_{i,t}^G + p_{i,t}^{B+} - p_{i,t}^{B-} + \hat{p}_{i,t}^{PV} - \hat{p}_{i,t}^L, \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \quad (4b)$$

2.1.5. Network security constraints

The maximum power flow limits are enforced by:

$$-\bar{p}^l \leq p_{ij,t} \leq \bar{p}^l, \forall ij \in \mathcal{L}, \forall t \in \mathcal{T} \quad (5)$$

2.2. Second-stage recourse control stage

Decisions on the recourse actions involving the active and reactive power outputs of DGs ($\hat{p}_{i,t}^G/\hat{q}_{i,t}^G$) are defined based on the real realizations of renewable energy ($p_{i,t}^{PV}$) and load demands ($p_{i,t}^L$), to guarantee the minimal recourse control cost and the system security. The proposed model aims to enhance operational flexibility and coordination, ensuring a less-conservative strategy under actual operating conditions.

2.2.1. Objective function

The recourse control cost (Λ_2) is determined by the cost of recourse active power output of the DGs and the cost of additional electricity purchased from the main grid.

$$\min \Lambda_2 = \sum_{i \in \mathcal{T}} \left(\sum_{i \in \mathcal{G}} \hat{c}_i \hat{p}_{i,t}^G + \hat{d}_i (\hat{p}_i^S - p_i^S) \right) \quad (6)$$

2.2.2. DGs recourse constraints

The DGs' constraints are composed of the recourse capability limits (7a), the reactive power output limits (7b), the active power output after recourse control (7c) as well as the ramp up/down rate limits of the active output after recourse control (7d)–(7e).

$$\hat{p}_{i,t}^G \leq \hat{p}_{i,t}^G \leq \bar{p}_{i,t}^G, \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \quad (7a)$$

$$\hat{q}_{i,t}^G \leq \hat{q}_{i,t}^G \leq \bar{q}_{i,t}^G, \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \quad (7b)$$

$$\hat{p}_{i,t}^G = p_{i,t}^G + \hat{p}_{i,t}^G, \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \quad (7c)$$

$$\hat{p}_{i,t}^G - \hat{p}_{i,t-1}^G \leq RU_i, \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \quad (7d)$$

$$\hat{p}_{i,t-1}^G - \hat{p}_{i,t}^G \leq RD_i, \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \quad (7e)$$

2.2.3. Power flow and voltage constraints

Constraints (8a)–(8b) model the active/reactive power balance equations. Constraints (8c)–(8e) consider the net active/reactive loads ($\hat{p}_{i,t}/\hat{q}_{i,t}$), and (8f) is the voltage magnitude drop in the lines.

$$\sum_{hi \in \mathcal{L}} \hat{p}_{hi,t} - \sum_{ij \in \mathcal{L}} \hat{p}_{ij,t} = \hat{p}_{i,t}, \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \quad (8a)$$

$$\sum_{hi \in \mathcal{L}} \hat{q}_{hi,t} - \sum_{ij \in \mathcal{L}} \hat{q}_{ij,t} = \hat{q}_{i,t}, \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \quad (8b)$$

$$\hat{p}_{i,t} = \hat{p}_{i,t}^G + p_{i,t}^{B+} - p_{i,t}^{B-} + p_{i,t}^{PV} - p_{i,t}^L, \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \quad (8c)$$

$$\hat{q}_{i,t} = \hat{q}_{i,t}^G - q_{i,t}^L, \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \quad (8d)$$

$$q_{i,t}^L = p_{i,t}^L \tan \delta_{i,t}, \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \quad (8e)$$

$$\hat{v}_{i,t} - \hat{v}_{j,t} = (r_{ij} \hat{p}_{ij,t} + x_{ij} \hat{q}_{ij,t}) / (1 - \phi_{ij}), \forall ij \in \mathcal{L}, \forall t \in \mathcal{T} \quad (8f)$$

The Lossy DistFlow formulation, with the help of parameter ϕ_{ij} [21], is employed to approximate the loss terms in voltage magnitude drop calculations within the traditional DistFlow model, as shown in (8f). This approach linearizes the voltage calculations while preserving a degree of accuracy.

2.2.4. Network security constraints

The maximum power flow limits and voltage magnitude limits are enforced by constraints (9a)–(9b).

$$-\bar{p}^l \leq \hat{p}_{ij,t} \leq \bar{p}^l, \forall ij \in \mathcal{L}, \forall t \in \mathcal{T} \quad (9a)$$

$$\underline{v} \leq \hat{v}_{i,t} \leq \bar{v}, \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \quad (9b)$$

Given the uncertainties of RESs and load demands, integrating the economic dispatch plan with a recourse control stage enhances operational flexibility. Compared to a single-stage approach, this model allows for dynamic adjustments by making recourse decisions in response to actual variations in uncertainties, and co-optimizes active and reactive power dispatch across different timescales, thereby minimizing the need for conservative strategies and ensuring system security against uncertain scenarios [9].

3. Confidence level-based information gap decision formulation

3.1. Deterministic model

A compact matrix-vector form of the optimization model presented in Section 2 is provided below for subsequent reformulation purposes.

$$\min_{\mathbf{x}, \mathbf{y}} \Lambda(\mathbf{x}, \mathbf{y}, \tilde{\xi}) = \mathbf{c}^T \mathbf{x} + \mathbf{d}^T \mathbf{y} \quad (10a)$$

$$\text{s.t. } \mathbf{Ax} \leq \mathbf{b} \quad (10b)$$

$$\mathbf{Ex} \leq \mathbf{F}\tilde{\xi} \quad (10c)$$

$$\mathbf{Gx} + \mathbf{Hy} \leq \mathbf{0} \quad (10d)$$

$$\mathbf{Ix} + \mathbf{Jy} \leq \mathbf{L}\tilde{\xi} \quad (10e)$$

$$\mathbf{My} \leq \mathbf{g} \quad (10f)$$

In the above deterministic formulation, the total operational cost is indicated by $\Lambda(\mathbf{x}, \mathbf{y}, \tilde{\xi})$, where \mathbf{x} refers to the first-stage decision variable including $[p_{i,t}^G, p_{i,t}^{B+}, p_{i,t}^{B-}, z_{i,t}]$, \mathbf{y} is the second-stage decision variable including $[\hat{p}_{i,t}^G, \hat{q}_{i,t}^G]$, and $\tilde{\xi} = \bar{\mathbf{p}}^L - \bar{\mathbf{p}}^{PV}$ denotes the uncertain vector related to load demands and renewable energy, which adopt their forecasted values in the deterministic model. The variable \mathbf{x} is constrained by (10b), corresponding to constraints (2)–(3), (4a), and (5). Constraint (10c) correlates \mathbf{x} with the forecasted value of uncertain vector $\tilde{\xi}$, corresponding to constraint (4b). Constraint (10d) builds the relationship between \mathbf{x} and \mathbf{y} , corresponding to constraint (7c). Constraint (10e) correlates \mathbf{x} and \mathbf{y} with the forecasted value of uncertain vector $\tilde{\xi}$, corresponding to constraints (8c)–(8e). Finally, the remaining constraints concerning \mathbf{y} are represented by (10f).

3.2. Information-gap decision model

DNs face significant uncertainties associated with RESs and load demands, necessitating a framework for uncertainty management. IGDT offers a non-probabilistic, non-fuzzy alternative that effectively handles uncertainties without the need to predefine the uncertainty extent [13]. The commonly used envelope-bound modelling by IGDT is shown below:

$$\Psi(\tilde{\xi}, \delta) = \{ \xi \mid |(\xi - \tilde{\xi})| / \tilde{\xi} \leq \delta, 0 \leq \delta \leq 1 \} \quad (11)$$

where the variable δ represents the uncertainty extent.

The uncertainty modelling $\Psi(\tilde{\xi}, \delta)$ highlights the gap between the forecasted values and unknown information. For a given $\tilde{\xi}$, an increasing δ enhances the risk-averse capability, leading to a more costly solution, as it requires more resources to ensure performance and reliability under uncertainty. A major achievement of IGDT is its risk-averse capability, which refers to the maximum uncertainty extent obtained by driving the actual system performance towards the desired target. As Fig. 1 shows, IGDT maximizes the uncertainty extent (safe region) to enable robust operational strategies. This is particularly advantageous in the energy

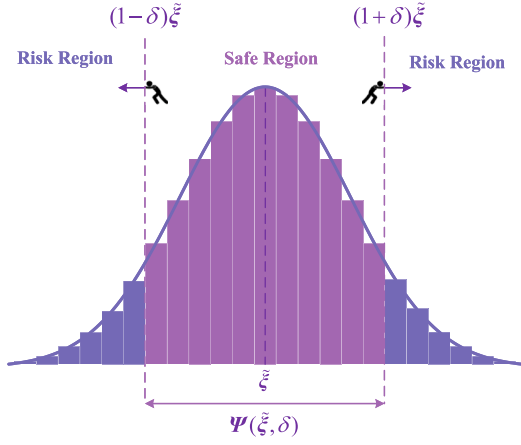


Fig. 1. The symmetric uncertainty set aligns perfectly with the shortest confidence interval of an uncertain parameter characterized by symmetric probability density function.

dispatch field when clear performance targets, such as the operational budget, are set by the DSO. Under the desired target, the risk-averse capability is maximized. Consequently, this paper applies IGDT to handle the uncertain parameters in (10), which can thus be reformulated as follows:

$$\max_{x,y} \delta \quad (12a)$$

$$\text{s.t. } \mathbf{Ax} \leq \mathbf{b} \quad (12b)$$

$$\mathbf{Ex} \leq \mathbf{F}\bar{\xi} \quad (12c)$$

$$\mathbf{Gx} + \mathbf{Hy}(\xi) \leq \mathbf{0}, \forall \xi \in \Psi(\bar{\xi}, \delta) \quad (12d)$$

$$\mathbf{Ix} + \mathbf{Jy}(\xi) \leq \mathbf{L}\xi, \forall \xi \in \Psi(\bar{\xi}, \delta) \quad (12e)$$

$$\mathbf{My}(\xi) \leq \mathbf{g}, \forall \xi \in \Psi(\bar{\xi}, \delta) \quad (12f)$$

$$\bar{\lambda} \leq (1 + \sigma)\Lambda_0 \quad (12g)$$

$$\bar{\lambda} = \mathbf{c}^T \mathbf{x} + \max_{\xi \in \Psi(\bar{\xi}, \delta)} \mathbf{d}^T \mathbf{y}(\xi) \quad (12h)$$

IGDT formulation in (12) is a two-level optimization model. At the upper level, the objective is to maximize the uncertainty extent while enforcing the constraints (12b)–(12g). The lower level is specified in (12g)–(12h), where σ represents the deviation factor that defines the acceptable degree of budget excess, and Λ_0 is the system performance under the expected scenario. Through this two-level framework, robust performance $\bar{\lambda}$ is deliberately driven towards the desired budget $(1 + \sigma)\Lambda_0$, ultimately enabling the optimal uncertainty extent.

The above uncertainty modelling requires little information on uncertainties and is well-suited for uncertain parameters exhibiting symmetric features. As shown in Fig. 1, such an uncertainty set can perfectly capture the scenarios most likely to occur. Uncertain parameters like renewable generation and load demands, however, typically do not follow a symmetric distribution [22]. As demonstrated in Fig. 2, this misalignment can result in issues such as some safe regions being overlooked where actual scenarios would have occurred with higher probability. As mentioned in Section 1, these scenarios can compromise DNs safe operation if they occur. Additionally, the safe regions can be unnecessarily large due to the inclusion of impossible scenarios, leading to conservatism.

3.3. Confidence level based information gap decision model

To mitigate the aforementioned drawbacks and better handle the asymmetric characteristics of uncertain parameters, we propose utilizing a confidence level-driven uncertainty set instead of the real-valued

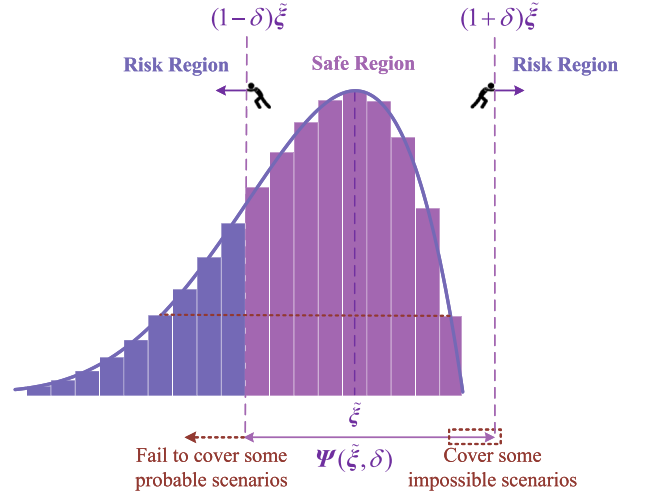


Fig. 2. For an uncertain parameter with asymmetric characteristics, the symmetric uncertainty set no longer matches the shortest confidence interval.

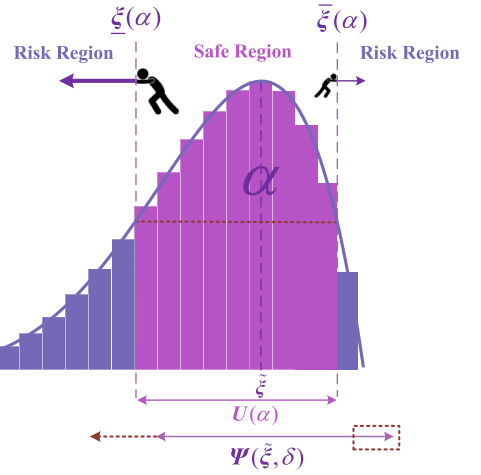


Fig. 3. The confidence level-driven uncertainty set captures the most probable uncertain scenarios, and thus matches the shortest confidence interval.

symmetric one. This approach transforms the uncertainty set into a sub-optimization problem:

$$U(\alpha) = \left\{ \xi \mid \xi \in [\underline{\xi}(\alpha), \bar{\xi}(\alpha)] = \arg \min_{\bar{\xi}, \underline{\xi}} \|\bar{\xi} - \underline{\xi}\| : \right. \\ \left. F_{\xi}(\bar{\xi}) - F_{\xi}(\underline{\xi}) \geq \alpha, 0 \leq \alpha \leq 1 \right\} \quad (13)$$

where α represents the confidence level of the uncertainty set; $\bar{\xi}(\alpha)$ and $\underline{\xi}(\alpha)$ are the upper and lower bounds of the set, respectively, functions of α ; $F_{\xi}(\cdot)$ denotes the cumulative distribution function (CDF) of ξ .

As illustrated in Fig. 3, the newly constructed uncertainty set captures the scenarios that are most probable, corresponding to the shortest confidence interval. This achieves the highest robustness with the least conservatism and mitigates the misalignment issue. Accordingly, the objective of IGDT shifts to maximizing α , allowing the uncertainty set to cover as many probable scenarios as possible within the desired budget. The improved IGDT formulation can be expressed as:

$$\max_{x,y} \alpha \quad (14a)$$

$$\text{s.t. } \mathbf{Ax} \leq \mathbf{b} \quad (14b)$$

$$E\mathbf{x} \leq F\bar{\xi} \quad (14c)$$

$$G\mathbf{x} + H\mathbf{y}(\xi) \leq \mathbf{0}, \forall \xi \in U(\alpha) \quad (14d)$$

$$I\mathbf{x} + J\mathbf{y}(\xi) \leq L\xi, \forall \xi \in U(\alpha) \quad (14e)$$

$$M\mathbf{y}(\xi) \leq \mathbf{g}, \forall \xi \in U(\alpha) \quad (14f)$$

$$\bar{\Lambda} \leq (1 + \sigma)\Lambda_0 \quad (14g)$$

$$\begin{aligned} \bar{\Lambda} &= \max_{\xi \in U(\alpha)} \Lambda(\mathbf{x}, \mathbf{y}(\xi), \xi) \\ &= \mathbf{c}^T \mathbf{x} + \max_{\xi \in U(\alpha)} \mathbf{d}^T \mathbf{y}(\xi) \end{aligned} \quad (14h)$$

By analysing the formulation in Section 2, it is evident that the largest operational cost occurs when the electricity supply is at its peak, i.e., when renewable generation and load demands reach their bounds $\xi^*(\alpha) = \bar{p}^L(\alpha) - p^{PV}(\alpha)$ within the uncertainty sets. Consequently, the two-level optimization model can be simplified into a single level one if the relatively complete assumption holds. Additionally, in standard IGDT, a single variable δ becomes less practical when multiple types of uncertain parameters are involved. In contrast, the proposed model represents multiple uncertainties through their respective probability distributions. Although the sets are mapped from the same α , each is linked to its own distribution, thereby achieving parameter decoupling compared to IGDT.

3.4. Construction of the ambiguity set

Introducing a confidence level α in (13) and (14) necessitates the CDFs, which may inherit the limitations associated with SP method. To address this challenge, we draw inspiration from DRO and propose constructing ambiguity sets of CDFs. By fully utilizing available historical data and exploiting the implicit information it contains, the worst-case distribution can serve as a surrogate for the precise distribution.

Consider a dataset \mathcal{D} of a random variable ξ with K different values $\xi_k, k = 1, \dots, K$, arranged in an increasing order. The CDF of ξ is defined as $F_\xi(x) = \Pr(\xi \leq x)$, so the cumulative probability of $\{\xi = \xi_k\}$ is $\theta_k = \Pr(\xi \leq \xi_k)$. According to the Law of Large Numbers, θ_k can be estimated as n_k^*/n as the total sample count $n \rightarrow \infty$, where n_k^* is the cumulative count corresponding to $\xi \leq \xi_k$. However, the limited size of \mathcal{D} introduces inaccuracy into the probability estimation. To quantify the imprecision in θ_k , the interval-valued probability is applied based on imprecise probability theory [23], which can be estimated using credible index corresponding to IDM [11]:

$$\begin{cases} \underline{\theta}_k = 0, \bar{\theta}_k = W^{-1}\left(\frac{1+\gamma}{2}\right), & n_k^* = 0 \\ \underline{\theta}_k = V^{-1}\left(\frac{1-\gamma}{2}\right), \bar{\theta}_k = W^{-1}\left(\frac{1+\gamma}{2}\right), & 0 \leq n_k^* \leq n \\ \underline{\theta}_k = V^{-1}\left(\frac{1-\gamma}{2}\right), \bar{\theta}_k = 1, & n_k^* = n \end{cases} \quad (15)$$

where γ is the confidence index to estimate the probability interval, and s is the equivalent sample size. $V(\cdot)$ and $W(\cdot)$ are the CDFs of Beta distribution, i.e., $\text{Beta}(n_k^*, s + n - n_k^*)$ and $\text{Beta}(s + n_k^*, n - n_k^*)$, respectively.

Due to the limited sample points in \mathcal{D} , a stair-step interpolation method is applied to obtain the probability intervals at non-sample points, which can be expressed as:

$$\begin{cases} \underline{\Pr}(\xi \leq \xi_\kappa) = \underline{\theta}_\kappa, \kappa \in [k, k+1) \\ \bar{\Pr}(\xi \leq \xi_\kappa) = \bar{\theta}_{k+1}, \kappa \in (k, k+1] \end{cases} \quad (16)$$

where κ indicates the non-sample points between two sample points k and $k+1$.

Finally, the ambiguity set is constructed by connecting the upper and lower bounds of these probability intervals, which can be expressed as:

$$\mathcal{A} = \{F_\xi \mid F_\xi(x) \in [\underline{\Pr}(\xi \leq x), \bar{\Pr}(\xi \leq x)]\} \quad (17)$$

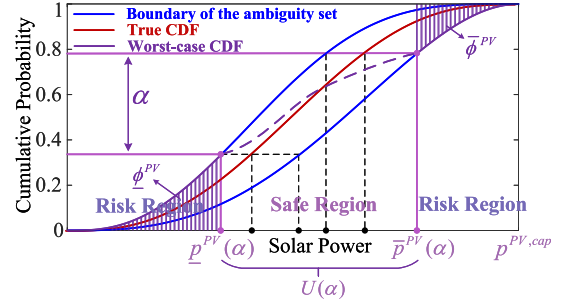


Fig. 4. Ambiguity set of solar power distribution. Given α , the optimal uncertainty set over the worst-case distribution (illustrated with purple dots) is the widest compared to those derived from alternative distributions (represented by black dots). This suggests that alternative distributions, which result in narrower uncertainty sets and thus better economic performance, are not the worst-case distributions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.5. Selection of the worst-case distribution

After counting the cumulative occurrences of different values of solar power p^{PV} , for example, (15) and (16) can be used to calculate the interval-valued cumulative probability and then construct the ambiguity set $\mathcal{A}_{p^{PV}}$, as depicted in Fig. 4.

When actual solar power remains within the safe region, the total operational cost is constrained within the preset budget while ensuring network security. Conversely, disturbances that extend beyond this region into the risk areas may require additional actions to address unexpected conditions, or even potentially compromise the network security.

It is assumed that the upside risk is caused by excessive solar power, whereas the downside risk is due to insufficient solar power. The risks can be quantified based on the concept of expected risk described in [11].

The downside risk (ϕ^{PV}) is quantified as follows:

$$\begin{aligned} \underline{\phi}^{PV} &= \int_0^{\underline{p}^{PV}} (\underline{p}^{PV} - x) f_{p^{PV}}(x) dx \\ &= \int_0^{\underline{p}^{PV}} F_{p^{PV}}(x) dx \end{aligned} \quad (18)$$

Considering the ambiguity in CDFs, the maximal downside risk is:

$$\max_{F_{p^{PV}} \in \mathcal{A}_{p^{PV}}} \underline{\phi}^{PV} = \max_{F_{p^{PV}} \in \mathcal{A}_{p^{PV}}} \int_0^{\underline{p}^{PV}} F_{p^{PV}}(x) dx \quad (19)$$

As illustrated in Fig. 4, the maximal downside risk equals to the shaded area between $F_{p^{PV}}$ and the interval $[0, \underline{p}^{PV}]$. Obviously, $\bar{F}_{p^{PV}}$ leads to the highest downside risk, thus representing the worst-case distribution when the actual solar power exceeds its left boundary of the uncertainty set. Similarly, $\underline{F}_{p^{PV}}$ represents the worst-case distribution when the real solar power exceeds the right boundary. The worst-case distribution of solar power $F_{p^{PV}}^*(\cdot)$ can be expressed as:

$$F_{p^{PV}}^*(x) = \begin{cases} \bar{F}_{p^{PV}}(x), & p^{PV} \leq \underline{p}^{PV} \\ \underline{F}_{p^{PV}}(x), & p^{PV} \geq \bar{p}^{PV} \end{cases} \quad (20)$$

With the construction of the ambiguity sets and the selection of the worst-case distribution pairs, the confidence level-driven uncertainty set in (13), which was originally based on precise CDFs, is now replaced by the one constructed using ambiguity sets, as (21) shows.

$$\begin{aligned} U(\alpha) &= \left\{ \xi \mid \xi \in [\underline{\xi}(\alpha), \bar{\xi}(\alpha)] = \arg \min_{\xi, \bar{\xi}} \|\xi - \bar{\xi}\|_1 : \right. \\ &\quad \left. \underline{F}_\xi(\bar{\xi}) - \bar{F}_\xi(\underline{\xi}) \geq \alpha, 0 \leq \alpha \leq 1 \mid \underline{F}_\xi, \bar{F}_\xi \in \mathcal{A} \right\} \end{aligned} \quad (21)$$

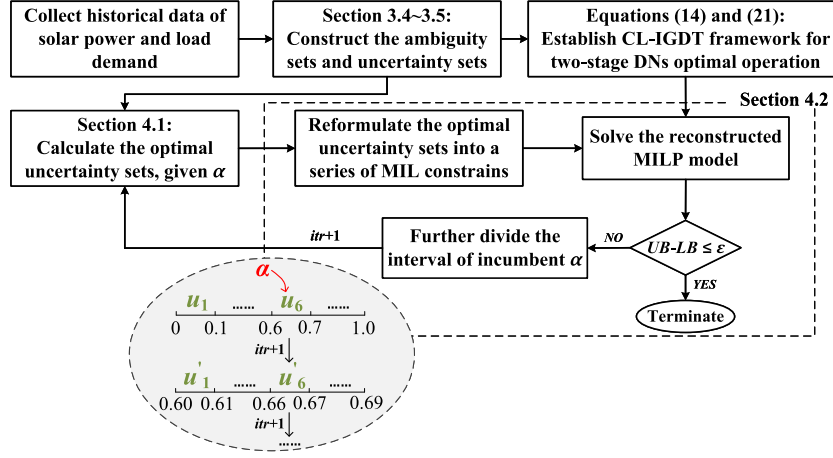


Fig. 5. Flowchart of the solution methodology.

4. Solution methodology

This section presents the methodology to solve the proposed optimization framework shown in (14) and (21). An interval search procedure is carried out as a preprocessing step to identify the shortest confidence interval, thereby enabling the offline determination of the optimal uncertainty set. The optimization formulation is thus transformed into a mixed integer linear programming (MILP) model. An iterative method is introduced to determine the upper and lower bounds of the optimal α until the required accuracy is achieved. Overall, the proposed methodology is outlined in Fig. 5.

4.1. Calculation of optimal uncertainty set

Explicitly expressing the uncertainty set in (21) is challenging, as the sub-optimization problem lacks closed-form distributions which are constructed by a series of discrete points, and thus deriving the optimality conditions is inapplicable. However, leveraging the structure of the proposed framework, an interval search approach is devised to calculate the uncertainty sets offline, which are then integrated into the original model. Again, we take solar power as an example.

Consider the ambiguity set of p_t^{PV} which consists of S ordered realizations. For each point $p_{t,s}^{PV}$ with $s = 1, \dots, S$, its corresponding lower and upper bounds of the CDFs, $\underline{F}_{p_t^{PV}}(\cdot)$ and $\overline{F}_{p_t^{PV}}(\cdot)$, are available.

Assume α is given, all candidate intervals $[p_t^{PV}, \overline{p}_t^{PV}]$ formed by the discrete points are examined by

$$\underline{F}_{p_t^{PV}}(\overline{p}_t^{PV}) - \overline{F}_{p_t^{PV}}(\underline{p}_t^{PV}) \geq \alpha \quad (22)$$

Among all feasible intervals, the one with the minimum length is selected. This procedure yields the shortest confidence interval of solar power in (21), which can be computed offline prior to solving the optimization problem. The resolution of the resulting uncertainty set is determined by the number of available sample points. As the shortest confidence intervals are constructed from discrete data points, the endpoints of these intervals can only take values from the available dataset.

4.2. Iteration solving procedure

The optimal uncertainty set depends on the optimal α , which remains unknown until the problem is solved. Fortunately, given that the range of the objective function is from 0 to 1, the structured nature of the optimization problem can be exploited to allow for an iterative method, where a larger objective value will lead to a higher worst-case cost. The solving procedure is shown in Fig. 5, and the details are as follows.

Step 1: The initial step is set as $itr = 1$. Assume that α takes values from ten equal intervals separated by discrete points, $0, 0.1, \dots, 1.0$. Binary variables $u_m, m = 1, \dots, 10$ are utilized to determine which interval α falls into, as shown below.

$$\begin{cases} \alpha = \sum_{m=1}^{10} \alpha_m \\ \underline{\alpha}_m^{(itr)} \cdot u_m \leq \alpha_m \leq \overline{\alpha}_m^{(itr)} \cdot u_m, \forall m \\ \sum_{m=1}^{10} u_m = 1 \end{cases} \quad (23)$$

where

$$\begin{aligned} \underline{\alpha}_m^{(itr)} &= 10^{-itr}(m-1) \\ \overline{\alpha}_m^{(itr)} &= 10^{-itr}m \end{aligned} \quad (24)$$

Step 2: Given the values of each known α_m , the corresponding optimal uncertainty set of solar power and load demands can be obtained offline by using the approach in Section 4.1, i.e., $U_t^{PV}(\alpha_m) = [p_t^{PV}(\alpha_m), \overline{p}_t^{PV}(\alpha_m)]$ and $U_t^L(\alpha_m) = [p_t^L(\alpha_m), \overline{p}_t^L(\alpha_m)], \forall t, \forall m = m \cup \{11\}$.

Step 3: The optimal uncertainty sets corresponding to different α_m are formulated as a series of constraints and then embedded into the original optimization problem. The big M technique is utilized to determine which set of constraints associated with the optimal α are active, as shown below.

$$\begin{cases} p_t^{PV}(\alpha_m) - M(1-u_m) \leq p_t^{PV} \\ \leq \overline{p}_t^{PV}(\alpha_m) + M(1-u_m), \\ p_t^L(\alpha_m) - M(1-u_m) \leq p_t^L \\ \leq \overline{p}_t^L(\alpha_m) + M(1-u_m) \end{cases} \quad (25)$$

Step 4: The original problem now consists of a series of mixed integer linear constraints, transforming the entire optimization problem into a MILP, which can be solved directly by commercial solvers. Since α assumes taking value from only ten intervals, the incumbent solution may lack precision. The boundaries of the interval into which the incumbent α falls $[\underline{\alpha}^{(itr*)}, \overline{\alpha}^{(itr*)}]$ serve as the lower bound LB and upper bound UB of the optimal α^* , respectively. To obtain more accurate result, we further divide the obtained interval into ten updated smaller intervals, and set the iteration itr to $itr + 1$.

$$\begin{cases} \alpha = \sum_{m=1}^{10} \alpha_m \\ (\underline{\alpha}^{(itr*)} + \underline{\alpha}_m^{(itr)}) \cdot u'_m \leq \alpha_m \leq (\overline{\alpha}^{(itr*)} + \overline{\alpha}_m^{(itr)}) \cdot u'_m, \forall m \\ \sum_{m=1}^{10} u'_m = 1 \end{cases} \quad (26)$$

Given ten new discrete values, the corresponding optimal uncertainty set can be calculated offline again. Embed them into the original optimization problem and resolve it, so the new information can be updated as $UB = \overline{\alpha}^{(itr*)}$, $LB = \underline{\alpha}^{(itr*)}$, and $itr = itr + 1$.

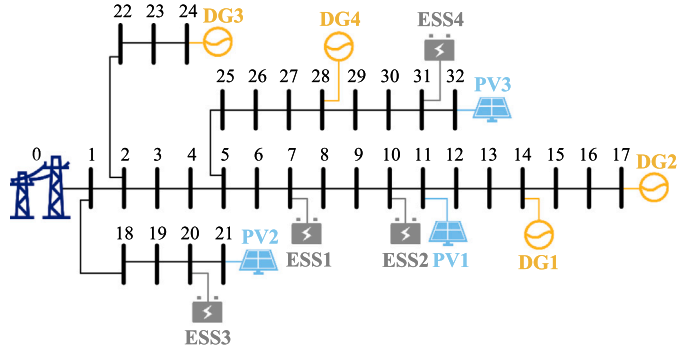


Fig. 6. Modified IEEE 33-bus system.

Table 2

DGs information.

Label	\bar{p}^G (kW)	\underline{p}^G (kW)	RU (kW)	RD (kW)	\bar{q}^G (kVar)	\underline{q}^G (kVar)	c (\$)
DG1	600	100	100	100	500	-400	4.0
DG2	400	100	90	90	300	-200	3.5
DG3	250	80	60	60	200	-150	2.5
DG4	50	10	15	15	40	-25	2.0

Table 3

ESSs information.

Label	\bar{p}^{B+} (kW)	\bar{p}^{B-} (kW)	\bar{e} (kW)	\underline{e} (kW)	η_B
ESS1	120	120	600	120	0.9
ESS2	40	40	200	40	0.9
ESS3	60	60	300	60	0.9
ESS4	80	80	400	80	0.9

Step 5: Further segmentation of α into smaller intervals is possible if $UB - LB \leq \epsilon$ does not satisfy, where ϵ indicates the specified accuracy. *Step 4* with smaller resolution will be repeated, or the optimal solution is assumed to be reached.

It is important to note that the achievable precision of α is inherently limited by the granularity of the available dataset, and its resolution depends on the sample size. For datasets with different sample sizes, users can adjust ϵ accordingly, i.e., a smaller ϵ may be feasible with richer data, while a larger ϵ may be more appropriate for coarser data.

5. Case study and discussion

The proposed framework was tested on the modified IEEE 33-bus system, as illustrated in Fig. 6. The system data are available in [24], and the data for DGs and ESSs are given in Tables 2 and 3, respectively. All simulations were conducted on a 64-bit PC with a 3.00 GHz CPU and 16 GB RAM in PyCharm. The simulation horizon was set to one day with a one-hour resolution. Solar power and load demand for each bus are considered uncertain sources. Approximately one year of historical data [25] was used to construct the ambiguity sets, and we set $\epsilon = 0.001$, which is supported by the granularity of our historical data. The confidence index γ and equivalent sample size s for IDM are set to 0.95 and 1, respectively [11].

5.1. Optimal operation results for distribution networks

To illustrate the basic results of the proposed method, the deterministic (DT) and IGDT-based approaches are chosen as benchmarks. The deviation factor σ is set as 30%, and all other parameters remain consistent. The optimal cost of DT is 216,773\$, which determines the desired budget for IGDT and CL-IGDT. The optimal uncertainty extent obtained by IGDT is 0.183. This result indicates that, by implementing the optimal operational scheme, whenever the actual solar power and load demands

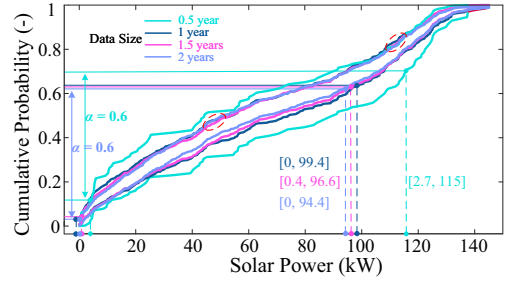


Fig. 7. Ambiguity sets of solar power distribution with different data sizes.

fall within the safe regions where deviations are within 18.3% of their predicted values, system safety is ensured and the total operational cost will remain within the preset budget; otherwise no guarantees can be made. For CL-IGDT, the optimal probability is 0.665. The same explanation applies, but the safe regions are defined by the shortest confidence intervals corresponding to this probability.

5.2. Ambiguity sets within proposed framework

According to the previous research in [11], as the sample size consistently increases, the ambiguity set shrinks and the worst-case distribution is expected to converge towards the true distribution. Consequently, the optimal uncertainty sets obtained should become progressively shorter, thereby reducing conservatism. Fig. 7 illustrates the ambiguity sets of solar power constructed using varying sizes of historical data and the optimal uncertainty set corresponding to each size, given $\alpha = 0.6$. Interestingly, as the data size increases, the ambiguity set enlarges slightly in some areas (e.g., those highlighted in ellipses), and the optimal uncertainty sets do not completely envelop the one obtained from a larger data size, which appears to contradict the theory to some extent. Upon further analysis, two reasons for this contradiction emerge.

Firstly, the distribution's shape can initially be rough when the data size is small. This explains the significant change in the shape of the ambiguity set initially, but eventually stabilizing and approaching its real shape as more data is incorporated. Secondly, the initial sample size may be insufficient, and thus the limited discrete points inadequately represent the confidence level of 0.6, causing a deviation in the uncertainty set. It is suggested that with a sufficiently large dataset, the actual distribution and the optimal uncertainty set will be accurately captured, thereby reducing conservatism and aligning with theoretical expectations.

5.3. Optimal uncertainty sets of different approaches

As shown in Fig. 8(a), the uncertainty set (safe region) obtained by CL-IGDT method is notably larger than that obtained by IGDT. Though larger uncertainty sets can lead to higher costs, it is probabilistically worthwhile to guarantee robustness at the expense of economic performance. This is because the safe region of the proposed method nearly (considering ambiguity) encompasses the scenarios with the highest probability of occurrence, whereas the uncertainty set obtained by IGDT is centered around its predictive value and includes some unnecessary scenarios (those highlighted in ellipse).

In contrast, Fig. 8(b) illustrates the shorter uncertainty set obtained by CL-IGDT than IGDT. In this case, the most probable scenarios are relatively centered around its predictive value. To probabilistically ensure economic performance, scenarios with low probabilities of occurrence are excluded from the safe region to shorten the uncertainty set. It is also worth noting that the differences between the uncertainty sets obtained by the two approaches are reduced compared to Fig. 8(a), as the asymmetry of the distribution is significantly mitigated. It is reasonable

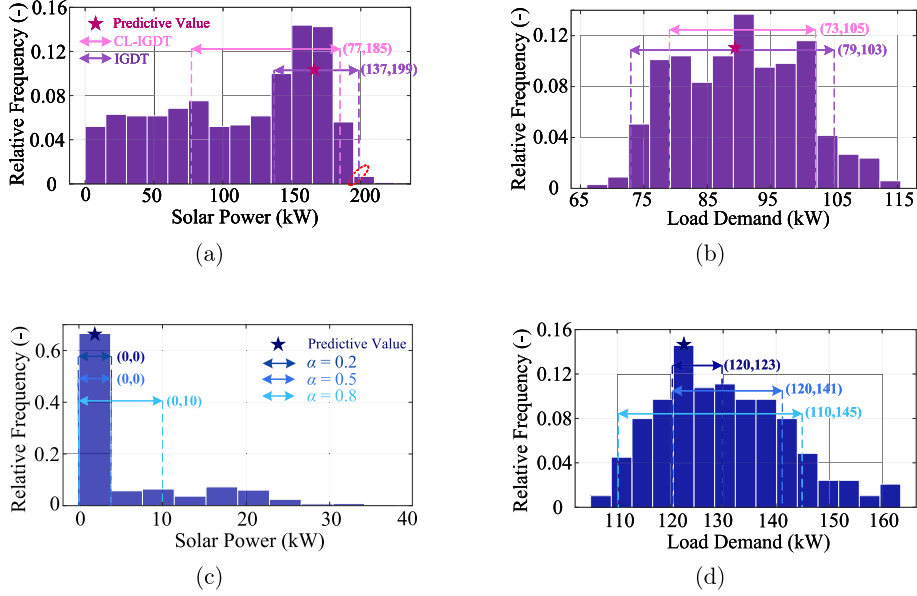


Fig. 8. Optimal uncertainty sets. (a) Solar power for time period 11:00–12:00; (b) load demand at node 20 for time period 1:00–2:00; (c) solar power for time period 8:00–9:00; (d) load demand at node 21 for time period 13:00–14:00.

to infer that if the distribution exhibits symmetric characteristics, such as a Gaussian distribution, the difference in uncertainty set will diminish further, potentially becoming nearly identical. Therefore, the trade-off between economic performance (smaller uncertainty set) and robustness (larger uncertainty set) can be guaranteed in a probabilistic manner, depending on the features of the data.

To further explore the superiority of the proposed uncertainty modelling, the optimal intervals under different α were examined. Fig. 8(c) shows that interval remains stable even though α increases from 0.2 to 0.5. While α reaches 0.8, this interval changes noticeably. The nonlinear expansion of the interval occurs since the solar power is predominantly centered around zero. For load demands without any distribution pattern as shown in Fig. 8(d), the interval also expands nonlinearly rather than linearly as in the IGDT method. This indicates that nonlinear robustness is considered in the proposed method.

5.4. Post-evaluation of different approaches

The ‘here-and-now’ decision x is robustly optimized and serves as the final scheme, whereas the ‘wait-and-see’ decisions y can dynamically adjust in response to the realization of the uncertain scenarios. To assess the risk-averse capability of the CL-IGDT solution, post-evaluation is conducted on x obtained by different approaches, using 50 randomly selected scenarios that were not included in the ambiguity sets. Additionally, to evaluate the impact of historical data size on performance, two cases with limited data—one month and six months—are considered, denoted as CL-IGDT (limited-1) and CL-IGDT (limited-2), respectively. The performance metric is the expected project budget (EPB) [26] defined in (27), and the reliability of the solutions, which refers to the percentage of feasible solutions.

$$EPB = c^T x + \frac{1}{NF} \sum_{n=1}^{NF} d^T y_n \quad (27)$$

where NF represents the number of common feasible solutions shared by different approaches.

5.4.1. Considering solar power and load demand uncertainties

As shown in Fig. 9(a), the DT method incurs the highest EPB and yields the lowest reliability. The optimal x^* contributes to the lowest

first-stage cost because it is optimized solely under expected scenarios. Without considering any potential uncertainty, the solution lacks risk-averse capability and requires substantial adjustments to compensate for the impact of the uncertain scenarios that were not taken into consideration during the decision process, thus resulting in an extremely high second-stage cost and the lowest reliability. In contrast, the robustness of the IGDT solution is achieved by investing 30% more budget. By considering possible uncertain scenarios, higher first-stage cost is incurred but a significant reduction in the second-stage cost and a substantial increase in the reliability of the solutions are expected.

The CL-IGDT approach further reduces EPB and enhances reliability by accounting for the asymmetric feature of uncertainties and leveraging implicit patterns within the historical data to optimize x . Compared to the IGDT, the proposed approach reduced the first-stage cost by 0.84% and the second-stage average cost by 6.7%. As discussed in Section 5.3, unnecessary robustness is avoided (which reduces the cost) while necessary robustness is maintained (which increases the cost). These two cases offset each other, so as to ensure the best trade-off between the risk-averse capability and economic performance (which can be demonstrated in Fig. 9(a) by the lower first-stage cost still contributing to lower second-stage cost and higher reliability). This explains why the improvement in economic performance, relative to the total operational costs, is relatively modest, while the reliability of the solutions is fairly enhanced by 8% compared to IGDT.

As data availability decreases, both economic performance and reliability of the proposed approach deteriorate. In particular, CL-IGDT (limited-1) incurs a \$382 higher EPB than IGDT but with the same reliability percentage due to the lack of sufficient data resulting in rough ambiguity sets as discussed in Section 5.2. This proves IGDT’s superiority in handling severe uncertainties without any information, while given adequate data, the proposed approach achieves a better solution.

To further validate the proposed method under symmetric uncertainty conditions, we conducted an additional simulation where the uncertainty distributions were assumed to be symmetric. Specifically, we fitted a Gaussian distribution to the historical data, and generated one year of scenarios from these fitted distributions as the input for CL-IGDT (to construct ambiguity sets) and for IGDT (using the expected values as forecasts). The performance of both methods was then evaluated on 50 randomly sampled scenarios. As shown in Fig. 9(b),

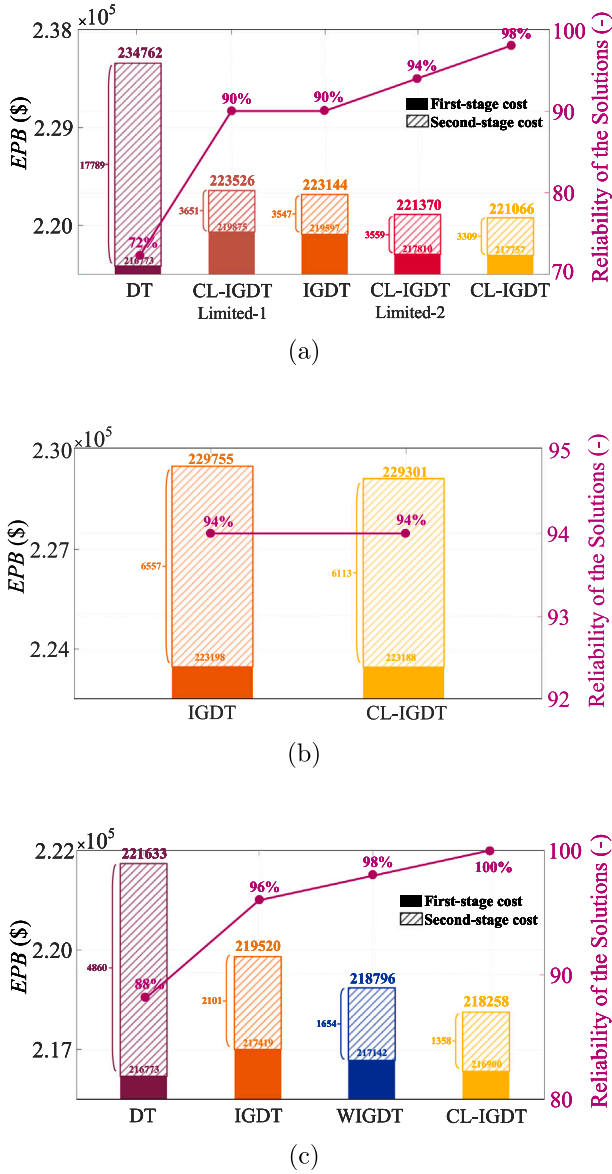


Fig. 9. Post-evaluation results by different approaches. (a) Considering solar power and load demand asymmetric uncertainties; (b) Considering solar power and load demand symmetric uncertainties; (c) Considering solar power asymmetric uncertainty only.

the EPB and the reliability of the solutions are nearly identical between CL-IGDT and standard IGDT. This is expected because when the underlying uncertainty is symmetric, the confidence-level-driven uncertainty set approaches the symmetric envelope-bound of IGDT, and the probabilistic risk-averse capability of CL-IGDT aligns with the robustness level achieved by IGDT. These results confirm that CL-IGDT generalizes the traditional IGDT framework. From Fig. 9(a) and (b), CL-IGDT maintains comparable performance in symmetric cases while offering superior risk-awareness under asymmetric uncertainties. The advantage of the proposed method is expected to become more pronounced as the asymmetric characteristics of the uncertainties become more severe.

5.4.2. Considering solar power uncertainty

To ensure a fair comparison with an improved IGDT approach—WIGDT [16], which fails to handle multiple uncertainties, the load demands are fixed at their expected values and treated as deterministic parameters. As shown in Fig. 9(c), the performance of WIGDT falls

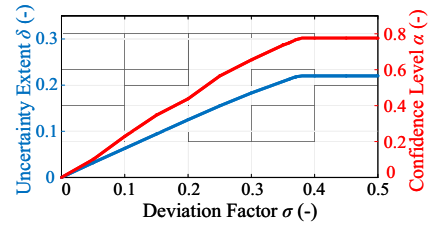


Fig. 10. Sensitivity analysis of deviation factor.

between IGDT and CL-IGDT. WIGDT accounts for the penetration levels of solar power by integrating them into the objective function as weights. Compared to IGDT, WIGDT offers greater flexibility, as it utilizes different variables to adjust uncertainty sets across different time periods. However, since WIGDT determines weights solely based on expected scenarios without leveraging historical solar power data, its performance remains inferior to the proposed method.

Additionally, a closer examination of Fig. 9(a) and (c) reveals that as multiple uncertainties are incorporated, the performance gap between IGDT and CL-IGDT widens, as reflected in the increase in economic performance difference from 0.57% to 0.93%, and in reliability difference from 4% to 8%. As discussed in Section 3.3, IGDT relies on a single variable δ to control uncertainty sets for different types of uncertain parameters, limiting its flexibility. The α in proposed method models the risk-averse capability in a probabilistic manner, ensuring that the uncertainty sets for different uncertain parameters consistently cover the most probable scenarios.

5.5. Sensitivity analysis

The impact of the deviation factor σ on the objective functions of IGDT and CL-IGDT methods is analyzed. Fig. 10 illustrates their corresponding values of δ and α , which indicate the risk-averse capability of the optimal solutions with regard to the varying σ . The risk-averse capability increases linearly as more budget is allocated, meaning that both approaches provide higher robustness at the expense of economical performance. However, as discussed in Section 5.3, the uncertainty sets of CL-IGDT increase nonlinearly even with a linear increase in α , unlike the linear response of IGDT method. Notably, when σ approaches around 0.4, the risk-averse capability plateaus for both approaches, signifying that the maximum level of robustness has been reached due to the limited availability of dispatchable resources within the DNs. This suggests that the risk-averse capability may further enhance if additional dispatchable resources are considered.

Furthermore, it is worth noting that δ and α determine the size of the uncertainty set in IGDT and CL-IGDT, respectively. A larger δ or α implies a more robust and costly decision-making. Within the general IGDT framework, the preset budget determines the maximum achievable robustness level. Through the sensitivity analysis, one can observe how the robustness level varies with the available budget, which distinguishes IGDT from traditional RO, where the uncertainty set size must be predetermined.

5.6. Accuracy of the optimal solutions

The iteration method in Section 4 is used to determine the upper and lower bounds of the objective function in (14a). To test the validity of the proposed method, different accuracy requirements were tested. As iterations progress, the incumbent α becomes more accurate, as Table 4 shows, thus leading to increasingly refined confidence level-driven uncertainty sets. Additionally, as the accuracy of α improves, the EPB decreases and the reliability of the solutions is enhanced. This indicates that the incumbent solution progressively moves towards the optimal α^* . Once the preset accuracy threshold is met, the incumbent solution is considered sufficiently optimal.

Table 4
Accuracy of the optimal solutions.

Iterations	Incumbent α	Range of α^*	EPB (\$)	Reliability	Time (s)
1	0.6	[0.60, 0.69]	223,453	94%	2.4
2	0.66	[0.660, 0.669]	223,434	98%	5.3
3	0.665	[0.6650, 0.6659]	223,421	98%	7.5
4	0.6651	[0.66510, 0.66519]	223,404	98%	10.1

5.7. Scalability test on the 69-bus system

To examine the scalability of the proposed framework and algorithm, a larger modified IEEE 69-bus system is implemented, as illustrated in Fig. 11. The network parameters can be referred to in [27], and the data for DGs and ESSs are provided in Tables 5 and 6, respectively.

To enable a clear comparison, information on the optimization models of the 33- and 69-bus systems is extracted from the Gurobi solver logs. Since each iteration of the proposed algorithm involves solving a MILP of identical size, Table 7 reports the model size in the first iteration. As shown, the numbers of constraints and continuous and binary variables for the 69-bus system are approximately twice those of the 33-bus system, while the number of nonzeros is nearly tripled. Both MILP models were solved by a branch-and-cut algorithm within Gurobi. For such algorithms, the increase in computational complexity is mainly associated with the number of binary variables. In our case, the number of binary variables used to control the operational modes of DGs and ESSs does not increase dramatically with the network scale and remains within a reasonable range, increasing by only 72 (from 130 to 202). The number of binary variables used to activate the uncertainty intervals remains the same for both networks. As a result, the computational time increases only slightly, from 2.4 s to 3.0 s, corresponding to an increase of approximately 25%.

The primary scalability bottleneck of the proposed framework lies in the number of binary variables rather than in the continuous variables or constraints, which grow approximately linearly with the network size. It is expected that, for even larger networks, the number of constraints and variables will continue to increase linearly. As long as the number of binary variables used to control DGs and ESSs remains within a manageable range, the computational time of the proposed algorithm is expected to remain acceptable, since the scalability of MILP is mainly constrained by the combinatorial effect of binary variables.

The iterative process for the 69-bus system is summarized in Table 8. Similar trends to those observed in the 33-bus system (Table 4) are

Table 5
DGs information for the IEEE 69-Bus System.

Label	\bar{p}^G (kW)	\underline{p}^G (kW)	RU (kW)	RD (kW)	\bar{q}^G (kVar)	\underline{q}^G (kVar)	c (\$)
DG1	600	100	100	100	500	-400	4.0
DG2	400	100	90	90	300	-200	3.5
DG3	250	80	60	60	200	-150	2.5
DG4	50	10	15	15	40	-25	2.0
DG5	50	10	15	15	40	-25	2.0
DG6	600	100	100	100	500	-400	4.0
DG7	1000	200	150	150	700	-700	5.0

Table 6
ESSs information for the IEEE 69-Bus System.

Label	\bar{p}^{B+} (kW)	\bar{p}^{B-} (kW)	\bar{e} (kW)	\underline{e} (kW)	η_B
ESS1	40	40	200	40	0.9
ESS2	80	80	400	80	0.9
ESS3	60	60	300	60	0.9
ESS4	120	120	600	120	0.9
ESS5	120	120	600	120	0.9
ESS6	180	180	1000	200	0.9
ESS7	60	60	300	60	0.9

Table 7
MILP model size for the 33- and 69-bus systems in the first iteration.

Network	Model	# Constraints	# Continuous	# Binary	# Nonzeros	Time(s)
33-bus system	MILP	8023	7811	130	74,355	2.4
69-bus system	MILP	16,579	16,235	202	269,235	3.0

Table 8
Accuracy of the optimal solutions on the 69-bus system.

Iterations	Incumbent α	Range of α^*	EPB (\$)	Reliability	Time (s)
1	0.6	[0.60, 0.69]	251,218	96%	3.0
2	0.61	[0.610, 0.619]	251,201	96%	6.2
3	0.611	[0.6110, 0.6119]	251,243	96%	9.2
4	0.6117	[0.61170, 0.61179]	251,140	96%	11.9

evident: the incumbent solution is progressively refined as the accuracy of α improves, leading to a more precise characterization of the uncertainty set. Meanwhile, the cumulative computational time over

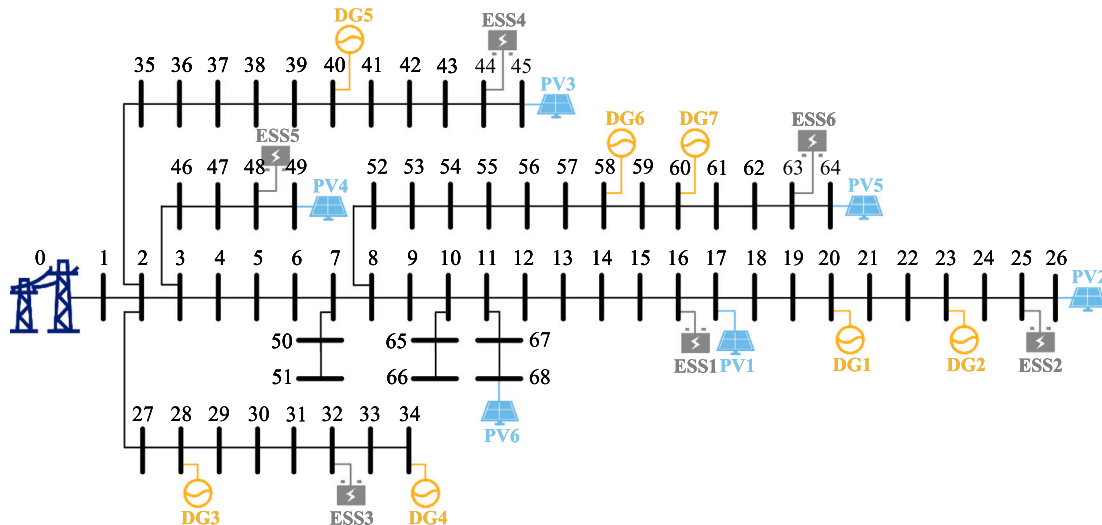


Fig. 11. Modified IEEE 69-bus system.

iterations increases slightly due to the modest growth in the number of binary variables. Overall, these results demonstrate the scalability and computational efficiency of the proposed framework and algorithm.

6. Conclusion

This paper proposes a CL-IGDT framework for the two-stage optimal operation of distribution networks, addressing renewable energy and demand asymmetric uncertainties. Results show that with sufficient historical data, the proposed method achieves the best trade-off between economic performance and robustness, reducing first-stage cost by 0.84%, second-stage average cost by 6.7%, and improving solution reliability by 8%. Otherwise, with insufficient data, confidence level-driven uncertainty sets become inaccurate, leading to performance deterioration. In such cases, IGDT is preferable for handling severe uncertainties without prior information. As additional budget is allocated, the risk-averse capability increases linearly (while uncertainty sets expand nonlinearly) but finally plateaus due to limited dispatchable resources. The proposed methodology is validated by its ability to achieve lower *EPB* and higher reliability over iterations.

The effectiveness and superiority of the proposed CL-IGDT framework rely on sufficient data, the presence of asymmetric uncertainties, and the unimodality of the underlying distributions. Unimodality guarantees the nestedness of the confidence-level-driven uncertainty sets and, consequently, the monotonicity of the recourse problem. While unimodality is a reasonable approximation for typical load and PV data in distribution networks, its validity may be challenged in some cases where distributions are multimodal. In addition, temporal and geographical correlations among uncertainties are not considered in the current framework. Incorporating such higher-fidelity modeling would enable more informed decision-making, albeit at the cost of more sophisticated reformulation and solution methodologies, an interesting direction for future research.

CRedit authorship contribution statement

Zhisheng Xiong: Writing – original draft, Software, Methodology, Investigation, Conceptualization. **Bo Zeng:** Methodology. **Peter Palensky:** Supervision, Funding acquisition. **Pedro P. Vergara:** Writing – review & editing, Validation, Supervision, Conceptualization.

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 work was supported by the [China Scholarship Council \(CSC\)](#) (Grant No. [202208360046](#)).

Data availability

Data will be made available on request.

References

- [1] L. Lukač, T. Antić, T. Capuder, Changing the paradigm of distribution networks planning and operation: a systematic review of the distributed energy resources impact, *Sustain. Energy Grids Netw.* 45 (2026) 102070, <https://doi.org/10.1016/j.segan.2025.102070>
- [2] W. Sun, Y. Peng, H. Wang, Y. Qiao, Low-carbon scheduling of mobile energy storage in distribution networks based on an equivalent reconfiguration method, *Sustain. Energy Grids Netw.* 42 (2025) 101649, <https://doi.org/10.1016/j.segan.2025.101649>
- [3] Z. Wu, P. Zeng, X.-P. Zhang, Q. Zhou, A solution to the Chance-Constrained Two-Stage stochastic program for unit commitment with wind energy integration, *IEEE Trans. Power Syst.* 31 (6) (2016) 4185–4196, <https://doi.org/10.1109/TPWRS.2015.2513395>
- [4] Y. Liu, N.-K.C. Nair, A Two-Stage stochastic dynamic economic dispatch model considering wind uncertainty, *IEEE Trans. Sustain. Energy* 7 (2) (2016) 819–829, <https://doi.org/10.1109/TSTE.2015.2498614>
- [5] S. Chen, C. Wang, Z. Zhang, Multitime scale active and reactive power coordinated optimal dispatch in active distribution network considering multiple correlation of renewable energy sources, *IEEE Trans. Ind. Appl.* 57 (6) (2021) 5614–5625, <https://doi.org/10.1109/TIA.2021.3100468>
- [6] H. Kiani, K. Hesami, A. Azarhooshang, S. Pirouzi, S. Safaee, Adaptive robust operation of the active distribution network including renewable and flexible sources, *Sustain. Energy Grids Netw.* 26 (2021) 100476, <https://doi.org/10.1016/j.segan.2021.100476>
- [7] Y. Zhang, X. Ai, J. Wen, J. Fang, H. He, Data-Adaptive robust optimization method for the economic dispatch of active distribution networks, *IEEE Trans. Smart Grid* 10 (4) (2019) 3791–3800, <https://doi.org/10.1109/TSG.2018.2834952>
- [8] C. Lu, N. Gu, W. Jiang, C. Wu, Sample-Adaptive robust economic dispatch with statistically feasible guarantees, *IEEE Trans. Power Syst.* 39 (1) (2024) 779–793, <https://doi.org/10.1109/TPWRS.2023.3267097>
- [9] Y. Yang, W. Wu, A distributionally robust optimization model for Real-Time power dispatch in distribution networks, *IEEE Trans. Smart Grid* 10 (4) (2019) 3743–3752, <https://doi.org/10.1109/TSG.2018.2834564>
- [10] J. Liu, Y. Chen, C. Duan, J. Lin, J. Lyu, Distributionally robust optimal reactive power dispatch with wasserstein distance in active distribution network, *J. Mod. Power Syst. Clean Energy* 8 (3) (2020) 426–436, <https://doi.org/10.35833/MPCE.2019.000057>
- [11] P. Li, M. Yang, Q. Wu, Confidence interval based distributionally robust Real-Time economic dispatch approach considering wind power accommodation risk, *IEEE Trans. Sustain. Energy* 12 (1) (2021) 58–69, <https://doi.org/10.1109/TSTE.2020.2978634>
- [12] Y. Li, M. Han, M. Shahidepour, J. Li, C. Long, Data-driven distributionally robust scheduling of community integrated energy systems with uncertain renewable generations considering integrated demand response, *Appl. Energy* 335 (2023) 120749, <https://doi.org/10.1016/j.apenergy.2023.120749>
- [13] Y. Ben-Haim, *Info-Gap Decision Theory: Decisions Under Severe Uncertainty*, Elsevier, 2006.
- [14] N. Rezaei, A. Ahmadi, A.H. Khazali, J.M. Guerrero, Energy and frequency hierarchical management system using information gap decision theory for islanded microgrids, *IEEE Trans. Ind. Electron.* 65 (10) (2018) 7921–7932, <https://doi.org/10.1109/TIE.2018.2798616>
- [15] M. Izadi, S. Hossein Hosseinian, S. Dehghan, A. Fakharian, N. Amjadi, Resiliency-Oriented operation of distribution networks under unexpected wildfires using Multi-Horizon Information-Gap decision theory, *Appl. Energy* 334 (2023) 120536.
- [16] M.-A. Nasr, E. Nasr-Azadani, H. Nafisi, S.H. Hosseinian, P. Siano, Assessing the effectiveness of weighted information gap decision theory integrated with energy management systems for isolated microgrids, *IEEE Trans. Industr. Inform.* 16 (8) (2020) 5286–5299, <https://doi.org/10.1109/TII.2019.2954706>
- [17] J. Zhao, C. Wan, Z. Xu, J. Wang, Risk-Based Day-Ahead scheduling of electric vehicle aggregator using information gap decision theory, *IEEE Trans. Smart Grid* 8 (4) (2017) 1609–1618, <https://doi.org/10.1109/TSG.2015.2494371>
- [18] M. Ahmadijorji, N. Amjadi, S. Dehghan, A robust model for multiyear distribution network reinforcement planning based on Information-Gap decision theory, *IEEE Trans. Power Syst.* 33 (2) (2018) 1339–1351, <https://doi.org/10.1109/TPWRS.2017.2732447>
- [19] S. Mobasheri, M. Rashidinejad, A. Abdollahi, M. MollahassaniPour, S. Dorahaki, A game-theoretic model for flexibility-constrained renewable energy communities in local energy trading with smart distribution networks, *Sustain. Energy Grids Netw.* 45 (2026) 102105, <https://doi.org/10.1016/j.segan.2025.102105>
- [20] P.P. Vergara, J.C. López, M.J. Rider, L.C.P. da Silva, Optimal operation of unbalanced Three-Phase islanded Droop-Based microgrids, *IEEE Trans. Smart Grid* 10 (1) (2019) 928–940, <https://doi.org/10.1109/TSG.2017.2756021>
- [21] E. Schweitzer, S. Saha, A. Scaglione, N.G. Johnson, D. Arnold, Lossy DistFlow formulation for single and multiphase radial feeders, *IEEE Trans. Power Syst.* 35 (3) (2020) 1758–1768, <https://doi.org/10.1109/TPWRS.2019.2954453>
- [22] M. Anvari, E. Proedrou, B. Schäfer, C. Beck, H. Kantz, M. Timme, Data-driven load profiles and the dynamics of residential electricity consumption, *Nat. Commun.* 13 (1) (2022) 4593.
- [23] P. Walley, Inferences from multinomial data: learning about a bag of marbles, *J. R. Stat. Soc. B: Stat. Methodol.* 58 (1) (1996) 3–34.
- [24] M. Baran, F. Wu, Network reconfiguration in distribution systems for loss reduction and load balancing, *IEEE Trans. Power Deliv.* 4 (2) (1989) 1401–1407, <https://doi.org/10.1109/61.25627>
- [25] H. Shengren, P.P. Vergara, E.M. Salazar Duque, P. Palensky, Optimal energy system scheduling using a constraint-aware reinforcement learning algorithm, *Int. J. Electr. Power Energy Syst.* 152 (2023) 109230, <https://doi.org/10.1016/j.ijepes.2023.109230>
- [26] X. Cao, J. Wang, B. Zeng, A chance constrained Information-Gap decision model for Multi-Period microgrid planning, *IEEE Trans. Power Syst.* 33 (3) (2018) 2684–2695, <https://doi.org/10.1109/TPWRS.2017.2747625>
- [27] O.D. Montoya, C.A. Ramos-Paja, L.F. Grisales-Noreña, An efficient methodology for locating and sizing PV generators in radial distribution networks using a mixed-integer conic relaxation, *Mathematics* 10 (15) (2022) 2626.