ENGINEERING Information Technology & Electronic Engineering
https://www.academax.com/EITEE; https://www.engineering.org.cn
https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=11344065
E-mail: eng.itee@zju.edu.cn

Engineering Journals
Available on IEEE Xplore®

Research Article
https://doi.org/10.1631/ENG.ITEE.2026.0024

An approach to characterizing the power system security region by integrating distributionally robust optimization and Transformer-based deep learning

Yuekai CHEN, Zhejing BAO✉, Miao YU

College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China

Abstract: Renewable generation and load uncertainty pose significant challenges to power system security, necessitating efficient approaches to characterizing high-dimensional security regions. To overcome the curse of dimensionality, uncertainty neglect, and undue conservatism in existing methods, this paper proposes an approach integrating distributionally robust optimization (DRO) and deep learning for security region characterization. First, to properly account for uncertainty while avoiding excessive conservatism, a DRO-based active search strategy is developed to identify critical boundary points, where diffusion-generated renewable scenarios and load-deviation samples constructed around typical demand profiles are jointly used to build a robust probabilistic ambiguity set. Subsequently, a Transformer-based model learns from these boundary points to reconstruct the full high-dimensional security region. The model's self-attention mechanism captures the global nonlinear dependencies among dimensions, enabling a precise and efficient boundary fit. Simulations on IEEE test systems confirm that the approach accurately characterizes high-dimensional security regions at a low computational cost, yielding a security region with strong robustness to renewable-load uncertainty. This work offers a new paradigm for security assessment and decision support in power systems under high uncertainty.

Key words: Security region; Distributionally robust optimization; Deep learning; Transformer model; Data-driven

1 Introduction

The secure operation of modern power systems is increasingly challenged by the high penetration of variable and uncertain renewable energy sources (Zhang ZY et al., 2023, 2024; Aryani and Song, 2024). Traditional approaches enforce security conditions as constraints to locate a single optimal operating point, providing no global view of the full set of feasible operating points. As a result, operational margins remain opaque and the geometry of the security boundary is not characterized. Moreover, the coupling between individual operating points and the feasible space is not visually intuitive, making it difficult for operators to develop comprehensive strategies.

To address this need, the power system security region (SR) concept was introduced (Wu F and Kumagai, 1982). The SR defines the set of all operating points that satisfy the power flow equations and operational security constraints, offering system operators a global view of operational margins for decision-making (Teng et al., 2024). SRs are commonly classified into steady-state security regions (SSRs) (Wu F and Kumagai, 1982) which capture static characteristics and dynamic security regions (DSRs) (Wu FF et al., 1988) which account for dynamic processes. Recently, the widespread integration of renewables has profoundly altered grid operating characteristics and reshaped steady-state security boundaries (Jin et al., 2023; Lin et al., 2023; Xiao et al., 2024). Under high renewable uncertainty, the dominant operational risks manifest at the power balance and network flow level on minute-to-hour timescales, making the SSR the primary object for scheduling and flexibility assessment. Consequently, accurate characterization of the SSR remains a central challenge and is the focus of this study. In what follows, SR refers to SSR.

Approaches for the characterization of SR are broadly categorized as analytical and numerical (search-and-fit). Analytical methods derive explicit boundary expressions (Dai et al., 2019; Lin et al., 2021; Su et al., 2021; Tinoco et al., 2021), offering high accuracy. Su et al. (2021) proposed a complete characterization for the SR of electricity-gas integrated energy systems. Without any model simplification, the method reveals that the region possesses complex geometric properties. However, despite high accuracy, their

✉ Zhejing BAO, zjbao@zju.edu.cn
Yuekai CHEN, https://orcid.org/0009-0009-9946-9688
Zhejing BAO, https://orcid.org/0000-0002-8678-3805
Miao YU, https://orcid.org/0000-0002-7638-5264
CLC number: TM71; TP18

Received: Jan. 21, 2026; Revision accepted: Apr. 16, 2026;
Crosschecked: Apr. 23, 2026; Published online: May 18, 2026

mathematical derivation becomes intractable for modern, high-dimensional, and nonlinear power systems. In contrast, numerical approaches first sample discrete points via vertex search (Nguyen et al., 2019; Monteiro et al., 2020; Avila et al., 2021; Li X et al., 2021; Lin et al., 2022) or point-wise simulation (Liu L et al., 2020; Gao et al., 2023) and then employ techniques like hyperplane fitting (Jiang T et al., 2021; Sun and Yu, 2023; Zhang S et al., 2024) or convex hull (Chen et al., 2019) to construct the SR boundary. Sun and Yu (2023) approximated the SR boundary using the hyperplane method, which linearizes the complex boundary into a set of tractable linear inequalities. However, it tends to generate numerous non-binding or redundant hyperplanes. Chen et al. (2019) utilized the convex hull method to construct an SR, in which a set of optimized boundary points is identified and then enclosed to generate a more accurate, piecewise linear approximation of the boundary. While adaptable, the accuracy of these methods depends heavily on the number of sample points and the performance of the fitting algorithm.

In recent years, researchers have attempted to explore artificial intelligence-based methods to characterize the SR. Wu XW et al. (2023) employed a neural network to approximate the feasible region of a district heating system and reformulated it as mixed-integer linear constraints. Feng et al. (2024) proposed a deep learning framework with an attention-based convolutional neural network to characterize a high-dimensional carbon emission operation region. Although these data-driven approaches are effective, they share a common limitation: they operate in a deterministic setting and do not explicitly account for the stochastic uncertainty from renewable sources.

To address the uncertainty introduced by renewable energy and load, robust optimization (RO) has been integrated into SR construction. For instance, Chen et al. (2019) constructed a robust SR by identifying an operating region that remains secure across a given set of wind power scenarios. Yorino et al. (2018) employed the bi-level optimization to identify the SR boundary under the worst-case realization of uncertainty, which is modeled as the deterministic bounded interval.

Despite these advances, existing methods face significant limitations, primarily due to the curse of dimensionality, high computational costs, and a tendency either to ignore uncertainty or to impose undue conservatism. Specifically, numerical fitting-based methods struggle to provide a global understanding of the high-dimensional SR, as they often rely on low-dimensional projections. Achieving satisfactory accuracy under such settings requires sampling an enormous number of boundary points, leading to a prohibitive computational burden, especially under renewable energy uncertainty. Moreover, approaches like traditional RO ensure security at the cost of excessive conservatism: their worst-case formulations often overshrink the SR, thereby sacrificing operational flexibility and economic efficiency.

To address these challenges, this paper develops a hybrid approach that integrates distributionally robust optimization (DRO) and deep learning (DL) to characterize high-dimensional SR of power systems under uncertainty. The main contributions are summarized as follows:

1. A novel framework for constructing an SR of power systems is proposed, consisting of two stages, a DRO-based active boundary point search and a DL-based high-dimensional boundary reconstruction, which provides a data-driven solution for SR characterization under uncertainty.

2. A DRO-based search strategy is proposed to identify critical boundary points of the SR, where the ambiguity set is constructed from diffusion-generated renewable scenarios and load-deviation samples around typical demand curves. The strategy balances robustness and conservatism, thereby overcoming the limitations of traditional methods that either neglect uncertainty or are overly conservative.

3. A scalable Transformer-based method for high-dimensional SR characterization is presented. Its self-attention mechanism captures global nonlinear interdependencies across dimensions, enabling an accurate reconstruction of SR. The approach eliminates the dependency on specific observation directions or low-dimensional visualizations, thus offering a unified, general-purpose solution for SR characterization in diverse power systems.

2 Methodology

2.1 Security region definition

The SR is defined as the set of all feasible power injections. Any point within this set satisfies the power flow equations and a set of operational security constraints, including nodal voltage limits, branch flow limits, and generator output bounds. The definition of SR is formulated as

$$\Omega_{\mathrm{ss}}=\{\boldsymbol{y}:\exists \boldsymbol{x}\in R,\ f(\boldsymbol{x})=\boldsymbol{y}\}. \tag{1}$$

In Eq. (1), Ω_{ss} represents the SR; $\boldsymbol{y}$ denotes the power injection vector; $\boldsymbol{x}$ represents the system state variable, such as voltage magnitudes and phase angles at each node, as well as power flows through each branch, depending on the specific power flow model adopted; $f(\boldsymbol{x})=\boldsymbol{y}$ denotes the power flow equation, representing the power balance relations in the network; R is the set of operational security constraints.

2.2 Mathematical model

2.2.1 Power flow constraints

The standard alternating current power flow (ACPF) equations are

$$\begin{cases} P_k=\sum\limits_{i=1}^{N_{\mathrm{node}}} V_k V_i (G_{ki}\cos\theta_{ki}+B_{ki}\sin\theta_{ki}), \\ Q_k=\sum\limits_{i=1}^{N_{\mathrm{node}}} V_k V_i (G_{ki}\sin\theta_{ki}-B_{ki}\cos\theta_{ki}), \end{cases} \tag{2}$$

where P_k and Q_k are the active and reactive power injected at node k, respectively; V_k and V_i are the voltage magnitudes at nodes k and i, respectively; G_{ki} and B_{ki} are the real and imaginary parts of the element in the k^{th} row and i^{th} column of the nodal admittance matrix, respectively; θ_{ki} is the voltage angle difference between nodes k and i; N_{node} is the number of nodes in the system.

Because the boundary point acquisition with the standard ACPF model in Eq. (2) is computationally prohibitive, using the simplified power flow model can enhance practicality and efficiency. It

should be noted that our proposed SR characterization framework is broadly applicable to different network models of power systems. In this study, a distribution network model is considered without compromising generality. This generality is evidenced by the following simulations: although the simplified power flow model is used to generate SR boundary points, high fitting accuracy can still be achieved when the trained model is evaluated on an independent ACPF-based test set. The distribution networks, often featuring radial or weakly-meshed topologies and high R/X ratios (where R represents the resistance and X represents the reactance), are well-suited for linearized algorithms. Among these, the widely-adopted DistFlow model is adopted, by which Eq. (2) is approximated as

$$\begin{cases} P_{ij}=\sum\limits_{k:j\to k}P_{jk}+P_j+r_{ij}l_{ij}, \\ Q_{ij}=\sum\limits_{k:j\to k}Q_{jk}+Q_j+x_{ij}l_{ij}, \\ v_j=v_i-2(r_{ij}P_{ij}+x_{ij}Q_{ij})+(r_{ij}^2+x_{ij}^2)l_{ij}, \\ v_i l_{ij}=P_{ij}^2+Q_{ij}^2, \end{cases} \tag{3}$$

where P_{ij} and Q_{ij} are the active and reactive power flowing from the parent node i to child node j at the ends of the branch, respectively; P_j and Q_j denote the net active and reactive power demands at node j, respectively; $\Sigma_{k:j\to k}$ denotes the sum over all the child nodes k, for which node j is the parent node; r_{ij} and x_{ij} are the resistance and reactance of the branch connecting nodes i and j, respectively; v_i denotes the squared voltage magnitude at bus i;l_{ij} denotes the squared current magnitude of branch (i, j).

2.2.2 Operational security constraints

The constraints of nodal voltage magnitude and branch current magnitude are

$$(V_i^{\min})^2\le v_i\le(V_i^{\max})^2, \tag{4}$$

$$0\le l_{ij}\le\bar{l}_{ij}, \tag{5}$$

where $V_i^{\min}$ and $V_i^{\max}$ are the lower and upper voltage magnitude limits at bus i. $\bar{l}_{ij}$ is the upper limit of the squared current magnitude of branch (i, j).

Also, the generator outputs must satisfy their capacity constraints:

$$P_g^{\min}\le P_g\le P_g^{\max}, \tag{6}$$

where P_g is the active power output of generator g, with $P_g^{\min}$ and $P_g^{\max}$ denoting its lower and upper output limits, respectively.

2.3 Framework of the SR characterization

The proposed data-driven framework for characterizing the power system SR under renewable and load uncertainty consists of four stages, as illustrated in Fig. 1.

Stage 1: DRO ambiguity set construction. A diffusion model generates high-fidelity, renewable generation scenarios from historical data, while load uncertainty is represented by deviation samples constructed around typical demand curves. These samples are combined to form an ambiguity set required for DRO.

Stage 2: boundary point search. A DRO model is formulated to identify the boundary points of the SR by solving the extreme value of a target variable, subject to power flow equations, operational security constraints, and the renewable-load uncertainty set.

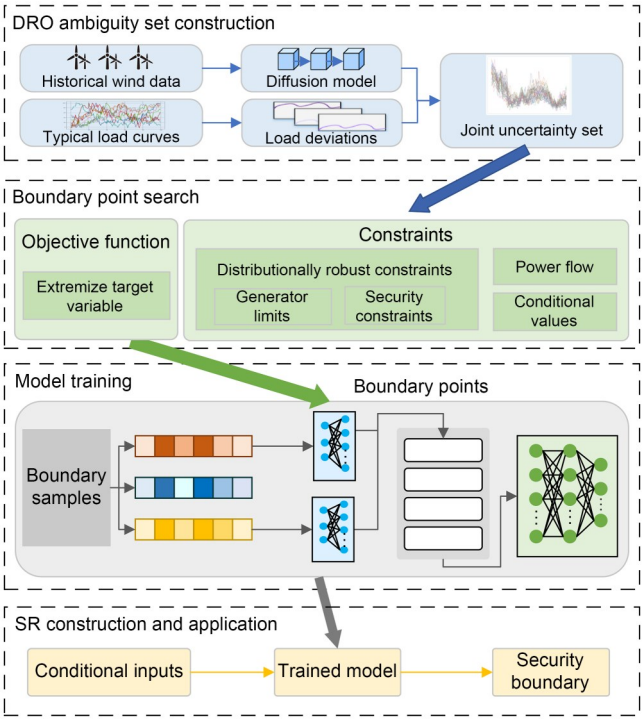

Fig. 1 Framework of the SR characterization by integrating DRO and DL

Stage 3: model training. A deep learning model is trained on the discrete boundary samples to learn the complex nonlinear mapping of the SR boundary, enabling the characterization of a continuous region.

Stage 4: SR construction and application. The trained model serves as an efficient computational tool that, for any given observation direction, can rapidly calculate the corresponding security boundary, thereby enabling online characterization of the entire SR.

3 Boundary point search strategy

A DRO-based search strategy is proposed to acquire a set of representative points that accurately characterize the SR boundary.

3.1 Construction of joint uncertainty samples

The DRO formulation requires representative samples to characterize source–load uncertainty. In this work, the uncertainty dataset is constructed from renewable generation scenarios and load-deviation samples (Rahimian and Mehrotra, 2019).

For renewable uncertainty, a diffusion model is employed to generate high-fidelity wind power scenarios from historical data. The core mechanism of the diffusion model consists of a forward diffusion process and a reverse denoising process. The forward diffusion process gradually adds Gaussian noise to the real historical wind power data until the signal transforms into pure noise. This mechanism enables the model to explicitly learn the complex, nonlinear probability distribution of wind power, thereby generating diverse and high-fidelity scenarios that preserve the statistical characteristics of the original data (Li S et al., 2024). The workflow of the diffusion-based scenario generation is illustrated in Fig. 2. Specifically, the model

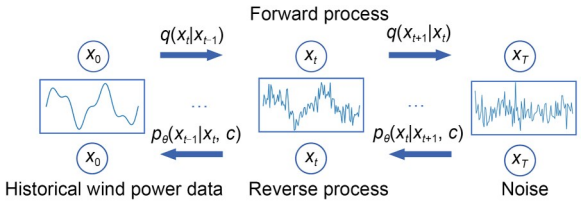

Fig. 2 Framework of the diffusion model for the wind power scenario generation (x_t, t=0, 1, ⋯, T, denotes the sample at diffusion step t, $q(\cdot)$ denotes the forward process, $p_\theta(\cdot)$ denotes the reverse process, c denotes the condition, and T denotes the total number of diffusion steps)

architecture proposed by Liu W et al. (2025) is adopted to generate a comprehensive set of wind power scenarios from historical data.

For load uncertainty, the typical load curve is taken as the nominal profile, and deviation fluctuations are superimposed to generate load uncertainty samples. These samples are combined with wind power scenarios to establish the source–load uncertainty dataset, which is used to construct the ambiguity set for the subsequent DRO-based boundary search.

3.2 Distributionally robust constraint formulation

To formulate the distributionally robust constraints under source–load uncertainty, let $\hat{\boldsymbol{\xi}}^n$ denote the n^{th} realization of the K-dimensional uncertainty vector, defined as

$$\hat{\boldsymbol{\xi}}^n=[\hat{\xi}_1^n, \hat{\xi}_2^n, \cdots, \hat{\xi}_K^n]^{\mathrm{T}}, n=1, 2, \cdots, N, \tag{7}$$

where N is the number of samples.

The empirical distribution is constructed as

$$\hat{\mathbb{P}}_N=\frac{1}{N}\sum_{n=1}^{N}\delta_{\hat{\xi}^n}, \tag{8}$$

where $\delta_{\hat{\xi}^n}$ denotes the Dirac measure concentrated at $\hat{\boldsymbol{\xi}}^n$.

Then, the Wasserstein ambiguity set is defined as

$$D=\{\mathbb{P}\in M(\Xi): d_{\mathrm{W}}(\mathbb{P}, \hat{\mathbb{P}}_N)\le\rho\}, \tag{9}$$

where $M(\Xi)$ denotes the set of all probability distributions supported on Ξ, $d_{\mathrm{W}}(\cdot,\cdot)$ is the Wasserstein metric, and ρ is the Wasserstein radius (Jiang YP et al., 2024). The support set Ξ of the uncertainty vectors $\boldsymbol{\xi}$ is specified as the following box uncertainty set:

$$\Xi=\{\boldsymbol{\xi}: \underline{\xi}_k\le\xi_k\le\overline{\xi}_k, k=1, 2, \cdots, K\}, \tag{10}$$

where ξ_k is the k^{th} component of $\boldsymbol{\xi}$, and $\underline{\xi}_k$ and $\overline{\xi}_k$ are its lower and upper bounds, respectively.

To balance the uncertainty-induced net power mismatch, an affine recourse policy is adopted for the adjustable generators, as shown in Eq. (11):

$$P_g(\boldsymbol{\xi})=\bar{P}_g-\alpha_g\boldsymbol{h}^{\mathrm{T}}\boldsymbol{\xi}, \sum_{g\in G_{\text{adj}}}\alpha_g=1, \alpha_g\ge0, \tag{11}$$

where $P_g(\boldsymbol{\xi})$ is the real-time active power output of generator g, $\bar{P}_g$ is its base-point output, and α_g is the participation factor. The term $\boldsymbol{h}^{\mathrm{T}}\boldsymbol{\xi}$ denotes the aggregated net power mismatch caused by uncertainties. The condition $\sum_{g\in G_{\text{adj}}}\alpha_g=1$ ensures full mismatch compensation by the adjustable generator set G_{adj}.

To preserve tractability, the branch active/reactive power flows and the squared current/voltage magnitudes are parameterized by the affine decision rules in Eq. (12):

$$\begin{cases}P_{ij}(\boldsymbol{\xi})=P_{ij}^0+\boldsymbol{A}_{ij}^{\mathrm{T}}\boldsymbol{\xi}, Q_{ij}(\boldsymbol{\xi})=Q_{ij}^0+\boldsymbol{B}_{ij}^{\mathrm{T}}\boldsymbol{\xi},\\ l_{ij}(\boldsymbol{\xi})=l_{ij}^0+\boldsymbol{C}_{ij}^{\mathrm{T}}\boldsymbol{\xi}, v_i(\boldsymbol{\xi})=v_i^0+\boldsymbol{D}_i^{\mathrm{T}}\boldsymbol{\xi},\end{cases} \tag{12}$$

where P_{ij}^0, Q_{ij}^0, l_{ij}^0, and v_i^0 are the nominal components, while $\boldsymbol{A}_{ij}^{\mathrm{T}}$, $\boldsymbol{B}_{ij}^{\mathrm{T}}$, $\boldsymbol{C}_{ij}^{\mathrm{T}}$, and $\boldsymbol{D}_i^{\mathrm{T}}$ are the corresponding recourse coefficient vectors associated with $\boldsymbol{\xi}$. Substituting Eqs. (11) and (12) into Eqs. (3)–(6) yields deterministic reformulations of the uncertainty-dependent constraints. For illustration, only the nodal voltage constraint is reformulated explicitly in inequality (13), while the remaining uncertainty-dependent constraints in Eqs. (4) and (5) can be derived analogously and are omitted for brevity.

$$\begin{aligned}(V_j^{\min})^2\le v_i^0+\boldsymbol{D}_i^{\mathrm{T}}\boldsymbol{\xi}-2r_{ij}(P_{ij}^0+\boldsymbol{A}_{ij}^{\mathrm{T}}\boldsymbol{\xi})\\ -2x_{ij}(Q_{ij}^0+\boldsymbol{B}_{ij}^{\mathrm{T}}\boldsymbol{\xi})+(r_{ij}^2+x_{ij}^2)(l_{ij}^0+\boldsymbol{C}_{ij}^{\mathrm{T}}\boldsymbol{\xi})\le(V_j^{\max})^2.\end{aligned} \tag{13}$$

The quadratic branch-flow relation $v_i l_{ij}=P_{ij}^2+Q_{ij}^2$ in Eq. (3) is the source of nonconvexity and is relaxed into the standard second-order-cone (SOC) form:

$$\left\|\begin{bmatrix}2P_{ij}(\boldsymbol{\xi})\\ 2Q_{ij}(\boldsymbol{\xi})\\ v_i(\boldsymbol{\xi})-l_{ij}(\boldsymbol{\xi})\end{bmatrix}\right\|_2\le v_i(\boldsymbol{\xi})+l_{ij}(\boldsymbol{\xi}), \tag{14}$$

where $\|\cdot\|_2$ denotes the Euclidean norm.

Substituting the affine state parameterization in Eq. (12) into inequality (14) yields the compact robust SOC constraint in inequality (15):

$$\|\boldsymbol{M}_{ij}\boldsymbol{\xi}+\boldsymbol{m}_{ij}\|_2\le\boldsymbol{r}_{ij}^{\mathrm{T}}\boldsymbol{\xi}+s_{ij}, \forall\boldsymbol{\xi}\in\Xi, \tag{15}$$

where $\boldsymbol{M}_{ij}$, $\boldsymbol{m}_{ij}$, $\boldsymbol{r}_{ij}$, and s_{ij} are constant terms determined by the nominal components and affine recourse coefficients. For the box uncertainty set Ξ, tractability is preserved through vertex enumeration. Since $\boldsymbol{\xi}$ is K-dimensional, Ξ contains 2^K vertices. Therefore, inequality (15) is equivalently enforced by checking the SOC condition at all vertices of Ξ, yielding

$$\|\boldsymbol{M}_{ij}\boldsymbol{\xi}^{(v)}+\boldsymbol{m}_{ij}\|_2\le\boldsymbol{r}_{ij}^{\mathrm{T}}\boldsymbol{\xi}^{(v)}+s_{ij}, \forall\boldsymbol{\xi}^{(v)}\in V(\Xi), \tag{16}$$

where $V(\Xi)$ denotes the vertex set of Ξ, and $\boldsymbol{\xi}^{(v)}$ is the v^{th} vertex.

To unify generator-output, nodal-voltage, and branch-current limits under uncertainties, the single-sided distributionally robust chance constraint (DRCC) in inequality (17) is introduced as

$$\inf_{\mathbb{P}\in D}\mathbb{P}\{\boldsymbol{a}_h^{\mathrm{T}}\boldsymbol{\xi}\le b_h\}\ge1-\delta_h, \forall h\in H, \tag{17}$$

where H is the index set of scalar security constraints, $\boldsymbol{a}_h$ and b_h are the coefficient vector and threshold of the h^{th} constraint, and δ_h is the risk level. This constraint ensures that each security constraint holds with a probability at least $1-\delta_h$ for all distributions in the ambiguity set D.

Following the tractable Wasserstein-based reformulation, the DRCC in inequality (17) is converted into the deterministic constraints in inequality (18):

$$\begin{cases} \rho w_h - \delta_h \eta_h \le \dfrac{1}{N}\sum\limits_{n=1}^{N} z_h^n, \ \forall h \in H, \\ z_h^n + \eta_h \le b_h - \boldsymbol{a}_h^{\mathrm{T}} \hat{\boldsymbol{\xi}}^n, \ \forall h \in H, \forall n, \\ z_h^n \le 0, \ \forall n, \forall h \in H, \\ \|\boldsymbol{a}_h\|_\infty \le w_h, \ w_h \ge 0, \ \eta_h \ge 0, \ \forall h \in H, \end{cases} \tag{18}$$

where z_h^n, w_h, and η_h are auxiliary variables. Specifically, z_h^n is associated with the n^{th} scenario of the h^{th} constraint, w_h bounds the dual norm $\|\boldsymbol{a}_h\|_\infty$, and η_h is a nonnegative slack variable. Hence, the original DRCC is reformulated as a finite set of tractable linear inequalities.

The ambiguity-set radius ρ controls the conservativeness of the DRO model and is determined according to Eq. (19) (Zhou et al., 2020):

$$\begin{cases} \rho = C\sqrt{\dfrac{1}{N}\ln\left(\dfrac{1}{1-\beta}\right)}, \\ C \approx 2\inf\limits_{\eta>0}\left(\dfrac{1}{2\eta}\left(1+\ln\left(\dfrac{1}{N}\sum\limits_{n=1}^{N} \mathrm{e}^{\eta\left\|\hat{\xi}^n-\hat{\boldsymbol{\mu}}\right\|_1^2}\right)\right)^{1/2}\right), \end{cases} \tag{19}$$

where $\|\cdot\|_1$ denotes the L_1 norm. $\hat{\boldsymbol{\mu}}$ denotes the sample mean, β is the prescribed confidence level, and C is a sample-dependent constant estimated by minimizing over the positive scalar auxiliary variable η, which is solved using the bisection method. Thus, the generated scenarios affect the DRO model through both the empirical distribution in Eq. (8) and the ambiguity radius in Eq. (19). Therefore, the original uncertainty-constrained problem is reformulated as a deterministic optimization.

3.3 Optimization-based search algorithm

Based on the deterministic reformulation in Section 3.2, the SR boundary is identified through an optimization-based search. In active distribution networks, the active power limits of the adjustable generators define the SR. For a target generator $g^* \in G_{\text{adj}}$, its output P_{g^*} is optimized, while the outputs of other generators in the set $G_{\text{adj}} \setminus \{g^*\}$ are fixed at prescribed conditioning variables in each run. By varying these conditional variables within their admissible ranges and repeatedly solving the resulting problem, discrete boundary points of the SR are obtained. Meanwhile, the nodal active and reactive load demands at bus j, denoted as P_j^{L} and Q_j^{L}, are fixed at the prescribed benchmark nodal active and reactive load values $P_j^{\text{L,pre}}$ and $Q_j^{\text{L,pre}}$, and the associated source–load uncertainty is represented by the DRO constraints in Section 3.2. The boundary-point search is formulated as

$$\begin{cases} \max & P_{g^*} \\ \text{s.t.} & \text{the first three constraints in Eq. (3)}, \\ & \text{Eqs. (11), (12), and (19)}, \\ & \text{inequalities (16) and (18)}, \\ & P_g = \bar{P}_g, \ \forall g \in G_{\text{adj}} \setminus \{g^*\}, \\ & P_j^{\text{L}} = P_j^{\text{L,pre}}, \ Q_j^{\text{L}} = Q_j^{\text{L,pre}}. \end{cases} \tag{20}$$

Solving problem (20) gives the extreme feasible value of P_{g^*} under a given conditioning vector, corresponding to one SR boundary point. Repeating this procedure over different conditioning values yields a representative set of boundary samples for the subsequent boundary-fitting model.

4 Deep learning model for boundary fitting

The above search algorithm can extract a set of boundary points that characterize the SR boundary. However, these discrete points cannot form a continuous and complete representation of the SR, making them difficult to apply directly for the online assessment. To address this issue, a Transformer-based DL method is presented, including its model architecture, core modules, training strategy, and evaluation metrics. Based on these boundary samples, a model capable of efficiently and accurately fitting the SR is ultimately constructed.

4.1 Model architecture

To achieve the construction of a high-dimensional SR from discrete boundary points, a deep neural network is designed to learn and fit the complex geometric dependencies among dimensions. The Transformer architecture, originally developed in the field of natural language processing (NLP) and renowned for its remarkable success, is innovatively introduced into the proposed framework. As illustrated in Fig. 3, the model mainly consists of four core modules: input masking and encoding, input embedding, Transformer encoder, and coordinate prediction.

An n-dimensional coordinate vector of an SR boundary point $[P_1, P_2, \cdots, P_n]$ can be analogized to a "sentence" of length n, where P_i denotes the active power output of the i^{th} generator and functions like a "word." Just as the meaning of a word in a sentence depends on its surrounding context, the feasible range of one coordinate within the SR strongly depends on the values of other coordinates. In this context, the Transformer's self-attention mechanism is capable of quantitatively capturing the dependency weights between any pair of coordinate dimensions, while the multi-head attention enables the model to learn from multiple feature subspaces and to capture complex interdependent relationships among them. Through this mechanism, the model overcomes the local-view limitations of traditional methods and learns the global geometric constraints of the SR boundary.

4.2 Core modules

4.2.1 Input masking and encoding

To enable the model to learn the intrinsic correlations among dimensions and to effectively adapt to the sequence processing mechanism of the Transformer, a self-supervised data augmentation and masking strategy is designed. This strategy expands each n-dimensional boundary point sample into n independent training instances, as illustrated in Fig. 4.

Specifically, when generating the i^{th} training instance, the i^{th} dimensional component P_i of the original boundary point coordinate vector $\boldsymbol{P}^{\text{point}}$ is "masked," producing two inputs: input mask vector $\boldsymbol{M}^{\text{mask}}$ with 1 at the target position i and 0 elsewhere and masked

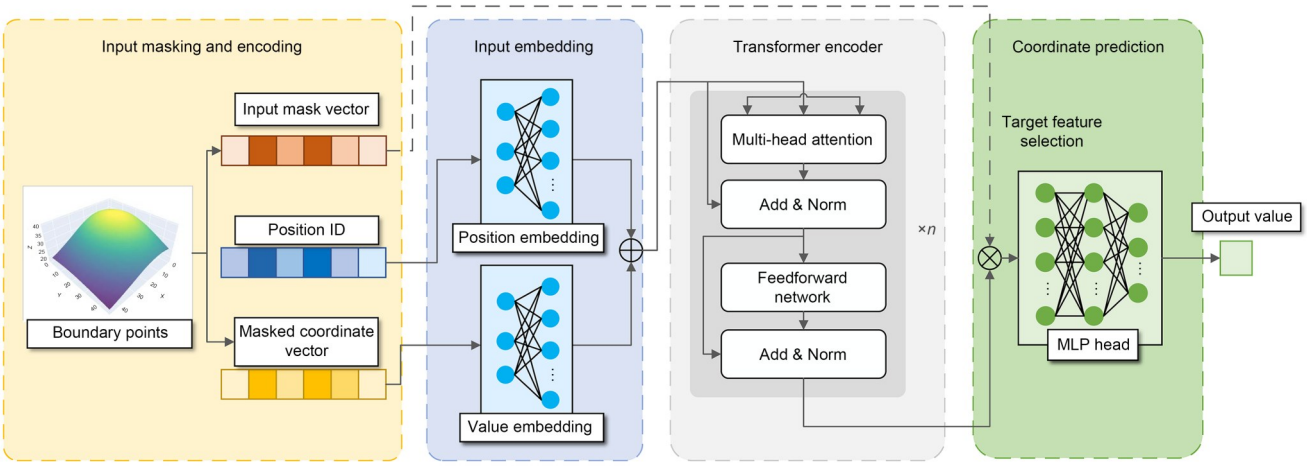

Fig. 3 Architecture of the Transformer-based SR boundary characterization model

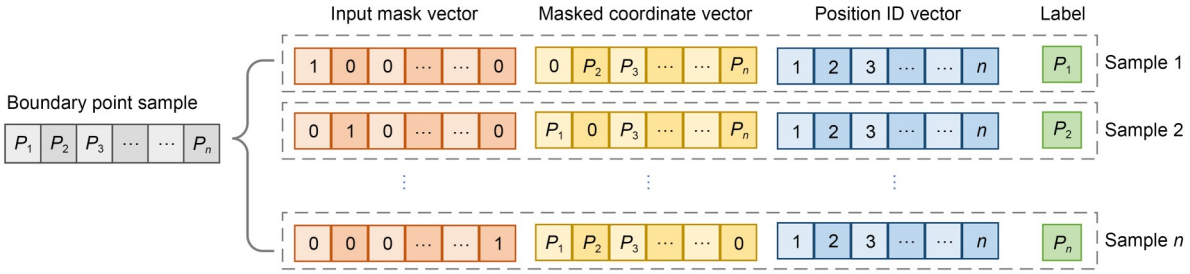

Fig. 4 Self-supervised sample augmentation and masking strategy

coordinate vector $\boldsymbol{P}^{\text{mask}}$ with the i^{th} dimension replaced by 0 and the remaining dimensions unchanged. Meanwhile, the original unmasked value P_i serves as the target label for this training instance.

In addition, since the Transformer architecture itself lacks the ability to process sequential order information, explicit positional information must be provided. Therefore, for each training instance, a position identifier (ID) vector $\boldsymbol{P}^{\text{ID}}$ is introduced, defined as an integer sequence $[1, 2, \cdots, n]$. This vector assigns a unique identifier to each coordinate dimension, enabling the model to learn dimension-specific feature representations in the subsequent embedding layer and thereby effectively distinguish different physical quantities. Through this strategy, the raw boundary points are transformed into a structured dataset ready for self-supervised training.

4.2.2 Input embedding

The input embedding module converts the pre-processed coordinate vectors into a high-dimensional embedding sequence $\boldsymbol{E} \in \mathbb{R}^{n \times d_{\text{model}}}$ for the Transformer encoder. First, to enrich feature representation, a linear layer projects each scalar in the masked coordinate vector $\boldsymbol{P}^{\text{mask}}$ into a d_{model}-dimensional space, generating the value embedding $\boldsymbol{E}_{\text{value}}$. Concurrently, a trainable embedding matrix maps each index in the position ID vector $\boldsymbol{P}^{\text{ID}}$ to a unique high-dimensional vector, forming the positional embedding $\boldsymbol{E}_{\text{pos}}$. This embedding provides the context information about the "identity" of each coordinate position. Finally, the input embedding sequence $\boldsymbol{E}$ is obtained by element-wise summation of the value embedding $\boldsymbol{E}_{\text{value}}$ and the positional embedding $\boldsymbol{E}_{\text{pos}}$, ensuring that the representation of each input element simultaneously encodes the information of both its coordinate value and position identity.

4.2.3 Transformer encoder

The Transformer encoder module is the core computational engine of the entire model, responsible for processing the high-dimensional feature sequences generated by the input embedding module. This module is composed of a stack of identical encoder layers, each consisting of a multi-head self-attention submodule and a point-wise feedforward network submodule, with residual connections and layer normalization applied to ensure the training stability.

The multi-head self-attention submodule performs global information exchange, dynamically updating each coordinate's features with the context from the entire input sequence. The feedforward network then applies a nonlinear transformation to each feature independently, thereby enhancing the model's representational capacity. Through multiple layers of iterative computation, the encoder outputs a refined boundary-aware feature sequence for the subsequent coordinate prediction.

4.2.4 Coordinate prediction

The coordinate prediction module decodes the boundary-aware feature sequence from the encoder to generate a precise prediction for the masked coordinate. First, the input mask vector $\boldsymbol{M}^{\text{mask}}$ selects the feature vector corresponding to the masked coordinate from the encoder's output sequence. Then, the extracted feature vector is fed into a prediction head composed of a multilayer perceptron (MLP), which nonlinearly maps the high-dimensional feature back to a one-dimensional space, producing the final predicted value at the masked position.

4.3 Model training

To enhance the training efficiency and generalization ability, a dynamic training strategy based on validation set performance is adopted. The training process aims to minimize the mean squared error (MSE) between the predicted values and the true labels. During training, the model's loss on the validation set is continuously monitored. If the validation loss ceases to decrease for a preset number of epochs, the learning rate is automatically reduced via an adaptive learning rate adjustment mechanism to help the model escape from local optima and reach a more precise solution. Concurrently, to prevent overfitting, an early stopping strategy is implemented. If the validation loss fails to improve for multiple epochs after a learning rate adjustment, the training is terminated prematurely, and the model weights that achieve the best performance on the validation set are saved as the final model.

4.4 Evaluation metrics

The performance of the proposed SR characterization approach is evaluated using the following three metrics: mean absolute percentage error (MAPE), root mean squared error (RMSE), and adjusted R-squared (ARS).

1. MAPE

MAPE is used to measure the relative accuracy of the model's predictions for the SR boundary points. A smaller value indicates a higher prediction accuracy. However, MAPE can excessively magnify errors when the true value is close to zero. MAPE is calculated as

$$\mathrm{MAPE}=\frac{1}{N_{\mathrm{test}}}\sum_{i=1}^{N_{\mathrm{test}}}\left|\frac{y_i-\hat{y}_i}{y_i}\right|\times 100\%, \tag{21}$$

where N_{test} is the total number of test samples, and y_i and $\hat{y}_i$ are the true and predicted coordinate values of the i^{th} sample, respectively.

2. RMSE

RMSE also measures the model's prediction accuracy. Because it is more sensitive to larger errors, it can better reflect the stability and deviation of the prediction results. A smaller value reflects a lower overall deviation between the predicted and true values. The expression of RMSE is

$$\mathrm{RMSE}=\sqrt{\frac{1}{N_{\mathrm{test}}}\sum_{i=1}^{N_{\mathrm{test}}}(y_i-\hat{y}_i)^2}. \tag{22}$$

3. ARS

To further evaluate the goodness of fit of the model from a statistical perspective, that is, its ability to capture the overall geometric shape of the SR boundary, ARS is introduced as a supplementary evaluation metric (Xie, 2021). ARS considers the number of input features when evaluating, introducing a penalty for model complexity. Therefore, it provides a more unbiased measure of the model's explanatory power. A value closer to 1 suggests a more precise fit and stronger explanatory power. The calculation formula is

$$\mathrm{ARS}=1-\frac{f_{\mathrm{RSS}}/(N_{\mathrm{test}}-n_{\mathrm{f}}-1)}{f_{\mathrm{TSS}}/(N_{\mathrm{test}}-1)}, \tag{23}$$

where f_{RSS} is the residual sum of squares between the predicted and actual boundaries; f_{TSS} is the total sum of squares of the actual boundaries; n_{f} denotes the number of input features used in the ARS evaluation.

5 Case studies

5.1 Simulation setup

To validate the effectiveness of the proposed DRO and DL-based approach for power grid SR characterization, case studies are conducted on the IEEE 33-bus and IEEE 123-bus systems shown in Fig. 5. Compared with the standard IEEE 33-bus system (Dolatabadi et al., 2021), the system in Fig. 5a is modified by placing a wind turbine at bus 8 and conventional generators at buses 7, 21, 24, and 30. Additionally, the loads at buses 3 and 6 are modeled as uncertain loads. Similarly, the IEEE 123-bus system (Bharati et al., 2023) in Fig. 5b is adapted by adding wind turbines at buses 7 and 53 and conventional generators at buses 18, 24, 46, 60, 77, and 105, and by treating the loads at buses 13 and 76 as stochastic. Since the standard IEEE 33-bus and 123-bus benchmarks lack controllable generation, conventional generators are added to describe the feasible operation region. Our approach is a general-purpose method that remains effective irrespective of the location of these generators within the network. The detailed parameters of the added generators are listed in Table 1. All simulations are performed on a workstation with an Intel® Core™ i9-14900K CPU and an NVIDIA GeForce RTX 4090 GPU. The algorithms are implemented in Python, and the models are built and trained using the PyTorch DL framework.

The historical wind power data used to train the scenario generation model consist of one year of offshore measurements from Eastern China sampled every 15 minutes. Based on this dataset, the diffusion model described in Section 3.1 is employed to generate 500 high-fidelity wind power scenarios. As this study focuses on the SR at a single time instance, only the values at 12:00 from all scenarios are extracted for the subsequent analysis. Representative generated scenarios together with the forecast are depicted in Fig. 6. The scenarios generated by the diffusion model follow the forecast trend and encompass it, validating the effectiveness of diffusion model-based scenario generation.

Both the renewable uncertainty and stochastic load deviations are considered in the case studies. The load uncertainty is represented by deviation samples around the corresponding typical load curves. These samples are then combined with the diffusion-generated renewable scenarios to form the joint uncertainty samples used in the DRO-based boundary search.

5.2 Performance and validation

5.2.1 Dataset construction

Taking the IEEE 33-bus system as an example, the construction of the dataset for model training and validation is illustrated. The generator outputs at buses 21 and 24 are set as conditional variables and the generator output at bus 7 as the target variable, denoted as PG_1, PG_2, and PG_3, respectively. By uniformly discretizing the conditional variables within their feasible ranges and performing boundary solving of the target variable for each sampled combination, a dataset containing all feasible boundary points is ultimately

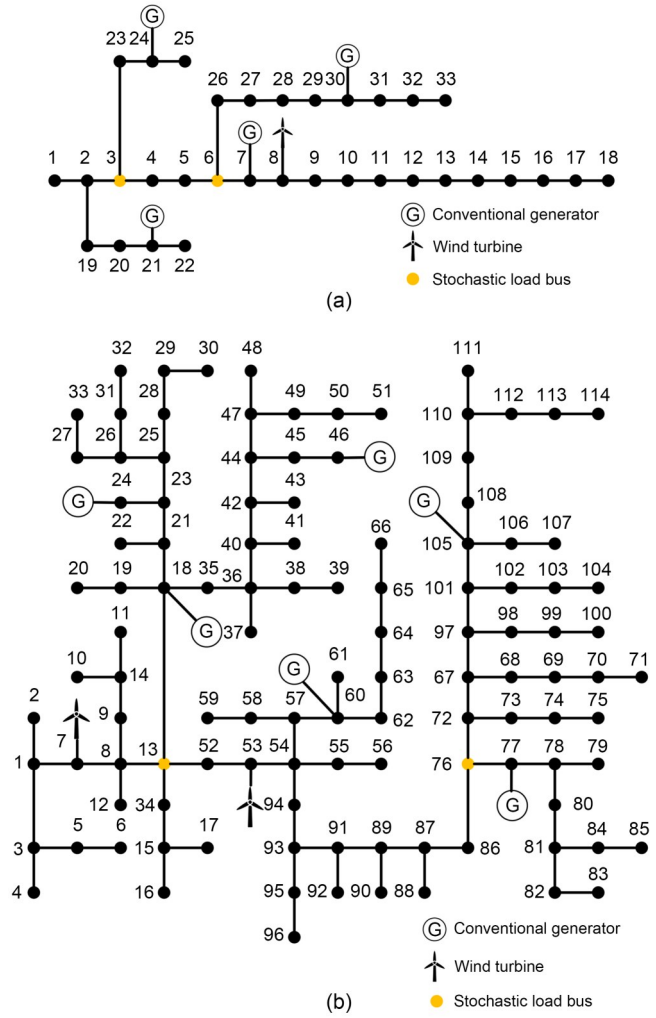

Fig. 5 IEEE 33-bus (a) and IEEE 123-bus (b) systems used in this study

Table 1 Parameters of the added conventional generators

System	Generator location (bus No.)	$P_g^{\max}$ (MW)	$P_g^{\min}$ (MW)
IEEE 33-bus	7, 21, 24, 30	2	0
IEEE 123-bus	18, 24, 46, 60, 77, 105	2	0

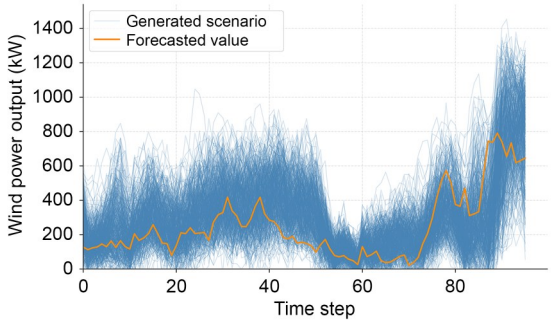

Fig. 6 Generated wind power scenarios using the diffusion model

generated, which is then randomly split into 80% for training and 20% for validation.

To evaluate the generalization of the model, an independent ACPF-based test set is used to assess the physical consistency of the generated boundary points. This test set is not only employed to assess the generalization performance of the model on out-of-sample data, but also serves as a benchmark to verify the consistency between the training data generated from the DistFlow power flow approximation model and the actual physical boundaries of the system.

5.2.2 Convergence analysis

The proposed approach is compared against four baseline models on the IEEE 33-bus system: convolutional neural network (CNN), graph neural network (GNN), MLP, and recurrent neural network (RNN). All models are trained and evaluated under identical conditions.

Fig. 7 displays the training loss (MSE) curves for the five models. The proposed Transformer-based model demonstrates a superior performance in both the convergence speed and the final accuracy. Its training loss drops rapidly within the first 50 epochs and stabilizes at the lowest level, outperforming the other baseline models, which show clear gaps in either convergence speed or the final precision. Furthermore, the proposed model exhibits strong generalization, as its validation loss closely tracks the training loss without significant divergence. This indicates that the model effectively avoids overfitting and confirms its superior learning efficiency and fitting capability.

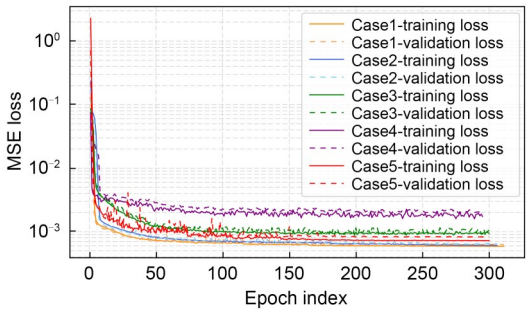

Fig. 7 MSE loss convergence for different models on the IEEE 33-bus system (Case 1: the proposed Transformer-based model. Case 2: a CNN-based model. Case 3: a GNN-based model. Case 4: an MLP-based model. Case 5: an RNN-based model)

5.2.3 Accuracy assessment and computational time

To show the performance on the unseen data, all models are evaluated on an independent ACPF-based test set. Table 2 summarizes the fitting accuracy. Table 3 reports the offline training time and the average inference time required for a boundary point.

Table 2 Accuracy comparison of different models on the test set of the IEEE 33-bus and IEEE 123-bus systems

Case index	MAPE (%)		RMSE		ARS ($\times 10^{-2}$)	
	IEEE 33-bus	IEEE 123-bus	IEEE 33-bus	IEEE 123-bus	IEEE 33-bus	IEEE 123-bus
1	0.29	0.38	0.39	0.57	99.99	99.99
2	0.68	0.75	0.87	0.89	99.97	99.97
3	1.97	1.92	2.27	2.34	99.88	99.75
4	2.40	2.51	3.35	3.67	99.66	99.61
5	1.28	1.47	1.63	1.87	99.89	99.91

Table 3 Computational time comparison of the IEEE 33-bus and IEEE 123-bus systems

Case index	Training time (s)		Inference time (ms)	
	IEEE 33-bus	IEEE 123-bus	IEEE 33-bus	IEEE 123-bus
1	128.63	118.69	0.5215	0.5276
2	94.57	89.67	0.3306	0.3298
3	134.67	129.56	0.5389	0.5316
4	58.40	54.14	0.2917	0.2899
5	114.06	112.98	0.5067	0.5012

The proposed Transformer-based model achieves the best accuracy across all metrics. Although CNN, MLP, and RNN are slightly faster, all compared models operate at the millisecond level in inference, so the speed difference is negligible for online applications. For this type of boundary-sensitive problem, even seemingly moderate prediction errors may cause noticeable deviations in the reconstructed feasible boundary and may further affect the reliability of security assessment near the boundary. From this perspective, the accuracy improvement of the proposed Transformer-based model is practically meaningful and justifies the slight increase in the computational cost.

In addition, the training time depends on the model architecture, input dimension, and sample size. Consequently, Table 3 shows similar training times for the IEEE 33-bus and IEEE 123-bus systems. The inference times for the two test systems are also nearly identical, as inference time is primarily determined by the model architecture, which remains constant across both systems. Thus, the results demonstrate the consistent computational efficiency of the proposed approach when applied to different grid scales.

5.2.4 Validation of power flow approximation

The proposed DRO-based boundary-point search is formulated on the DistFlow model, for which a tractable reformulation is available, whereas no directly comparable tractable counterpart exists under the standard ACPF model. Hence, the ACPF model is used only for ex-post validation. The results show that the model trained on DistFlow-based boundary points still achieves very high fitting accuracy on an independent ACPF-based test set, indicating that the DistFlow-determined boundaries are highly consistent with the actual physical boundaries and sufficiently accurate for subsequent DL model training.

Fig. 8 shows the three-dimensional SR boundary surface for both the IEEE 33-bus and IEEE 123-bus systems, as reconstructed by the proposed model. The space enclosed by this surface and the three coordinate planes, i.e., the *XY*-, *XZ*-, and *YZ*-planes, constitutes the determined SR. For a clearer assessment, Fig. 9 depicts a two-dimensional slice for the IEEE 33-bus and IEEE 123-bus systems, comparing the fitted boundary (orange line) with the true ACPF-based boundary points (blue dots). The high degree of overlap visually confirms the accuracy of the proposed approach.

5.3 Scalability and flexibility

5.3.1 Scalability and observational flexibility

A significant advantage of the proposed approach lies in its flexibility with respect to observation directions. While a single

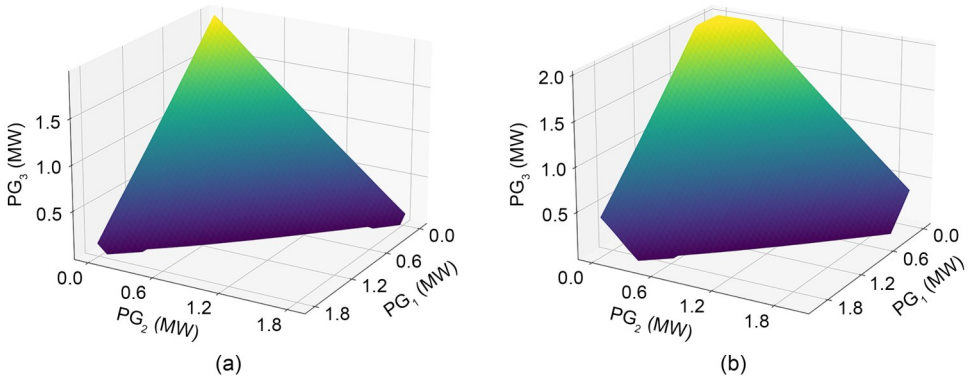

Fig. 8 Visualization of the constructed SRs: (a) IEEE 33-bus system; (b) IEEE 123-bus system

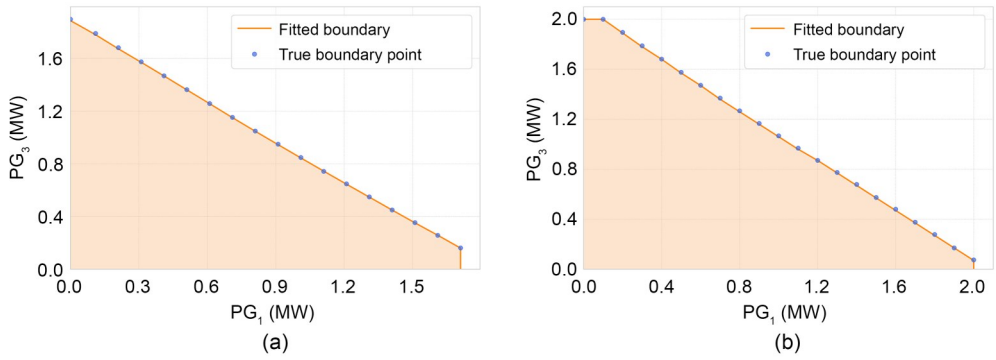

Fig. 9 Two-dimensional SR accuracy comparison: (a) IEEE 33-bus system; (b) IEEE 123-bus system

target variable is optimized at the boundary point search stage, the trained DL model enables multi-dimensional cross-prediction. Taking the three-dimensional SR as an example, the same model can perform predictions in $XY \rightarrow Z$, $XZ \rightarrow Y$, and $YZ \rightarrow X$ directions by applying different conditional inputs and input mask vector $\boldsymbol{M}^{\text{mask}}$. The capability is verified and the results are presented in Table 4.

Table 4 Model accuracy for different test systems and observation directions

Direction	MAPE (%)	
	IEEE 33-bus system	IEEE 123-bus system
$XY \rightarrow Z$	0.25	0.28
$XZ \rightarrow Y$	0.30	0.52
$YZ \rightarrow X$	0.32	0.34
Average	0.29	0.38

As shown in Table 4, the proposed approach exhibits extremely high fitting accuracy across different observation directions for both the IEEE 33-bus and IEEE 123-bus systems. These results clearly demonstrate two principal merits of the proposed approach: (1) the approach has good scalability, enabling effective application to power systems of different scales; (2) the trained model is multi-functional, allowing for flexible, multi-perspective observation and evaluation of SR.

5.3.2 High-dimensional analysis

Characterizing high-dimensional SR has long been a bottleneck for traditional methods because of their invisibility. To assess the proposed approach's performance in a high-dimensional space, a five-dimensional SR for the IEEE 123-bus system is constructed and evaluated.

The study chooses the active power outputs of five generators as the observation variables to create a five-dimensional boundary point dataset. An independent ACPF-based test set of 300 five-dimensional boundary points is generated. During evaluation, one dimension is randomly selected and predicted from the other four dimensions.

As shown in Table 5, the low MAPE and RMSE values indicate that the approach can successfully learn the complex geometry of the five-dimensional SR. It overcomes the reliance of traditional approaches on dimensionality reduction mapping, demonstrating that the approach is a scalable analytical tool with significant advantages in handling high-dimensional SR.

Table 5 Model fitting accuracy for a five-dimensional SR of the IEEE 123-bus system

Observation dimension	MAPE (%)	RMSE	ARS ($\times 10^{-2}$)
Five-dimensional	0.51	0.61	99.99

5.4 Robustness analysis

For the sake of visualization, the robustness of the proposed SR boundary characterization approach is analyzed.

5.4.1 Comparison of uncertainty handling

Handling renewable and load uncertainty is a core challenge for power system security. As a tool that intuitively illustrates the system's operating boundaries, the practical value of the SR guiding schedule decisions depends directly on its ability to effectively and reasonably account for such uncertainty. Three approaches are compared to demonstrate that our proposed DRO-based approach provides a better balance between robustness and conservatism. Fig. 10 illustrates the SR of the IEEE 33-bus and IEEE 123-bus systems constructed by the three approaches:

Case 1: the proposed approach.

Case 2: a deterministic approach using forecast values.

Case 3: a traditional RO approach from the reference (Chen et al., 2019) for the worst-case scenario.

The hierarchical relationship among the three SRs can be clearly observed in Fig. 10. The deterministic approach ignores the uncertainty, resulting in an overly optimistic boundary that encloses the largest area. Conversely, the SR derived by the traditional RO approach is the smallest and the most conservative; while it offers the strongest security guarantee by defending against the worst-case scenario, it does so at the cost of an overly compressed SR that severely limits the operational flexibility. Our proposed DRO-based approach achieves a balance between these two extremes. This approach rigorously accounts for the uncertainty by constructing

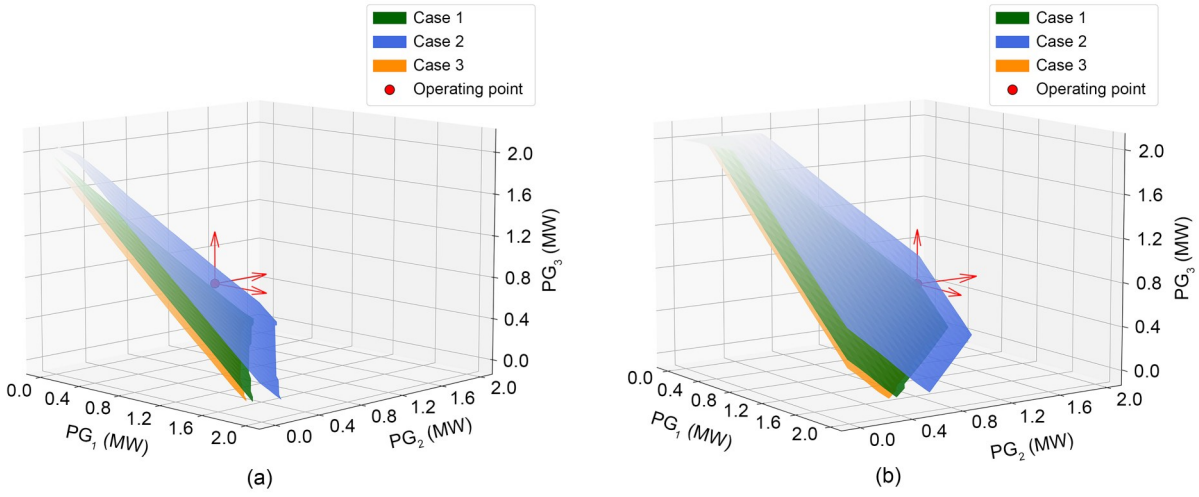

Fig. 10 SRs under different uncertainty handling approaches: (a) IEEE 33-bus system; (b) IEEE 123-bus system

an ambiguity set of probability distributions, while avoiding over-reaction to extreme events with very low probabilities.

5.4.2 Analysis of key parameters

Two key tunable parameters that control the DRO-based SR boundary are analyzed: the confidence level β for the ambiguity set and the risk parameter δ for the chance constraint.

The confidence level β determines the size of the data-driven ambiguity set and is directly related to the approach's conservatism. For ease of illustration, the following analysis uses a two-dimensional SR from a specific observation perspective. Fig. 11 shows the SR boundaries of the IEEE-33 bus and IEEE-123 bus systems at different confidence levels with a fixed risk parameter δ=10%. As β rises from 80% to 99%, the SR boundary contracts inward, and the feasible area decreases. A higher confidence level yields a larger ambiguity set containing more plausible uncertainty distributions. As the ambiguity set radius ρ increases, more low-probability extreme scenarios are included, and the feasible region found by the search model will correspondingly decrease.

The risk parameter δ defines the maximum violation risk the system can tolerate under the worst-case probability distribution. Fig. 12 shows the SR boundaries of the IEEE-33 bus and IEEE-123 bus systems for different risk parameters at a fixed confidence level β=95%. The SR shrinks as δ decreases from 20% to 1%. A lower risk parameter implies that the uncertainty-related chance constraints must be satisfied with a higher probability under the worst-case distribution. To satisfy these stricter constraints, the optimization model must remain feasible under a wider range of uncertainty realizations, which ultimately leads to a contraction of the SR.

Therefore, the confidence level β and risk parameter δ are two key tunable parameters in the proposed DRO-based framework. They provide a flexible mechanism to adjust the SR, balancing the trade-off between the excessive conservatism of the RO-based approach and the over-optimism of the deterministic one.

5.4.3 Application of the proposed SR approach

Taking the operating point (red dot) in Fig. 10 as an example, although it appears secure in the deterministic approach, it becomes insecure once uncertainty is considered. Furthermore, in the actual operations, the system must retain a security margin to withstand unforeseen disturbances, and traditional SR solving methods struggle to incorporate an explicitly quantified margin. The core advantage of the proposed approach lies in its adjustability. Based on the system's operating state and security criteria, decision-makers can flexibly construct a "customized" SR with an embedded, quantified security margin by jointly adjusting the confidence level and the risk parameter. The resulting region is robust yet not overly conservative, offering a practical boundary for scheduling decisions.

6 Conclusions and future work

In this paper, a general approach is proposed to characterize the high-dimensional SR of power systems under uncertainties from

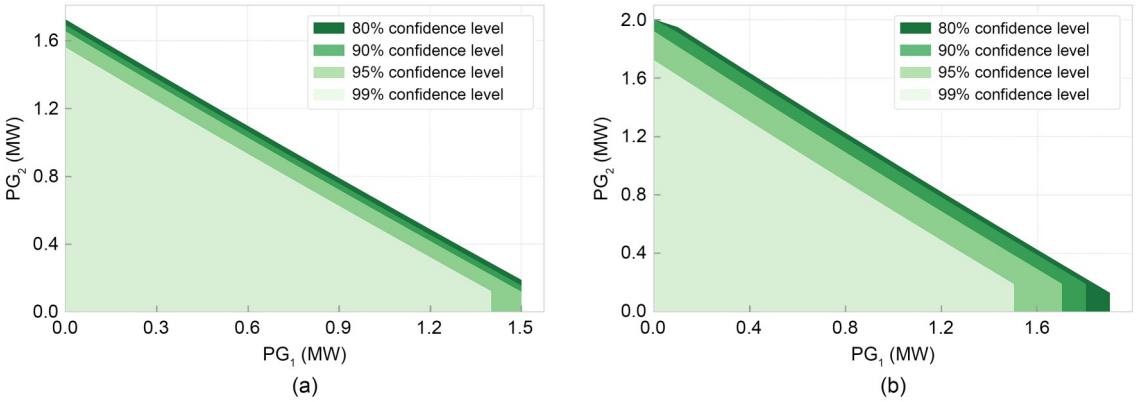

Fig. 11 SRs at different confidence levels: (a) IEEE 33-bus system; (b) IEEE 123-bus system

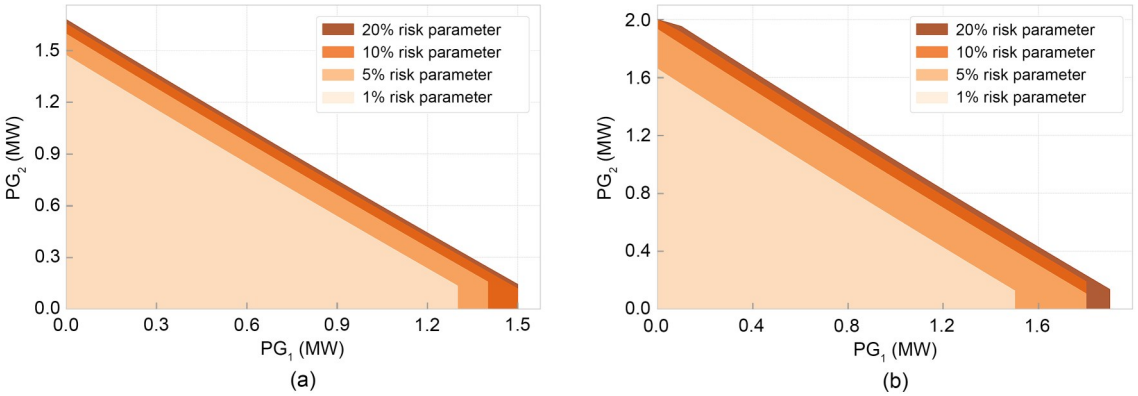

Fig. 12 SRs at different risk parameters: (a) IEEE 33-bus system; (b) IEEE 123-bus system

high renewable energy penetration, using a framework consisting of DRO-based boundary point search and DL-based boundary fitting. The Transformer-based DL model accurately fits the SR, overcoming the curse of dimensionality common to traditional methods in high-dimensional spaces and showing good universality. A comparison of uncertainty handling methods and a sensitivity analysis of key parameters confirm that our approach effectively balances robustness and conservatism. Its inherent tunability provides a practical and economically valuable decision support tool for grid operators.

In our future work, the approach may be generalized to other networks, such as microgrids and integrated energy systems. Moreover, its core concept can be applied to characterize other operational domains, including energy-carbon or economic feasibility regions.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. U22B2098).

Author contributions

Yuekai CHEN conceptualized the study, developed the methodology, implemented the simulations, performed the validation and formal analysis, created the visualizations, and drafted the paper. Zhejing BAO handled the writing (review and editing), supervised the work, provided resources, and acquired funding. Miao YU handled writing (review and editing) and supervised the work.

Conflict of interest

All the authors declare that they have no conflict of interest.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declaration on the use of generative AI tools

During the preparation of this work, the authors used ChatGPT in order to improve language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

Aryani DR, Song H, 2024. A review on power system security issues in the high renewable energy penetration environment. *J Electr Eng Technol*, 19:4649-4665. https://doi.org/10.1007/s42835-024-02028-x

Avila OF, Passos Filho JA, Peres W, 2021. Steady-state security assessment in distribution systems with high penetration of distributed energy resources. *Electr Power Syst Res*, 201:107500. https://doi.org/10.1016/j.epsr.2021.107500

Bharati AK, Ajjarapu V, Du W, et al., 2023. Role of distributed inverter-based-resources in bulk grid primary frequency response through HELICS based SMTD co-simulation. *IEEE Syst J*, 17(1):1071-1082. https://doi.org/10.1109/JSYST.2022.3218117

Chen S, Wei Z, Sun G, et al., 2019. Convex hull based robust security region for electricity-gas integrated energy systems. *IEEE Trans Power Syst*, 34(3):1740-1748. https://doi.org/10.1109/TPWRS.2018.2888605

Dai W, Yang Z, Yu J, et al., 2019. Security region of renewable energy integration: characterization and flexibility. *Energy*, 187:115975. https://doi.org/10.1016/j.energy.2019.115975

Dolatabadi SH, Ghorbanian M, Siano P, et al., 2021. An enhanced IEEE 33 bus benchmark test system for distribution system studies. *IEEE Trans Power Syst*, 36(3): 2565-2572. https://doi.org/10.1109/TPWRS.2020.3038030

Feng J, Ren ZY, Jiang YP, et al., 2024. Committed carbon emissions operation regions of power system: concept and method. *Proc CSEE*, 44(22):8846-8859 (in Chinese). https://doi.org/10.13334/j.0258-8013.pcsee.231021

Gao Y, Ren Z, Jiang Y, et al., 2023. Analysis method of committed carbon emission operational region for electricity-hydrogen coupling system. *Electr Power Autom Equip*, 43(12):29-36 (in Chinese). https://doi.org/10.16081/j.epae.202310028

Jiang T, Zhang R, Li X, et al., 2021. Integrated energy system security region: concepts, methods, and implementations. *Appl Energy*, 283:116124. https://doi.org/10.1016/j.apenergy.2020.116124

Jiang YP, Ren ZY, Lu CH, et al., 2024. A region-based low-carbon operation analysis method for integrated electricity-hydrogen-gas systems. *Appl Energy*, 355:122230. https://doi.org/10.1016/j.apenergy.2023.122230

Jin Y, Acquah MA, Seo M, et al., 2023. Optimal EV scheduling and voltage security via an online bi-layer steady-state assessment method considering uncertainties. *Appl Energy*, 347:121356. https://doi.org/10.1016/j.apenergy.2023.121356

Li S, Xiong H, Chen Y, 2024. DiffCharge: generating EV charging scenarios via a denoising diffusion model. *IEEE Trans Smart Grid*, 15(4):3936-3949. https://doi.org/10.1109/TSG.2024.3360874

Li X, Zhang LW, Jiang T, et al., 2021. General algorithm for exploring security region boundary in power systems using Lagrange multiplier. *Proc CSEE*, 41(15):5139-5152 (in Chinese). https://doi.org/10.13334/j.0258-8013.pcsee.200860

Lin W, Yang Z, Yu J, et al., 2021. Tie-line security region considering time coupling. *IEEE Trans Power Syst*, 36(2):1274-1284. https://doi.org/10.1109/TPWRS.2020.3015483

Lin W, Jiang H, Yang Z, 2022. Tie-line security regions in high dimension for renewable accommodations. https://doi.org/10.48550/arXiv.2201.01019

Lin W, Jiang H, Jian HJ, et al., 2023. High-dimension tie-line security regions for renewable accommodations. *Energy*, 270:126887. https://doi.org/10.1016/j.energy.2023.126887

Liu L, Wang D, Hou K, et al., 2020. Region model and application of regional integrated energy system security analysis. *Appl Energy*, 260:114268. https://doi.org/10.1016/j.apenergy.2019.114268

Liu W, Wang CG, Cao Y, et al., 2025. A method for generating wind power output scenarios based on improved conditional generative diffusion model. *Electr Power Syst Res*, 247:111779. https://doi.org/10.1016/j.epsr.2025.111779

Monteiro MR, Alvarenga GF, Rodrigues YR, et al., 2020. Network partitioning in coherent areas of static voltage stability applied to security region enhancement. *Int J Electr Power Energy Syst*, 117:105623. https://doi.org/10.1016/j.ijepes.2019.105623

Nguyen HD, Dvijotham K, Turitsyn K, 2019. Constructing convex inner approximations of steady-state security regions. *IEEE Trans Power Syst*, 34(1):257-267. https://doi.org/10.1109/TPWRS.2018.2868752

Rahimian H, Mehrotra S, 2019. Distributionally robust optimization: a review. https://doi.org/10.48550/arXiv.1908.05659

Su J, Chiang HD, Zeng Y, et al., 2021. Toward complete characterization of the steady-state security region for the electricity-gas integrated energy system. *IEEE Trans Smart Grid*, 12(4):3004-3015. https://doi.org/10.1109/TSG.2021.3065501

Sun D, Yu Y, 2023. Accurate identification of critical boundary hyperplanes of practical steady-state security region in distribution grids. *IEEE Trans Smart Grid*, 14(6): 4312-4321. https://doi.org/10.1109/TSG.2023.3262693

Teng F, Zhang YX, Yang TK, et al., 2024. Distributed optimal energy management for We-Energy considering operation security. *IEEE Trans Netw Sci Eng*, 11(1):225-235. https://doi.org/10.1109/TNSE.2023.3295079

Tinoco RAG, Passos Filho JA, Peres W, et al., 2021. A new particle swarm optimization-based methodology for determining online static security regions. *Int Trans Electr Energy Syst*, 31(3):e12790. https://doi.org/10.1002/2050-7038.12790

Wu F, Kumagai S, 1982. Steady-state security regions of power systems. *IEEE Trans Circ Syst*, 29(11):703-711. https://doi.org/10.1109/TCS.1982.1085091

Wu FF, Tsai YK, Yu YX, 1988. Probabilistic steady-state and dynamic security assessment. *IEEE Trans Power Syst*, 3(1):1-9. https://doi.org/10.1109/59.43173

Wu XW, Zhang B, Nielsen MP, et al., 2023. Neural network based feasible region approximation model for optimal operation of integrated electricity and heating system. *CSEE J Power Energy Syst*, 9(5):1808-1819. https://doi.org/10.17775/cseejpes.2022.09040

Xiao J, Li C, She B, et al., 2024. Distribution system security region with energy storage systems. *Energy*, 313:133841. https://doi.org/10.1016/j.energy.2024.133841

Xie W, 2021. On distributionally robust chance constrained programs with Wasserstein distance. *Math Program*, 186(1-2):115-155. https://doi.org/10.1007/s10107-019-01445-5

Yorino N, Abdillah M, Sasaki Y, et al., 2018. Robust power system security assessment under uncertainties using bi-level optimization. *IEEE Trans Power Syst*, 33(1):352-362. https://doi.org/10.1109/TPWRS.2017.2689808

Zhang S, Gu W, Zhang XP, et al., 2024. Steady-state security region of integrated energy system considering thermal dynamics. *IEEE Trans Power Syst*, 39(2):4138-4153. https://doi.org/10.1109/TPWRS.2023.3296080

Zhang ZY, Yang ZB, Yau DKY, et al., 2023. Data security of machine learning applied in low-carbon smart grid: a formal model for the physics-constrained robustness. *Appl Energy*, 347:121405. https://doi.org/10.1016/j.apenergy.2023.121405

Zhang ZY, Liu MX, Sun MY, et al., 2024. Vulnerability of machine learning approaches applied in IoT-based smart grid: a review. *IEEE Int Things J*, 11(11):18951-18975. https://doi.org/10.1109/JIOT.2024.3349381

Zhou A, Yang M, Wang M, et al., 2020. A linear programming approximation of distributionally robust chance-constrained dispatch with Wasserstein distance. *IEEE Trans Power Syst*, 35(5):3366-3377. https://doi.org/10.1109/TPWRS.2020.2978934