Computing and Informatics, Vol. 45, 2026, 1272–1304, doi: 10.31577/cai_2026_5_1272

OPTIMAL ECONOMIC DISPATCHING IN DISTRIBUTION NETWORK INVOLVING DATA AUGMENTATION AND POWER SOURCE-LOAD UNCERTAINTIES

Yunjia Wang, Ning Pang, Hongguang Yu, Yan Zhang, Guozhen Ma, Shiyao Hu

Economic Research Institute
State Grid Hebei Electric Power Co., Ltd.
Shijiazhuang, China

Yang Liu*

College of Electrical Engineering, Sichuan University
Chengdu, China
e-mail: Yang.Liu@scu.edu.cn

Abstract. The uncertainties of wind power generation and load significantly impact on the economy of distribution network dispatching. Therefore, operating an optimal dispatching strategy can improve the economic performance of the power grid. In modern data-driven optimization approaches, the proper volume of data generates remarkably optimal results. However, due to the difficulties of data collection, transmission, and storage, wind power data is naturally scarce, which severely deteriorates the performance of the optimization. Additionally, the uncertainty existing on the load side leads to more complicated scenarios in optimization, which further prevents the optimization from achieving the optimal dispatching strategy. Therefore, to solve the aforementioned issues, this paper presents an economic dispatching optimization approach for a distribution network while considering wind power data augmentation and uncertainties of both wind power and load. Firstly, the synthetic minority over-sampling technique (SMOTE) is adopted to generate a proper number of wind power generation output data samples. And then, a variational autoencoder (VAE) is further employed to generate adequate data samples

* Corresponding author

with a similar distribution to the original data samples. Secondly, the deep embedded clustering algorithm (DECA) deeply characterizes and clusters the augmented data samples to figure out typical wind power generation output scenarios over-represented in the samples. Thirdly, for each typical scenario, an ambiguity set is constructed using an imprecise Dirichlet data-driven model (IDDM); And then, the details, including the objective functions and constraints of the economic dispatching for the distribution network are analyzed and modeled. Finally, a two-stage distributed robust optimization (DRO) model considering the uncertainties of both wind power generation output and load for achieving the optimal economic dispatching strategy for the distribution network is established and solved; The case studies demonstrate the effectiveness of the presented approach.

Keywords: Data augmentation, uncertainty, ambiguity set, distributionally robust optimization, economic dispatching, distribution network

Mathematics Subject Classification 2010: 90C15, 90C90

1 INTRODUCTION

In recent years, renewable energy represented by wind power and photovoltaic power has been developed rapidly. Especially due to advantages such as the richness of wind resources and the compact land footprint, wind turbines have been deployed widely in modern power systems [1]. Moreover, because of verticalization and miniaturization of wind turbine, the penetration of distributed wind power has been increased sharply in distribution networks, which significantly turns the traditional distribution networks into active ones [2, 3]. Although the power supply becomes more adaptable and convenient in an active distribution network, it should be realized that partially shifting the energy supply source from the main grid to distributed generation introduces a series of problems that must be resolved in an active distribution network [4, 5, 6, 7, 8]. Krahmer et al. [4] reported that due to the increasing number of distributed renewable energy sources in distribution networks, the stability analysis is a vital aspect for network operators. Therefore, the authors focused on characterizing Q(V) feature of the distribution network and further parameterized the reactive power control of the distributed energy resources to maintain the stability of the network. Nirbhavane et al. [5] also pointed out that the distribution system voltage stability analysis with distributed energy resources allows high-penetration of renewable energy necessary for the sustainability of the system. Ge and Wu [6] pointed out the distribution network security issue caused by the distributed energy resources, especially wind power. Therefore, the authors employed a convolutional neural network to implement the online security analysis for the distribution network with involving the features of node voltage, current, power, and the location of distributed power generation. Hu et al. [7] reported that the distributed power

generations significantly impact on both the reliability and planning of the distribution network. Yang et al. [8] further reported that the traditional system planning algorithms or approaches have limited abilities to cope with modern planning tasks for distribution networks with distributed wind power generation.

From the above research, issues caused by distributed power generation, especially wind power generation, can be clearly identified. However, several studies [9, 10, 11, 12, 13] suggested that although the issues indeed exist in the fields of stability, security, and planning of distribution networks, the dispatching stage of the distribution network can provide compatibility to even out the issues to some extent. The distribution network benefits from the adjusting ability of the flexible resources within the network during the dispatching. However, the researchers also pointed out that the extensive fluctuations of wind power generation output may exceed the regulation capability of the distribution network. This point leads to significant impact on the economy of dispatching. Therefore, figuring out an optimal dispatching strategy is highly essential for improving the performance of the distribution network. Wang et al. [12] presented a day-ahead and intraday joint optimal dispatch method for an active distribution network. In their work, the energy storage system is considered to even out the fluctuations of the renewable energy sources. By optimizing the dispatching of the centralized energy storage, distributed energy storage, and the other adjustable resources, their method shows significant improvement of the dispatching economy. Liu et al. [14] presented a dispatching model for distribution network with involving multiple active management measures. The authors also built up an optimal model using mixed-integer linear programming. Regarding the energy fluctuations, these researches solved the issues in a multiple time resolution way. In each short time interval, the fluctuation is regarded as a predictable and fixed value. However, it is known that wind power generation output is highly uncertain. Therefore, these deterministic optimizations show limited performances. As a result, Liu and Nair [15] and Soares et al. [16] presented stochastic optimization-based distribution network dispatching strategies involving the uncertainty of wind power. However, their models highly rely on the probability distribution function of uncertain wind power which is extremely difficult to be explicitly achieved from the wind power generation output data samples. In order to eliminate the reliance on the probability distribution, Liu et al. [17] presented a robust optimization-based dispatching strategy. In their work, a concept of uncertainty set which is directly constructed from the wind power generation output data samples is employed. However, although the uncertainty set is independent of the probability distribution function, its performance still depends on an accurate wind forecasting error which also cannot be accurately obtained. Therefore, Chen et al. [18] and Wu et al. [19] further presented adjustable robust optimization (ARO) in enabling the optimal dispatching strategy. By employing an adjustable parameter, their uncertainty sets are able to control the occurrences of the worst-case scenarios, finding a compromise between the economy and robustness. However, the authors admitted that the performances of ARO are limited by accuracy of the adjustable parameter. Its improper value results in worse economy of the dispatching.

Recently, several researchers employed distributionally robust optimization (DRO) to establish the optimal economic dispatching model. DRO does not assume that the uncertain variable is subject to any explicit distribution. However, it works by relying on an ambiguity set which contains all the possible probability distributions. And then, DRO optimizes according to the distribution and finds the most optimal solution for the worst distribution. Therefore, constructing a proper ambiguity set has become a critical point. Saberi et al. [20] presented a Wasserstein metric-based ambiguity set for implementing the efficient energy management. The authors claimed the uncertainty issue can be fully addressed based on their DRO model. However, it should be noticed that Wasserstein distance-based ambiguity set always encounters the difficulty of determining the proper radius of Wasserstein sphere. It is also reported that the ambiguity set is extremely sensitive to the outliers, which results in over-conservativeness in the optimization. Zou et al. [21] presented a DRO based electrical infrastructure resilience enhancement strategy. In their research, a moment-based ambiguity set is adopted and its effectiveness can be clearly figured out. However, the moment-based ambiguity set only constrains very limited features of the data samples. Therefore, it may contain quite a number of distributions which are extremely different from the real distribution of the data samples. Consequently, the optimization frequently encounters the over-conservativeness and computational difficulty issues. Researches [13, 19] summarized that the construction of an ambiguity set is still a challenging problem to be solved in future. Additionally, they also suggested that the accuracy of depicting the uncertainty using the ambiguity set is highly related to the number of samples in the set. It should be noted that due to the difficulties of data collection, transmission, and storage, wind power data is naturally scarce, which prevents the original wind power data samples from constructing the accurate ambiguity set. Although there are several data synthetization algorithms for example oversampling [22], undersampling [23], and interpolation [24], they lack the ability to precisely simulate the data distribution of the original data samples, which exhibit constrained performances.

Motivated by aforementioned researches, this paper presents a DRO based optimal economic dispatching strategy for the distribution network. To improve the accuracy of the ambiguity set, this paper first expands the volume of the scarce original wind power generation output samples. SMOTE [24] is employed to implement the first-step data augmentation, in which a proper number of samples are synthesized. And then, VAE [25] is further adopted to implement the second-step data augmentation, in which adequate data samples with similar distribution to the original samples are generated. Secondly, DECA deeply characterizes and clusters the augmented wind power generation output data samples to figure out the typical scenarios. Afterwards, for each typical scenario, an IDDM based ambiguity set is constructed to accurately depict the uncertainty of wind power generation output. And then, the details including the objective functions and constraints of the economic dispatching for distribution network are analyzed and modeled. Finally, a two-stage DRO model considering uncertainties of both wind power generation

output and load is established and solved for figuring out the optimal dispatching strategy. The experimental results show the effectiveness of the presented approach.

2 GENERATION OF IDDM BASED AMBIGUITY SETS USING DATA AUGMENTATION AND DECA

Wind power generation output data samples are usually scarce, which prevents them from being utilized in depth. Therefore, this section firstly employs SMOTE to implement the first-step data augmentation which synthesizes a proper number of samples to expand the dataset. However, it should be realized that SMOTE frequently encounters issues for example overfitting and difficulty in processing high-dimensional data samples. As a result, VAE is further employed to implement the second-step data augmentation, which generates more data samples. Benefitting from the learning capability of VAE, the distribution of the generated samples has great potential to be close to that of the original data samples.

2.1 SMOTE and VAE Based Data Generation

Traditionally, SMOTE is frequently adopted to synthesize samples of minority classes. In this section, as the number of wind power generation output samples is usually small, SMOTE is employed to synthesize the samples based on the entire dataset. The steps of the synthesization are listed below.

Step 1: Let $\mathbf{H} = \{\mathbf{h_i} \,|i = 1, 2, \ldots, q\}$ denote the original wind power generation output data samples; $\mathbf{h_i}$ denote one individual data sample. Firstly, the number of the synthesized data samples N_{SMOTE} is determined in advance.

Step 2: Randomly select $\mathbf{h_i}$ from $\mathbf{H}$ based on the uniform distribution. And then, according to Euclidean distance, for each $\mathbf{h_i}$, computing a number of n_i neighboring data samples $\{\mathbf{h_j} \,|j = 1, 2, \ldots, n_i\}$, $\sum_{i=1}^{q} n_i = N_{\text{SMOTE}}$.

Step 3: For $\mathbf{h_i}$ and each $\mathbf{h_j}$, generate a random value $\lambda \sim \text{Uniform}(0, 1)$. Therefore, the new synthesized data sample can be computed using Equation (1):

$$\mathbf{h_{new_SMOTE}} = \mathbf{h_i} + \lambda \times (\mathbf{h_j} - \mathbf{h_i}). \tag{1}$$

Step 4: Repeat Steps 1 to 3 until N_{syn} synthesized data samples are generated. SMOTE synthesis terminates.

However, SMOTE is characterized by a propensity for overfitting and limited effectiveness in handling high-dimensional data. Therefore, a limited number of N_{syn} data samples are first synthesized by SMOTE in the first-step data augmentation. In terms of generating more data samples which can restore the original data distribution, VAE is further employed for implementing the second-step data augmentation.

VAE is a generative model which learns the probability distribution of the input data samples. Different from traditional autoencoders, it does not output a fixed encoding. Instead, it maps the input data samples to a certain distribution. Afterwards, it obtains the encoding by sampling from the distribution and then reconstructs the data. It can learn the underlying statistical features of the data samples, which further generates new data samples with a similar distribution to those of the original ones. VAE consists of an encoder and a decoder, which are two neural network components. Therefore, to generate new data samples using VAE, the main steps can be summarized as two phases, including the training phase and the data sample generation phase listed below.

In the training phase:

Step 1: In the forward pass, let $\mathbf{h}$ denote an input data sample; ϕ denote the parameters of the encoder neural network; N denote Normal distribution; $\mathbf{I}$ denote identity matrix. Therefore, the encoder generates an approximate posterior distribution shown by Equation (2).

$$\mathbf{q}_\phi(\mathbf{z}|\mathbf{h}) = N(\mathbf{z}; \mu, \sigma^2\mathbf{I}). \tag{2}$$

Step 2: The encoder implements the reparameterization trick. It samples a random noise $\varepsilon \sim N(\mathbf{0}, \mathbf{I})$. And then, a latent $\mathbf{z}$ encoding can be computed using Equation (3).

$$\mathbf{z} = \boldsymbol{\mu} + \boldsymbol{\sigma} \odot \boldsymbol{\varepsilon}. \tag{3}$$

Step 3: $\mathbf{z}$ is input into the decoder. And then, the decoder outputs a likelihood distribution vector $\mathbf{p}_\theta(\mathbf{h}|\mathbf{z})$, where θ denotes the parameters of the decoder neural network.

Step 4: Let L_1 denote the reconstruction loss; L_2 denote the Kullback-Leibler (KL) divergence loss; J denote the dimensionality of the latent space; μ_j and σ_j denote mean and variance of the j^{th} dimension. Therefore, the total loss $\mathrm{L}_{\mathrm{Total}}$ can be computed using Equation (4).

$$\mathrm{L}_{\mathrm{Total}} = \mathrm{L}_1 + \mathrm{L}_2 = -\mathrm{E}_{\mathbf{q}_\phi(\mathbf{z}|\mathbf{h})}[\log \mathbf{p}_\theta(\mathbf{h}|\mathbf{z})] - \frac{1}{2}\sum_{j=1}^{J}\left(1 + \log\left(\sigma_j^2\right) - \mu_j^2 - \sigma_j^2\right). \tag{4}$$

Step 5: In backpropagation, compute the gradient of $\mathrm{L}_{\mathrm{Total}}$ with respect to the encoder parameters ϕ and the decoder parameters θ. And then, the parameters can be updated using Equations (5) and (6), in which ℓ denotes the learning rate.

$$\phi = \phi - \ell \nabla_\phi \mathrm{L}, \tag{5}$$

$$\theta = \theta - \ell \nabla_\theta \mathrm{L}. \tag{6}$$

Step 6: Repeat Step 1 to Step 5, until the training phase terminates.

In the data sample generation phase:

Step 1: Sample a random encoding $\mathbf{z}_{\text{syn}}$ from the prior distribution $\mathbf{p}(\mathbf{z}) = N(\mathbf{0}, \mathbf{I})$.

Step 2: Input $\mathbf{z}_{\text{syn}}$ to the decoder. And then the decoder outputs the likelihood distribution parameter $\mathbf{p}$.

Step 3: A new generated data sample $\mathbf{h}_{\text{new_VAE}}$ can be achieved by sampling from $\mathbf{p}_\theta(\mathbf{h}|\mathbf{z}_{\text{syn}})$.

Step 4: Repeat Step 1 to Step 3, until enough data samples are generated. The generation phase terminates.

Finally, let $\mathbf{H}_{\text{SMOTE}}$ denote the set of SMOTE based synthesized wind power generation output data samples; $\mathbf{H}_{\text{VAE}}$ denote the set of VAE based generated wind power generation output data samples. The set of the augmented wind power generation output data samples $\mathbf{G}$ can be denoted by Equation (7).

$$\mathbf{G} = \mathbf{H} \cup \mathbf{H}_{\text{SMOTE}} \cup \mathbf{H}_{\text{VAE}}. \tag{7}$$

The set $\mathbf{G}$ can be regarded as the original input samples for the next section.

2.2 DECA Based Data Clustering in Enabling Typical Scenario Identification

In order to identify the typical scenarios of the wind power generation output, clustering or classification algorithms are commonly used. However, the conventional algorithms frequently encounter inaccuracy issues. Therefore, this paper adopts DECA to implement the clustering of data samples, which can achieve several typical wind power generation output scenarios represented by different clusters. Basically, DECA implements data dimensionality reduction on the original data samples. And then, it partitions the samples in a lower dimension feature space. Therefore, it improves the clustering accuracy while it also maintains the efficiency of the clustering algorithm [26]. DECA mainly consists of two components including the stacked autoencoder (SAE) and k-means based clustering. SAE comprises multiple autoencoders, each of which contains an encoder and a decoder. The encoder reduces the dimension of the original input samples $\mathbf{G} = \{\mathbf{g_i}|i = 1, 2, \ldots, n\}$ to the lower dimension samples $\mathbf{D} = \{\mathbf{d_i}|i = 1, 2, \ldots, n\}$, where n is the number of the samples. Contrarily, the decoder can restore each dimensionally reduced sample $\mathbf{d_i}$ to $\widehat{\mathbf{g}}_\mathbf{i}$ which should be as close as possible to the original input sample $\mathbf{g_i}$. In terms of measuring the error between $\widehat{\mathbf{g}}_\mathbf{i}$, this section employs mean squared error (MSE) as the loss function represented by Equation (8).

$$\text{Loss}_{\text{MSE}} = \sum_{i=1}^{n} \left\| \mathbf{g_i} - \widehat{\mathbf{g}}_\mathbf{i} \right\|^2. \tag{8}$$

To minimize the loss function, Adam back propagation algorithm [26] is also employed to iteratively implement the adjustments of the weights and biases in SAE.

The adjustment in each iteration can be represented by Equation (9):

$$\mathbb{T}_{\mathbb{N}} = \nabla_{o_{\mathbb{N}}} \mathrm{Loss}_{\mathrm{MSE}}(\mathbf{o}_{\mathbb{N}}), \mathbf{o}_{\mathbb{N}} = [\mathbb{Q}_{\mathbb{N}}, \mathbf{p}_{\mathbb{N}}], \tag{9}$$

where $\mathbb{Q}$ is the weight matrix; $\mathbf{p}$ is the bias vector; $\mathbb{N}$ is the round of the iterations. Adam implements the adjustments based on mean value $J_{\mathbb{N}}$ and variance value $F_{\mathbb{N}}$ of gradients. The adjustment process is shown by Equation (10):

$$\begin{cases} \mathbb{Q}_{\mathrm{N}} = B_1 J_{\mathrm{N}-1} + (1 - B_1)\mathbb{T}_{\mathrm{N}}, \\ p_{\mathbb{N}} = B_2 F_{\mathrm{N}-1} + (1 - B_2)\mathbb{T}_{\mathrm{N}}^2, \end{cases} \tag{10}$$

where B_1 and B_2 are the adjustment parameters. The iteratively adjustments of SAE parameters are shown by Equation (11):

$$\mathbf{o}_{\mathbb{N}+1} = \mathbf{o}_{\mathbb{N}} - \frac{X}{\sqrt{\hat{p}_{\mathbb{N}}} + \sigma}\hat{\mathbb{Q}}_{\mathbb{N}}, \quad \hat{\mathbb{Q}}_{\mathbb{N}} = \frac{\mathbb{Q}_{\mathbb{N}}}{1 - B_1}, \quad \hat{p}_{\mathbb{N}} = \frac{p_{\mathbb{N}}}{1 - B_2}. \tag{11}$$

where σ is a small constant number which ensures the denominator is not zero; X is the learning rate. The trained SAE is able to implement the dimensionality reduction for the original input samples to generate the dimensionally-reduced samples $\mathbf{D} = \{\mathbf{d_i}|i = 1, 2, \ldots, n\}$.

Further, DECA employs k-means algorithm to cluster $\mathbf{D} = \{\mathbf{d_i}|i = 1, 2, \ldots, n\}$. Therefore, based on the clustering result, the reference points for subsequent probability assignments can be provided. Let $\mathbf{Cluster} = \{\mathbf{cluster_j}|j = 1, 2, \ldots, k\}$ denote the result, a highly brief introduction of k-means algorithm is shown in Table 1.

Pseudo code of k-means clustering algorithm
Input: $\mathbf{D} = \{\mathbf{d_i}|i = 1, 2, \ldots, n\}$
Output: cluster result $\mathbf{Cluster} = \{\mathbf{cluster_j}|j = 1, 2, \ldots, k\}$
1. A number of k initial centroids $\mathbf{c_j}, j = 1, 2, \ldots, k$ are randomly selected from $\mathbf{D}$.
2. Compute the Euclidean distance $\mathrm{dis}(\mathbf{d_i}, \mathbf{c_j})$ between $\mathbf{d_i}$ and $\mathbf{c_j}$ using $\mathrm{dis}(\mathbf{d_i}, \mathbf{c_j}) = \|\mathbf{d_i} - \mathbf{c_j}\|_2$
3. For each $\mathbf{d_i}$, its belonged cluster can be determined according to $\mathbf{cluster} = \arg\min_{j=1}^{k} \|\mathbf{d_i} - \mathbf{c_j}\|_2$
4. Update the centroids according to $\mathbf{c_j^{new}} = \frac{\sum_{\mathbf{d_i} \in \mathbf{cluster_j}} \mathbf{d_i}}{n_j}$ where n_j is the number of samples in the j^{th} cluster.
5. Repeat Steps 2 to 4, until the centroids are not changed.
6. Algorithm terminates.

Table 1. The pseudo code of k-means clustering algorithm

Additionally, to determine the cluster number k of k-means algorithm, this paper employs Davison-Bouldin Index (DBI) [25] and Silhouette Coefficient (SC) [25]. DBI is able to identify the tightness of one sample to the other samples within the same cluster and the separation of the sample among the other samples in different

clusters. Smaller DBI presents tighter interior of the cluster and better separation among clusters. SC is able to identify the closeness of a sample to its own cluster and its separation from the other clusters. Larger SC indicates a relatively closer relationship between the sample and cluster. According to Equation (12), the initial centroid number k can be finally determined.

$$k = \arg \max_{k_i \in [k_{\min}, k_{\max}]} [\mathrm{SC}(k_i)] \cap \arg \max_{k_i \in [k_{\min}, k_{\max}]} [\mathrm{DBI}(k_i)]. \tag{12}$$

Based on the determined cluster number k and k-means algorithm, a number of k clusters and their centroids $\mathbf{c_j}, j = 1, 2, \ldots, k$ can be finally achieved. And then, DECA employs t-distribution which is shown by Equation (13) to define the probability of each sample $\mathbf{d_i}$ belonged to the centroid $\mathbf{c_j}$.

$$Q_i^j = \frac{\left(\frac{\left\| \mathbf{d_i} - \mathbf{c_j} \right\|_2}{\alpha} + 1 \right)^{-\frac{1+\alpha}{2}}}{\sum_{j=1}^{k} \left(\frac{\left\| \mathbf{d_i} - \mathbf{c_j} \right\|_2}{\alpha} + 1 \right)^{-\frac{1+\alpha}{2}}}, \tag{13}$$

where Q_i^j is the initial probability distribution of $\mathbf{d_i}$ belonged to $\mathbf{c_j}$; α is the degrees of freedom parameter of which the value is frequently set as 1. It also suggests that to define a target auxiliary distribution P_i^j can further improve the parameter adjustments for DECA. Therefore, in this paper, P_i^j is defined according to Equation (14):

$$P_i^j = \frac{\left(Q_i^j\right)^2 / F_j}{\sum_{j=1}^{k} \left(Q_i^j\right)^2 / F_j}, F_j = \sum_{i=1}^{n} Q_i^j, \tag{14}$$

where F_j is the soft assignment sum for the j^{th} cluster. And then, KL divergence is adopted to minimize the error between Q_i^j and P_i^j according to Equation (15):

$$\mathrm{LOSS}_{\mathrm{KL}} = \mathrm{KL}(P_i^j \left\| Q_i^j \right.) = \sum_{i=1}^{n} \sum_{j=1}^{k} P_i^j \log \frac{P_i^j}{Q_i^j}. \tag{15}$$

In the iterations, based on Adam back propagation, the encoder weights and cluster centroids are alternately optimized to minimize $\mathrm{LOSS}_{\mathrm{MSE}}$ and $\mathrm{LOSS}_{\mathrm{KL}}$ until DECA terminates. Finally, according to the optimized centroids, the optimal clusters $\mathbf{Cluster} = \{\mathbf{cluster_j} | j = 1, 2, \ldots, k\}$ can be achieved. In each cluster, the decoder of SAE is able to restore each dimensionally reduced samples to the original wind power generation output data samples. Therefore, the sample-restored clusters $\mathbf{S} = \{\mathbf{S_j} | j = 1, 2, \ldots, k\}$ each of which depicts one typical scenario of the wind power generation output can be obtained.

2.3 Construction of IDDM Based Ambiguity Set

To precisely characterize the uncertainty of wind power generation output, this section constructs the IDDM based ambiguity set for each cluster $\mathbf{S_j}$ in $\mathbf{S}$. Let x_κ denote the number of wind power generation data samples in a certain cluster $\mathbf{S_j}$. Therefore, there are a number of x_κ possible states of the wind power output variables $u^{\zeta}_{\mathrm{WPG}} = [u^1_{\mathrm{WPG}}, u^2_{\mathrm{WPG}}, \cdots, u^{x_\kappa}_{\mathrm{WPG}}]$, $\zeta = 1, \ldots, x_\kappa$ in $\mathbf{S_j}$. Let ϖ^ζ denote the occurrence probability of each state. Therefore, the posterior Dirichlet density function based on the Bayesian theory can be derived using Equation (16):

$$\begin{cases} \chi(\varpi\,|\mathcal{T}) = \frac{\Gamma(1+\mathcal{T})\cdot\prod_{\zeta=1}^{x_\kappa}(\varpi^\zeta)^{r^\zeta+\mathcal{T}^\zeta-1}}{\prod_{\zeta=1}^{x_\kappa}\Gamma(r^\zeta+\mathcal{T}^\zeta)}, \\ \sum_{\zeta=1}^{x_\kappa} r^\zeta = 1, \forall r^\zeta \in [0,1], \end{cases} \tag{16}$$

where $\Gamma(\bullet)$ is the Gamma function; r^ζ is the ζ^{th} weight factor; $\mathcal{T}^\zeta$ is the number of occurrences for the ζ^{th} states; $\mathcal{T} = \mathcal{T}^1 + \mathcal{T}^2 + \cdots + \mathcal{T}^{x_\kappa}$ is the observations of each states. Therefore, based on r^ζ, the expected value of the posterior distribution can be obtained using Equation (17):

$$E(\varpi^\zeta) = (r^\zeta + \mathcal{T}^\zeta)(\mathcal{T}+1)^{-1}. \tag{17}$$

Moreover, according to Equation (17), uncertainty intervals of probabilities of occurrences for states can be computed using Equation (18):

$$\varpi^\zeta = \left[\underline{E}(\varpi^\zeta), \overline{E}(\varpi^\zeta)\right] = \left[\mathcal{T}^\zeta/\mathcal{T}+1, \mathcal{T}^\zeta+1/\mathcal{T}+1\right], \tag{18}$$

where $\overline{E}(\varpi^\zeta)$ and $\underline{E}(\varpi^\zeta)$ are the upper bound and lower bound of the expected values of the posterior distribution when $r^\zeta = 0$ and $r^\zeta = 1$. Therefore, the confidence interval $I = [\underline{\varpi^\zeta}, \overline{\varpi^\zeta}]$ based on the confidence coefficient γ can be constructed using Equation (19):

$$I = \begin{cases} \underline{\varpi^\zeta} = 0, \quad \overline{\varpi^\zeta} = \beta_2^{-1}\left(\frac{1+\gamma}{2}\right), \quad \mathcal{T}^\zeta = 0, \\ \underline{\varpi^\zeta} = \beta_1^{-1}\left(\frac{1-\gamma}{2}\right), \quad \overline{\varpi^\zeta} = 1, \quad \mathcal{T}^\zeta = x_\kappa, \\ \underline{\varpi^\zeta} = \beta_1^{-1}\left(\frac{1-\gamma}{2}\right), \quad \overline{\varpi^\zeta} = \beta_2^{-1}\left(\frac{1+\gamma}{2}\right), \quad 0 < \mathcal{T}^\zeta < x_\kappa, \end{cases} \tag{19}$$

where β_1 and β_2 are the cumulative distribution function (CDF) of the probability density function (PDF) of Beta distribution $B(m^\zeta, 1 + x_\kappa - \mathcal{T}^\zeta)$ and $B(m^\zeta + 1, x_\kappa - \mathcal{T}^\zeta)$. It should be aware that the confidence interval $[\underline{\varpi^\zeta}, \overline{\varpi^\zeta}]$ may not cover the real estimated interval of the wind power generation output variables u_{WPG} of which its real estimated interval $[u^l_{\mathrm{WT}}, u^u_{\mathrm{WT}}]$ should be measured using Equation (20):

$$\begin{aligned} u_{\mathrm{WT}}^{l} &= u_{\mathrm{WT}}^{1} - \frac{1-\gamma}{2}, \\ u_{\mathrm{WT}}^{u} &= u_{\mathrm{WT}}^{x_\kappa} + \frac{1-\gamma}{2}. \end{aligned} \tag{20}$$

Based on both confidence interval $[\underline{\varpi^{\zeta}}, \overline{\varpi^{\zeta}}]$ and real estimated interval $[u_{\mathrm{WT}}^{l}, u_{\mathrm{WT}}^{u}]$, the ambiguity set can be denoted by Equation (21):

$$\mathrm{U} = \left\{ P\left(u_{\mathrm{WT}} \leqslant u_{\mathrm{WT}}^{\zeta}\right) \in \left[\underline{\varpi^{\zeta}}, \overline{\varpi^{\zeta}}\right] \left| P \in \mathrm{U}\left(\left|u_{\mathrm{WT}}^{l}, u_{\mathrm{WT}}^{u}\right|\right) \right\}, \tag{21}$$

where U represents the ambiguity set of wind power output variables.

Finally, based on Equation (21) and research [27] the lower bound $\underline{u}_{\mathrm{WT}}$ and upper bound $\bar{u}_{\mathrm{WT}}$ of U can be determined using Equation (22):

$$[\underline{u}_{\mathrm{WT}}, \bar{u}_{\mathrm{WT}}] = [(1-\gamma)/2, (1+\gamma)/2]. \tag{22}$$

3 TWO-STAGE DRO MODEL FOR DISTRIBUTION NETWORK

In this section, this paper presents a two-stage DRO model to enable optimal economic dispatching for a distribution network. A typical distribution network with distributed power generation, including wind power and gas turbines, is shown in Figure 1.

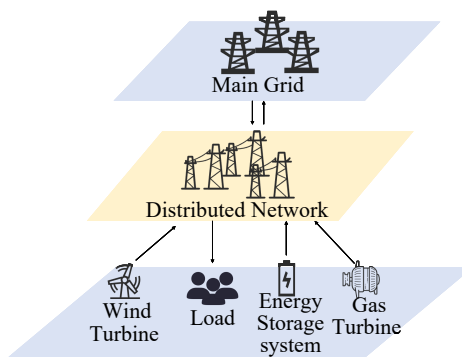


Figure 1. A typical distribution network with distributed power generations

Briefly, a modern distribution network contains distributed power generation, energy storage, and load. It also connects to the bulk power system. Therefore, according to the conditions of its own distributed power generation, the distribution network can buy or sell power energy from or to the main grid. As a result, the uncertainties of both wind power generation output and load significantly impact

the power generation and consumption; moreover, the economics of dispatching for the distribution network.

3.1 The Construction of Uncertainty Set and Ambiguity Set for Load and Wind Power Generation Output

In terms of presenting uncertainty of load in the distribution network, this paper employs the box-type uncertainty set $[\underline{u}_{\mathrm{Lo}}, \overline{u}_{\mathrm{Lo}}]$. Additionally, the adjustable robust parameters Γ_{Lo} and Γ_{WT} are adopted to adjust the conservativeness of the uncertainty set and the ambiguity set. The ambiguity set of wind power generation output is constructed using the presented IDDM approach. Finally, the uncertainty set and the ambiguity set of load and wind power generation output can be indicated using Equation (23):

$$\begin{cases} \sum_{t=1}^{T} \left(\dfrac{P_{\mathrm{Lo}}(t) - P_{\mathrm{Lo}}^{\mathrm{pre}}(t)}{\underline{u}_{\mathrm{Lo}}} \nu_{\mathrm{Lo}}^{+} + \dfrac{P_{\mathrm{Lo}}^{\mathrm{pre}}(t) - P_{\mathrm{Lo}}(t)}{\overline{u}_{\mathrm{Lo}}} \nu_{\mathrm{Lo}}^{-} \right) \leq \Gamma_{\mathrm{Lo}}, \\ \sum_{t=1}^{T} \left(\dfrac{P_{\mathrm{WT}}(t) - P_{\mathrm{WT}}^{\mathrm{pre}}(t)}{\underline{u}_{\mathrm{WT}}} \nu_{\mathrm{WT}}^{+} + \dfrac{P_{\mathrm{WT}}^{\mathrm{pre}}(t) - P_{\mathrm{WT}}(t)}{\overline{u}_{\mathrm{WT}}} \nu_{\mathrm{WT}}^{-} \right) \leq \Gamma_{\mathrm{WT}}, \end{cases} \tag{23}$$

where $\underline{u}_{\mathrm{WT}}$ is the lower bound of load uncertainty set; $\bar{u}_{\mathrm{Lo}}$ is the upper bound of load uncertainty set; v_{Lo}^{+} and v_{Lo}^{-} are auxiliary variables of which the value are 0 or 1; $P_{\mathrm{Lo}}^{\mathrm{pre}}(t)$ is the predicted value of load at time t; $P_{\mathrm{WT}}^{\mathrm{pre}}(t)$ is the predicted value of wind power generation output at time t; v_{WT}^{+} and v_{WT}^{-} are auxiliary variables of which the value are 0 or 1; T is 24 which denotes the dispatching horizon.

3.2 Objective Function and Constraints of Economic Dispatching Model

The objective function of the economic dispatching model is to minimize the total dispatching cost with considering the uncertainties of wind power generation output and load. It contains the costs of two stages including day-ahead stage and real-time stage, which is shown by Equation (24):

$$\min C_{\mathrm{all}} = C_{\mathrm{da}}(x) + C_{\mathrm{rt}}(u, y), \tag{24}$$

where C_{all} is the total dispatching cost which should be minimized; x is the decision variable in the first stage; u is the uncertain variable; y is the decision variable in the second stage.

In Equation (24), C_{da} is the day-head dispatching cost of which the details are presented by Equation (25):

$$\begin{cases} C_{\mathrm{da}} = C_{\mathrm{Grid}}^{\mathrm{da}} + C_{\mathrm{MT}} + C_{\mathrm{loss}}, \\ C_{\mathrm{Grid}}^{\mathrm{da}} = \sum_{t}^{\mathrm{T}} \lambda_{\mathrm{Grid}}^{\mathrm{da}} P_{\mathrm{Grid}}^{\mathrm{da}}(t) \Delta t, \\ C_{\mathrm{MT}} = \sum_{t}^{\mathrm{T}} \left[a_{\mathrm{MT}} P_{\mathrm{MT}}(t) + b_{\mathrm{MT}} \right] \Delta t, \\ C_{\mathrm{Loss}} = \sum_{t=1}^{\mathrm{T}} \sum_{i,j=1}^{\mathrm{N}} a_{\mathrm{Loss}} I_{ij}^{2}(t) R_{ij}, \end{cases} \tag{25}$$

where $C_{\mathrm{Grid}}^{\mathrm{da}}$ is the day-ahead interaction cost with bulk power system; $\lambda_{\mathrm{Grid}}^{\mathrm{da}}$ is the day-ahead purchasing or selling price of power; $P_{\mathrm{Grid}}^{\mathrm{da}}(t)$ is the amount of purchased or sold power in the day-ahead stage; C_{MT} is the cost of gas turbine output; a_{MT} and b_{MT} are the cost coefficients of power generation; $P_{\mathrm{MT}}(t)$ is the gas turbine output at time t; a_{Loss} denotes the cost of network loss; $I_{ij}(t)$ is the current of brunch ij at time t; R_{ij} is the resistance of brunch ij.

In Equation (24), C_{rt} is the real-time dispatching cost of which the details are presented by Equation (26):

$$\begin{cases} C_{\mathrm{rt}} = C_{\mathrm{MT}}^{\mathrm{up}} + C_{\mathrm{MT}}^{\mathrm{down}} + C_{\mathrm{WPG}} + C_{\mathrm{Load}} + C_{\mathrm{Grid}}^{\mathrm{rt}}, \\ C_{\mathrm{MT}}^{\mathrm{up}} + C_{\mathrm{MT}}^{\mathrm{down}} = \sum_{t=1}^{\mathrm{T}} \left[\lambda_{\mathrm{MT}}^{\mathrm{up}} P_{\mathrm{MT}}^{\mathrm{up}}(t) + \lambda_{\mathrm{MT}}^{\mathrm{down}} P_{\mathrm{MT}}^{\mathrm{down}}(t) \right] \Delta t, \\ C_{\mathrm{WPG}} = \sum_{t=1}^{\mathrm{T}} \lambda_{\mathrm{WPG}} \left[P_{\mathrm{WT}}^{\mathrm{pre}}(t) - P_{\mathrm{WPG}}^{\mathrm{rt}}(t) \right] \Delta t \\ C_{\mathrm{Load}} = \sum_{t=1}^{\mathrm{T}} \lambda_{\mathrm{Load}} \left[P_{\mathrm{Lo}}^{\mathrm{pre}}(t) - P_{\mathrm{Load}}^{\mathrm{rt}}(t) \right] \Delta t, \\ C_{\mathrm{Grid}}^{\mathrm{rt}} = \sum_{t=1}^{\mathrm{T}} \lambda_{\mathrm{Grid}}^{\mathrm{rt}} P_{\mathrm{Grid}}^{\mathrm{rt}}(t) \Delta t, \end{cases} \tag{26}$$

where $C_{\mathrm{MT}}^{\mathrm{up}}$ and $C_{\mathrm{MT}}^{\mathrm{down}}$ are the real-time ramp-up and down costs of the gas turbine; C_{WPG} is the cost of wind curtailment; C_{Load} is the cost of load shedding; $C_{\mathrm{Grid}}^{\mathrm{rt}}$ is the cost of the power interaction with bulk power system; $\lambda_{\mathrm{MT}}^{\mathrm{up}}$ and $P_{\mathrm{MT}}^{\mathrm{up}}(t)$ are the unit ramp-up cost and ramp-up power amount of the gas turbine; $\lambda_{\mathrm{MT}}^{\mathrm{down}}$ and $P_{\mathrm{MT}}^{\mathrm{down}}(t)$ are the unit ramp-down cost and ramp-down power amount of the gas turbine; λ_{WPG} is the unit wind curtailment cost; $P_{\mathrm{WPG}}^{\mathrm{rt}}(t)$ is the real wind power generation output; λ_{Load} is the unit load shedding cost; $P_{\mathrm{Load}}^{\mathrm{rt}}(t)$ is the real load; $\lambda_{\mathrm{Grid}}^{\mathrm{rt}}$ is the real-time purchasing or selling price of power; $P_{\mathrm{Grid}}^{\mathrm{rt}}(t)$ is the amount of purchased or sold power in the real-time stage.

Moreover, the constraints related to the economic dispatching model are listed below.

3.2.1 Power Flow Constraints

The power flow constraints are indicated by Equations (26), (27), (28), (29), (30), (31) and (32):

$$P_j(t) = P^j_{\text{Load}}(t) - P^j_{\text{Grid}}(t) - P^j_{\text{MT},w}(t),\ \forall j \in N, \tag{27}$$

$$Q_j(t) = Q^j_{\text{Load}}(t) - Q^j_{\text{Grid}}(t),\ \forall j \in N, \tag{28}$$

$$P_j(t) = \sum_{i\in\phi(j)} P_{ij}(t) - \sum_{k\in\theta(j)} P_{jk}(t),\ \forall j \in N, \tag{29}$$

$$Q_j(t) = \sum_{i\in\phi(j)} Q_{ij}(t) - \sum_{k\in\theta(j)} Q_{jk}(t),\ \forall j \in N, \tag{30}$$

$$V_i^2(t) - V_j^2(t) = -2\left[P_{ij}(t)R_{ij} + Q_{ij}(t)X_{ij}\right] + I_{ij}^2(t)(R_{ij}^2 + X_{ij}^2),\ \forall ij \in N, \tag{31}$$

$$I_{ij}^2(t)V_j^2(t) = P_{ij}^2(t) + Q_{ij}^2(t),\ \forall ij \in N, \tag{32}$$

where N denotes number of nodes of the distribution network; $\phi(j)$ and $\theta(j)$ are the set of source nodes and set of terminal nodes of branches connected to node j respectively; $P_j(t)$ and $Q_j(t)$ are the active power and reactive power injected at node j at time t respectively; $P_{ij}(t)$ and $Q_{ij}(t)$ represent the active power and reactive power of brunch ij at time t respectively; $P_{jk}(t)$ and $Q_{jk}(t)$ represent the active power and reactive power of brunch jk at time t respectively; $P^j_{\text{MT},w}(t)$ is the active power injected at node j at time t by the w^{th} gas turbine; $V_i(t)$ represents the voltage of node i at time t; $V_j(t)$ represents the voltage of node j at time t; $I_{ij}(t)$ represents the current of brunch ij at time t; $P_{\text{Grid},j}(t)$ and $Q_{\text{Grid},j}(t)$ represent the active power and reactive power injected by the transformer at time t respectively; $P_{\text{Load},j}(t)$ and $Q_{\text{Load},j}(t)$ represent the active power load and reactive power load at node j at time t respectively; R_{ij} and X_{ij} are the resistance and reactance of brunch ij respectively. Because of the non-linear features in Equation (32), this paper adopts second-order cone relaxation to transform the equation to linear constraint shown by Equation (33).

$$\left\| \begin{array}{c} 2P_{ij}(t) \\ 2Q_{ij}(t) \\ I_{ij}^2(t) - V_j^2(t) \end{array} \right\|_2 \le V_j^2(t) + I_{ij}^2(t). \tag{33}$$

The branch current and voltage limit constraints of the distribution network are shown by Equations (34) and (35):

$$0 \le V_j^2(t) \le V_{\max}^2, \tag{34}$$

$$0 \le I_{ij}^2(t) \le I_{\max}^2, \tag{35}$$

where $I_{\max}$ is the upper limit of the brunch current; $V_{\max}$ is upper limit of the node voltage.

3.2.2 Gas Turbine Output Constraints

In the day-ahead stage, the gas turbine output constraints are shown by Equation (36):

$$\begin{cases} P_{\mathrm{MT},w}^{\min} \le P_{\mathrm{MT},w}(t) \le P_{\mathrm{MT},w}^{\max}, \\ P_{\mathrm{MT},w}(t-1) - P_{\mathrm{MT},w}(t) \le r_{\mathrm{MTd}} \cdot \Delta t, \\ P_{\mathrm{MT},w}(t) - P_{\mathrm{MT},w}(t-1) \le r_{\mathrm{MTu}} \cdot \Delta t, \end{cases} \tag{36}$$

where $P_{\mathrm{MT},w}^{\max}$ is the upper limit of the w^{th} gas turbine output; $P_{\mathrm{MT},w}^{\min}$ is the lower limit of the w^{th} gas turbine output; Δt is the unit time interval; r_{MTd} is the ramp-down rate of the gas turbine; r_{MTu} is the ramp-up rate of the gas turbine.

In the real-time stage, the gas turbine output constraints are shown by Equation (37):

$$P_{\mathrm{MT},w}^{\min} \le P_{\mathrm{MT},w}(t) + P_{\mathrm{MT},w}^{\mathrm{up}}(t) - P_{\mathrm{MT},w}^{\mathrm{down}}(t) \le P_{\mathrm{MT},w}^{\max}. \tag{37}$$

3.2.3 Energy Storage Constraints

In the day-ahead stage, the energy storage constraints are shown by Equation (38):

$$\begin{cases} S_{\mathrm{ESS}}^{\mathrm{Cha}}(t) + S_{\mathrm{ESS}}^{\mathrm{Dis}}(t) \le 1, \\ S_{\mathrm{ESS}}^{\mathrm{Cha}}(t) P_{\mathrm{Cha}}^{\min} \le P_{\mathrm{Cha}}(t) \le S_{\mathrm{ESS}}^{\mathrm{Cha}}(t) P_{\mathrm{Cha}}^{\max}, \\ S_{\mathrm{ESS}}^{\mathrm{Dis}}(t) P_{\mathrm{Dis}}^{\min} \le P_{\mathrm{Cha}}(t) \le S_{\mathrm{ESS}}^{\mathrm{Dis}}(t) P_{\mathrm{Dis}}^{\max}, \\ E_{\mathrm{ESS}}(t) = (1-\delta_{\mathrm{ESS}}) E_{\mathrm{ESS}}(t-1) + \left[\eta_{\mathrm{ESS}}^{\mathrm{Cha}} P_{\mathrm{Cha}}(t-1) - P_{\mathrm{Dis}}(t-1)/\eta_{\mathrm{ESS}}^{\mathrm{Dis}}\right], \\ E_{\mathrm{ESS}}^{\min} \le E_{\mathrm{ESS}}(t) \le E_{\mathrm{ESS}}^{\max}, \end{cases} \tag{38}$$

where $S_{\mathrm{ESS}}^{\mathrm{Cha}}(t)$ is a 0-1 variable which denotes the state of energy storage charging; $S_{\mathrm{ESS}}^{\mathrm{Dis}}(t)$ is a 0-1 variable which denotes the state of energy storage discharging; $P_{\mathrm{Cha}}(t)$ denotes the current state of energy storage; $P_{\mathrm{Cha}}^{\max}$ is the energy storage maximum charging limit; $P_{\mathrm{Cha}}^{\min}$ is the energy storage minimum charging limit; $P_{\mathrm{Dis}}^{\max}$ is the energy storage maximum discharging limit; $P_{\mathrm{Dis}}^{\min}$ is the energy storage minimum discharging limit; $E_{\mathrm{ESS}}(t)$ represents the state of the energy storage at time t; δ_{ESS} is the self-loss power of energy storage; $\eta_{\mathrm{ESS}}^{\mathrm{Cha}}$ is the charging efficiency of the energy storage; $\eta_{\mathrm{ESS}}^{\mathrm{Dis}}$ is the discharging efficiency of the energy storage; $E_{\mathrm{ESS}}^{\max}$ denotes the energy storage upper limit; $E_{\mathrm{ESS}}^{\min}$ denotes the energy storage lower limit. In the

real-time stage, the energy storage constraints are shown by Equation (39).

$$\begin{cases} P_{\text{Cha}}^{\min} \le P_{\text{Cha}}(t) + P_{\text{Cha}}^{\text{up}}(t) - P_{\text{Cha}}^{\text{down}}(t) \le P_{\text{Cha}}^{\max}, \\ P_{\text{Dis}}^{\min} \le P_{\text{Dis}}(t) + P_{\text{Dis}}^{\text{up}}(t) - P_{\text{Dis}}^{\text{down}}(t) \le P_{\text{Dis}}^{\max}, \\ E_{\text{ESS}}^{\text{rt}}(t) = (1-\delta_{\text{ESS}})E_{\text{ESS}}^{\text{rt}}(t-1) + \left\{ \eta_{\text{ESS}}^{\text{Cha}} \left[P_{\text{Cha}}(t) + P_{\text{Cha}}^{\text{up}}(t) - P_{\text{Cha}}^{\text{down}}(t) \right] \right. \\ \qquad \left. - \dfrac{P_{\text{Dis}}(t) + P_{\text{Dis}}^{\text{up}}(t) - P_{\text{Dis}}^{\text{down}}(t)}{\eta_{\text{ESS}}^{\text{Dis}}} \right\} \Delta t, \\ E_{\text{ESS}}^{\min} \le E_{\text{ESS}}^{\text{rt}}(t) \le E_{\text{ESS}}^{\max}. \end{cases} \tag{39}$$

3.2.4 Power Purchase and Sale Constraints

The constraints in both the day-ahead stage and the real-time stage are the same, which is indicated by Equation (40):

$$P_{\text{Grid}}^{\min} \leqslant P_{\text{Grid}}(t) \leqslant P_{\text{Grid}}^{\max}, \tag{40}$$

where $P_{\text{Grid}}^{\max}$ and $P_{\text{Grid}}^{\min}$ are the power transmission capacity limit of the distribution network.

3.3 DRO Based Optimization Model and Its Solution of Economic Dispatching Model

The objective function and constraints presented in Section 3.2 can be formulated to a two-stage DRO model shown by Equation (41):

$$\begin{cases} \min\limits_{x} \left[\mathbb{C}^{\text{T}} x + \max\limits_{u} \min\limits_{y} \left(\mathbb{D}^{\text{T}} y + \mathbb{E}^{\text{T}} u \right) \right] \\ \text{s.t.} \\ \quad \mathbf{A}x = \mathbb{G}, \ \mathbf{B}x \le \mathbb{K}, \\ \quad \mathbf{C}x + \mathbf{D}y = \mathbb{I}, \quad \mathbf{E}x + \mathbf{F}y \le \mathbb{J}, \\ \quad \mathbf{G}y \le u, \end{cases} \tag{41}$$

where $\mathbb{C}, \mathbb{D}, \mathbb{E}, \mathbb{G}, \mathbb{K}, \mathbb{I}, \mathbb{J}, A, B, C, D, E, F, G$ are constant matrices and vectors. The model is a coupled min-max-min structure. The min function located in outer layer is to minimize the total dispatching cost including costs of both day-ahead stage and real-time stage. The max-min function located in inner layer is to find out the worst scenario and the optimal real-time adjusting cost in the scenario. To solve

the two-stage model with coupled inner and outer layers, Big-M method and duality theory [27] are adopted. The approaches are able to decouple the model into master problem and sub problem which can be further directly solved. Let r denote the iteration number, according to research [27], the decoupled master problem can be shown by Equation (42).

$$\begin{cases} \min\limits_{x}(\mathbf{C}^{\mathrm{T}}x+\theta) \\ \text{s.t.} \\ \quad \theta \geq \mathbb{D}^{\mathrm{T}}y+\mathbb{E}^{\mathrm{T}}u, \\ \quad \mathbf{A}x=\mathbb{G},\ \mathbf{B}x\leq\mathbb{K}, \\ \quad \mathbf{C}x+\mathbf{D}y=\mathbb{I}, \\ \quad \mathbf{E}x+\mathbf{F}y\leq\mathbb{J}, \\ \quad \mathbf{G}y\leq u, \end{cases} \quad \text{and} \quad \begin{cases} \theta \geq \mathbb{D}^{\mathrm{T}}y_{t+1}+\mathbb{E}^{\mathrm{T}}u_{t-1}, \\ \mathbf{C}x+\mathbf{D}y_{t+1}=\mathbb{I}, \\ \mathbf{E}x+\mathbf{F}y_{t+1}\leq\mathbb{J}, \\ \mathbf{G}y_{t+1}\leq u_{t+1}, \end{cases} \tag{42}$$

And the decoupled sub problem can be shown by Equation (43):

$$\begin{cases} \max\limits_{u,\alpha,\beta,\tau}\left[u^{\mathrm{up}}\varepsilon^{+}+u^{\mathrm{down}}\varepsilon^{-}+u^{\mathrm{pre}}(1-\varepsilon^{+}-\varepsilon^{-})+\mathbb{I}^{\mathrm{T}}\alpha-x^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\alpha+\mathbb{J}^{\mathrm{T}}\beta-x^{\mathrm{T}}\mathbf{E}^{\mathrm{T}}\beta\right] \\ \text{s.t.} \\ \quad \mathbf{D}^{\mathrm{T}}\alpha+\mathbf{F}^{\mathrm{T}}\beta+\mathbf{G}^{\mathrm{T}}\tau=\mathbb{D}, \\ \quad \beta\leq 0\ \tau\leq 0 \\ \quad \varepsilon=\mathbb{E}+\tau, \\ \quad -M(1-u_t^{+})+\varepsilon_t\leq\varepsilon_t^{+}\leq M(1-u_t^{+})+\varepsilon_t, \\ \quad -M(1-u_t^{-})+\varepsilon_t\leq\varepsilon_t^{-}\leq M(1-u_t^{-})+\varepsilon_t, \\ \quad -Mu_t^{+}\leq\varepsilon_t^{+}\leq Mu_t^{+}, \\ \quad -Mu_t^{-}\leq\varepsilon_t^{-}\leq Mu_t^{-}, \\ \quad u_t^{+}+u_t^{-}\leq 1, \\ \quad \sum\limits_{t}(u_t^{+}+u_t^{-})\leq\Gamma, \end{cases} \tag{43}$$

In Equation (43), α, β, τ, and ε are dual variables; u^{up} is the upper boundary of u; u^{down} is the lower boundary of u; u^{pre} is the predicted value of u; ε^{+} is the positive value of ε; ε^{-} is the negative value of ε; M is a big number; u_t^{-} and u_t^{+} are 0-1 variables; According to column and constraint generation (C&CG) [27], the problems can be solved iteratively.

4 EXPERIMENTAL RESULT

4.1 Parameters Employed in Experiments

The following experiments are carried out using PC with AMD Ryzen 7 5800H, 16 GB memory, Windows 10 operating system, and MATLAB R2020a. The decoupled DRO model is solved using CPLEX. Keras [28] on Python 3.9 is employed to implement DECA. The detailed structure of the distribution network studied by this paper is shown in Figure 2.

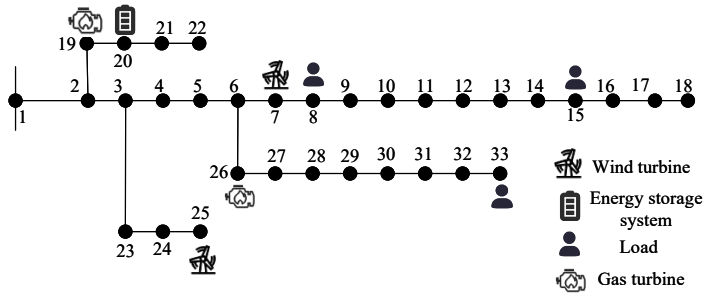


Figure 2. Detailed structure of the studied distribution network

Wind turbines 1 and 2 connect to buses 7 and 25 of the distribution network. The number of the original wind power generation output samples is 2 434 in total. The uncertainty of the wind power generation output is depicted using the presented IDDM based ambiguity set. Gas turbines 1 and 2 are connected to bus 19 and 26; Loads 1, 2, and 3 are connected to bus 8, 15, and 33 of the distribution network. The uncertainty of load is depicted using the box-type uncertainty set. Figure 3 depicts the input scenarios adopted for the proposed economic dispatch model, including probabilistic wind and load forecasts as well as the time-varying electricity prices.

The predicted values of wind power generation outputs and loads for day-head dispatching are shown in Figure 3 a). Both encoder and decoder in VAE are four-layer multilayer perceptron neural networks. The numbers of neurons in each layer of the encoder are 24, 512, 256, and 8 respectively. The activation function of all neurons of the encoder is ReLU. The numbers of neurons in each layer of the decoder are 8, 256, 512, and 24 respectively. The activation function of all neurons of the decoder is ReLU. Both encoder and decoder of SAE in DECA are five-layer multilayer perceptron neural networks. The numbers of neurons in each layer of the encoder are 24, 256, 128, 64, and 3 respectively. The activation function of all neurons of the encoder is ReLU. The numbers of neurons in each layer of the decoder are 3, 64, 128, 256, and 24 respectively. ReLU is employed as the activation functions of the neurons in the first four layers, while Sigmoid is used as the activation functions of the neurons located in the fifth layer. In DECA, the parameters are configured

as $B_1 = 0.9, B_2 = 0.999, X = 0.001$. The details of the interaction electricity prices between the distribution network and the main grid are shown in Figure 3 b).

The energy conversion efficiencies in the distribution network are all set as 0.8; The parameters of the gas turbines are listed in Table 2. The parameters of the energy storage system are listed in Table 3; The parameters of the distribution network are configured according to research [29].

Gas turbine	$P_{\max}$	$P_{\min}$	r_{MTu}	r_{MTd}	a	b
Gas turbine 1	2 500	0	200	200	0.25	0
Gas turbine 2	2 500	0	200	200	0.25	0

Table 2. The parameters of gas turbines

Item	**Value**
$E_{\mathrm{ESS}}^{\max}$/kW	1 500
$E_{\mathrm{ESS}}^{\min}$/kW	100
$P_{\mathrm{Cha}}^{\max} \backslash P_{\mathrm{Cha}}^{\min}$/kW · h	600\0
$P_{\mathrm{Dis}}^{\max} \backslash P_{\mathrm{Dis}}^{\min}$/kW · h	600\0
$\eta_{\mathrm{Cha}} \backslash \eta_{\mathrm{Dis}}$	0.98\0.98
δ	0.01

Table 3. Energy storage system parameters

4.2 Performance Validation of the Generated Data Samples

To demonstrate the effectiveness of the data augmentation approach, the original wind power generation output dataset, which contains a number of 2 434 data samples, is employed. In the first-step data augmentation, SMOTE synthesizes a number of 566 samples. Then, in the second-step data augmentation, VAE participates and generates a number of 2 000 samples based on the original data samples and SMOTE-synthesized data samples. Further, a standalone VAE is also employed to generate a number of 2 566 data samples directly for comparison. A comparison of the cumulative probability distributions is conducted among the original data samples, the samples generated by SMOTE and VAE-based data augmentation, and the samples generated by the standalone VAE.

Figure 4 shows that the differences in terms of data distributions between the original data samples and the generated data samples are inevitable. The data augmentation approaches still have limited performances to emulate the practical data samples. However, it can be observed that the cumulative probability distribution of the data samples generated by the standalone VAE is significantly distant from that of the original data samples. In contrast, the data samples generated using SMOTE and VAE shows a relatively similar performance in the perspective

of the cumulative probability distribution. The result suggests that the combination of SMOTE and VAE can find a compromise between the data volume and the data quality. The presented data augmentation approach is able to generate convincible data samples in enabling the improvement of the data-driven based applications.

4.3 Experimental Result of DECA Based Typical Scenario Identification

In order to evaluate the effectiveness of the presented DECA based typical scenario identification approach, the number of 5 000 wind power generation output samples conducted in the last experiment are employed as the input of DECA. To further optimize the clustering result, SC and DBI are also employed to determine the optimal cluster number for the input data samples.

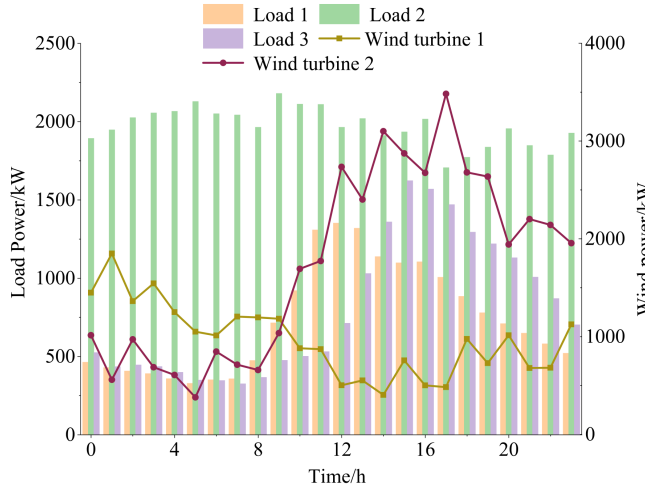
In Figure 5, it can be observed that SC achieves its maximum value and DBI achieves its minimum value when the cluster number k is 2. It suggests that there are a convincible number of 2 classes existing in the data samples. Therefore, in this paper, k=2 is selected as the optimal number of clusters. The parameters of DECA are tuned by Adam iteratively until the value of $\mathrm{LOSS_{KL}}$ is converged. Figure 6 demonstrates the convergence of $\mathrm{LOSS_{KL}}$.

Figure 6 shows that along with the increasing of the iterative number, $\mathrm{LOSS_{KL}}$ converges. Although there are some fluctuations during the convergence, when the iterative number reaches to 342 and $\mathrm{LOSS_{KL}} = 0.0098$, the iteration terminates. At this point, two classes can be achieved. The first class contains a number of 2 513 samples and the second class contains a number of 2 487 samples. The result suggests that there are two typical wind power generation scenarios. In terms of further participating in the economic dispatching optimization, the dimensionality-reduced samples of the two classes are restored to the high-dimensional samples shown in Figure 7 a) and Figure 7 b).

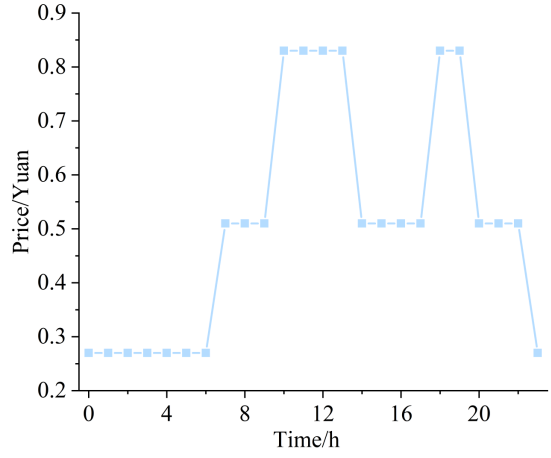
Based on the class centroids, Figure 7 a) and Figure 7 b) indicate that the wind power generation outputs in the first class remain comparatively low between 00:00 and 08:00. And then, the outputs keep ascending and reach to a peak value at 18:00. Afterwards, the outputs are reduced and become relatively stable. In contrast, the wind power generation outputs in the second class maintain a high level during the period 00:00 to 08:00. And then, the outputs reduce and maintain a low level for the rest of time, though there are several negligible peak values. The results suggest that two typical classes exhibit a remarkable difference. The results also suggest that two typical scenarios represented by the two classes can be clearly identified, which proves the effectiveness of the presented DECA.

4.4 Performance of IDDM Based Ambiguity Set

To demonstrate the effectiveness of the IDDM based ambiguity, different numbers of wind power generation output samples from two typical scenarios are selected to build up the sets. The details of the samples are shown in Table 4.



a) Predicted values of wind power generation outputs and loads



b) Interaction electricity prices between distribution network and main grid

Figure 3. Scenario data for the proposed dispatching model

Scenario	Number 1	Number 2	Number 3	γ
Class 1	100	1 000	2 513	95 %
Class 2	100	1 000	2 487	95 %

Table 4. Numbers of samples from two typical scenarios

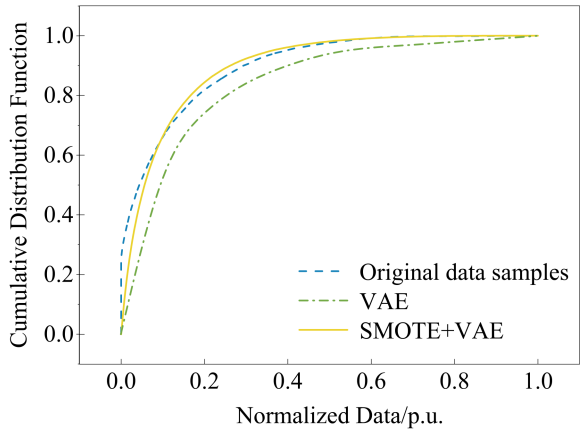


Figure 4. Cumulative probability distribution comparisons among three approaches

Two ambiguity sets with different numbers of samples are constructed for two typical scenarios. The results are shown in Figure 8 a) and Figure 8 b).

In Figure 8, it can be observed that smaller numbers of samples result in broader boundaries of the ambiguity sets. In contrast, larger numbers of samples can effectively narrow the boundaries. This is because the increasing number of samples enlarges the total observations, which narrows the confidence interval boundaries and consequently enhances the estimation accuracy of the ambiguity set for de-

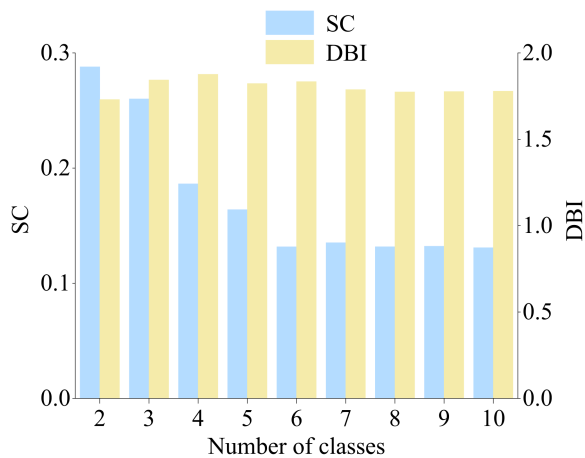


Figure 5. Convergence of $\mathrm{LOSS}_{\mathrm{KL}}$

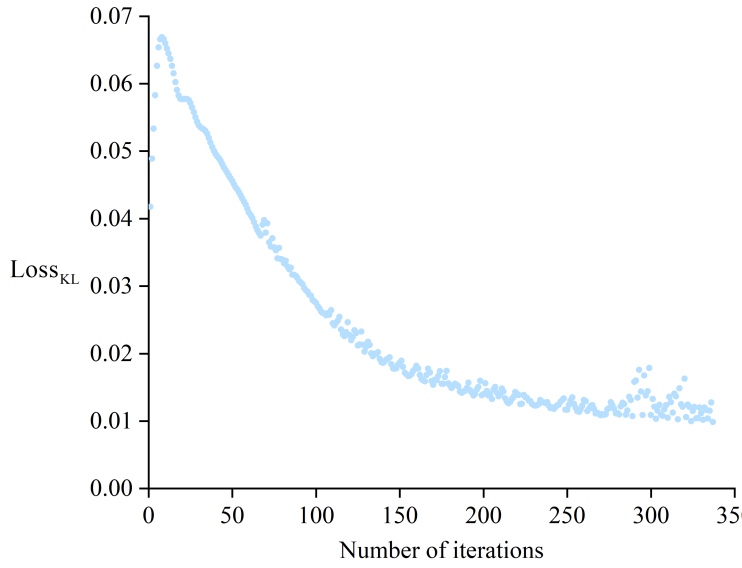


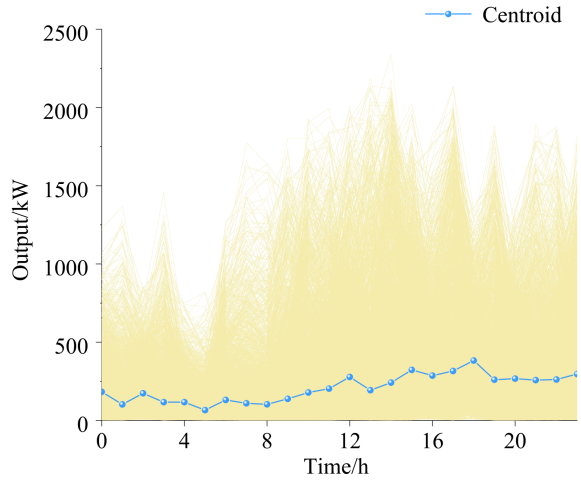
Figure 6. Convergence of $\mathcal{L}_{\mathrm{KL}}$

picting the event occurrence probability. The results also indicate that under the conditions of scarce wind power generation output samples, the presented SMOTE and VAE based data augmentation method can effectively enhance the precision of ambiguity sets in characterizing wind power generation output.

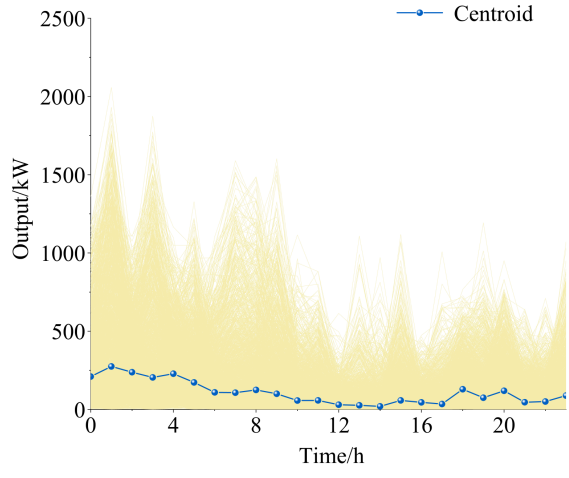
4.5 Evaluation of DRO Based Economic Dispatching

In this section, let the ambiguity set constructed using the samples from the first typical scenario depict the uncertainty of wind turbine 1; the ambiguity set constructed using the samples from the second typical scenario depicts the uncertainty of wind turbine 2. The ambiguity sets with increasing numbers of samples are employed by DRO. The optimization results are listed in Table 5. In terms of comparison, ARO is also adopted. The uncertainty set employed by ARO is box-type, in which the prediction error is 0.3. Additionally, the adjustable robust parameters Γ_{WPG} and Γ_{Load} of both ARO and DRO are set to 10. The comparisons of the dispatching costs are also listed in Table 5.

Table 5 indicates that in terms of total cost, day-ahead cost, and real-time cost, DRO outperforms ARO. The results suggest that the presented IDDM based ambiguity set can characterize the uncertainty more precisely than the typical box-type ambiguity set. And also, DRO specifically focuses on utilizing the probability

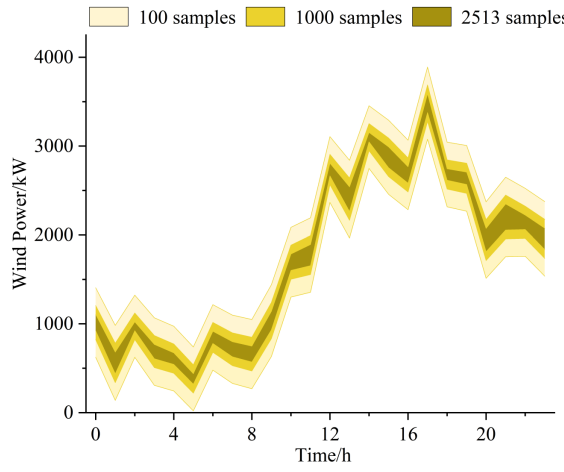


a) Class 1

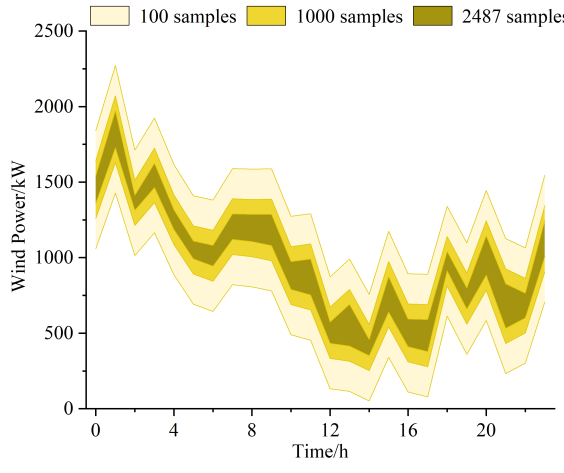


b) Class 2

Figure 7. Samples of two typical wind power generation scenarios



a) Wind turbine 1



b) Wind turbine 2

Figure 8. IDDM based ambiguity sets with different numbers of data samples

Item	Total cost (Yuan)	Day-ahead cost (Yuan)	Real-time cost (Yuan)
ARO	24 255.09	21 835.21	2 420.78
DRO (200 samples)	24 116.08	21 703.14	2 412.94
DRO (2 000 samples)	23 559.09	21 724.52	1 834.57
DRO (5 000 samples)	23 118.23	21 578.48	1 539.75

Table 5. Dispatching costs of ARO and DRO

distributions hidden in the samples, so the costs can be minimized under the worst-case distribution within the ambiguity set. It hedges against the risk arising from the misjudgment of the probability distribution. Therefore, even if an imperfect distributional estimation may occur, less adjustment cost is needed to implement the energy balance, which results in a reasonably good result. However, ARO focuses on "plan then adapt", which adjusts the plan as long as an uncertainty event occurs. Therefore, it performs more conservatively than the DRO does. The table also indicates that along with the increasing number of wind power generation output samples, the total costs can be significantly reduced using DRO. The reason is the boundaries of the ambiguity sets are significantly narrowed by IDDM, which leads to a more accurate characterization of the wind power generation output. Therefore, the adjustment cost due to the imprecise wind power forecasting in the real-time stage is commensurately lowered. Meanwhile, the final optimal economic dispatching strategy is listed in Figure 9.

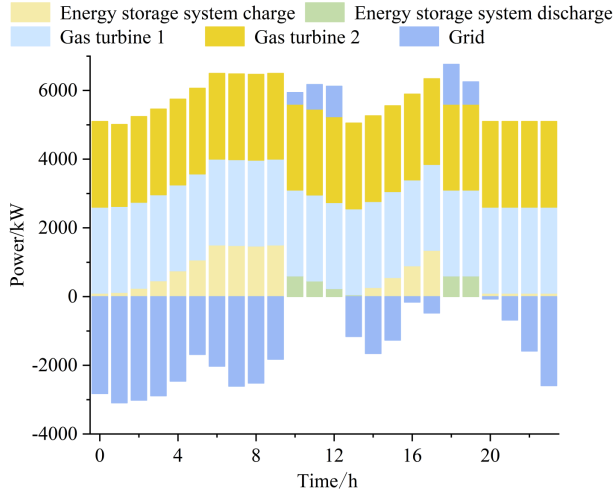


Figure 9. Optimal economic dispatching strategy of distribution network

Figure 9 shows that two gas turbines keep working with their full capacities to satisfy the load demand of the distribution network all day. And also, in terms of achieving a better economic benefit, when the wind generation is high and electricity demand is low, the distribution network sells the surplus electrical power to the main grid. Meanwhile, the extra power which cannot be consumed or sold is stored in the energy storage system. During the periods of low wind power generation output but high load demand (10:00–12:00, 18:00–19:00), the energy storage system discharges its stored electricity to satisfy the load demand.

5 CONCLUSION

In order to implement economic dispatching for a distribution network in a power system, this paper presents an optimal dispatching approach involving data augmentation and modeling of power source-load uncertainties. The approach firstly employs SMOTE and VAE to solve the issue of insufficient wind power generation output data samples. And then, DECA is presented to provide the identification of typical wind power generation output scenarios, which further facilitates the modeling of the uncertainty. Afterwards, this paper presents IDDM based ambiguity set construction method. The method successfully narrows the boundaries of the ambiguity set which can precisely characterize the wind generation power output uncertainty. Finally, based on the aforementioned approaches and methods, a DRO model enabling the optimal economic dispatch for the distribution network is built. Based on the solution of the model, the experimental results indicate that:

1. The SMOTE and VAE based data augmentation approach is able to enrich the wind power generation output data. SMOTE provides a preliminary expansion of the data samples. VAE learns the data distribution within the dataset and further generate relatively high-quality data samples. Based on the combination of two algorithms, the excessive impacts of their inherent limitations on the data quality can be significantly mitigated.
2. The presented DECA is able to implement the dimensionality reduction of the high-dimensional wind power generation output data samples, while it also provides a convincible performance for identifying the typical scenarios inside the data samples. Compared to conventional algorithms, DECA is able to fully characterize the data samples, and enables more comprehensive mining of their features.
3. The presented IDDM based ambiguity set is able to fully utilize the probabilities of the uncertainties hidden in the data samples. And also, along with the increasing number of data samples, IDDM can significantly narrows the boundaries of the ambiguity set, which limits the probabilities of the uncertainties in a narrower probability band. Therefore, the uncertainties can be characterized more precisely.

4. DRO based economic dispatching model is able to provide enough compatibility for modeling the comprehensive units and devices participated in the distribution network. And also, DRO performs highly effectively at handling uncertainties on both the power generation and load sides. Therefore, it significantly improves the economic benefit of the distribution network dispatching.

However, there still exists several problems in our work, which will be further addressed in future work:

1. Data generation may potentially affect the data distribution, especially in certain dataset with extremely unbalanced classes. Therefore, the observations of certain states in the generated samples may be distorted, which further impacts the precision of the ambiguity set. How to refine the distribution of generated data samples to be closer to that of the original data samples is a critical focus.
2. In the current stage of research, the sample sizes generated by the SMOTE and VAE stages respectively have not been systematically considered. Future study should investigate how to optimize the effectiveness of the two-stage sample generation process across different datasets.
3. This paper provides limited work on optimizing the dispatching strategy for an individual distribution network. Considering the interconnections among the distribution networks, subsequent research will conduct in-depth research on collaborative optimal dispatch for multiple distribution networks.

Acknowledgements

This work was supported by the Economic Research Institute of State Grid Hebei Electric Power Co., Ltd. under the project "The Research of Optimal Operation for Distribution Network Considering Uncertainty of Renewable Energy and New Types of Distributed Load" [Grant No. KJ2024-023].

REFERENCES

[1] Basit, A.—Hansen, A. D.—Altin, M.—Sørensen, P. E.—Gamst, M.: Compensating Active Power Imbalances in Power System with Large-Scale Wind Power Penetration. Journal of Modern Power Systems and Clean Energy, Vol. 4, 2016, No. 2, pp. 229–237, doi: 10.1007/s40565-015-0135-x.

[2] Ding, T.—Liu, S.—Yuan, W.—Bie, Z.—Zeng, B.: A Two-Stage Robust Reactive Power Optimization Considering Uncertain Wind Power Integration in Active Distribution Networks. IEEE Transactions on Sustainable Energy, Vol. 7, 2016, No. 1, pp. 301–311, doi: 10.1109/TSTE.2015.2494587.

[3] Keane, A.—Cuffe, P.—Diskin, E.—Brooks, D.—Harrington, P.—Hearne, T.—Rylander, M.—Fallon, T.: Evaluation of Advanced Operation

and Control of Distributed Wind Farms to Support Efficiency and Reliability. IEEE Transactions on Sustainable Energy, Vol. 3, 2012, No. 4, pp. 735–742, doi: 10.1109/TSTE.2012.2200000.

[4] Krahmer, S.—Ecklebe, S.—Schegner, P.—Röbenack, K.: Application of Stability Analysis of Q(V)-Characteristic Controls Related to the Converter-Driven Stability in Distribution Networks. IEEE Transactions on Industry Applications, Vol. 60, 2024, No. 3, pp. 5002–5011, doi: 10.1109/TIA.2024.3360023.

[5] Nirbhavane, P. S.—Corson, L.—Rizvi, S. M. H.—Srivastava, A. K.: TPCPF: Three-Phase Continuation Power Flow Tool for Voltage Stability Assessment of Distribution Networks with Distributed Energy Resources. IEEE Transactions on Industry Applications, Vol. 57, 2021, No. 5, pp. 5425–5436, doi: 10.1109/TIA.2021.3088384.

[6] Ge, J.—Wu, D.: A Novel Intelligence Analysis Method Based on Graph Convolutional Neural Networks for Online Security of Power Distribution Networks. 2024 11th International Forum on Electrical Engineering and Automation (IFEEA), 2024, pp. 550–554, doi: 10.1109/IFEEA64237.2024.10878514.

[7] Hu, D.—Sun, L.—Ding, M.: Integrated Planning of an Active Distribution Network and DG Integration in Clusters Considering a Novel Formulation for Reliability Assessment. CSEE Journal of Power and Energy Systems, Vol. 9, 2023, No. 2, pp. 561–576, doi: 10.17775/CSEEJPES.2020.00150.

[8] Yang, Y.—Qiu, J.—Zhang, C.: Distribution Network Planning Towards a Low-Carbon Transition: A Spatial-Temporal Carbon Response Method. IEEE Transactions on Sustainable Energy, Vol. 15, 2024, No. 1, pp. 429–442, doi: 10.1109/TSTE.2023.3294532.

[9] Chen, Y.—Li, H.—Liu, Y.—Wu, S.—Li, X.—Qi, M.: Two-Stage Trading Mechanism in Enabling the Design and Optimization of Flexible Resources Interaction in Smart Grid. Computing and Informatics, Vol. 43, 2024, No. 6, pp. 1483–1515, doi: 10.31577/cai_2024_6_1483.

[10] Chen, X.—Liu, Y.—Wu, L.: Towards Improving Unit Commitment Economics: An Add-on Tailor for Renewable Energy and Reserve Predictions. IEEE Transactions on Sustainable Energy, Vol. 15, 2024, No. 4, pp. 2547–2566, doi: 10.1109/TSTE.2024.3426337.

[11] Zhao, A. P.—Li, S.—Xie, D.—Wang, Y.—Li, Z.—Hu, P. J. H.—Zhang, Q.: Hydrogen as the Nexus of Future Sustainable Transport and Energy Systems. Nature Reviews Electrical Engineering, Vol. 2, 2025, pp. 447–466, doi: 10.1038/s44287-025-00178-2.

[12] Wang, G.—Wang, C.—Feng, T.—Wang, K.—Yao, W.—Zhang, Z.: Day-Ahead and Intraday Joint Optimal Dispatch in Active Distribution Network Considering Centralized and Distributed Energy Storage Coordination. IEEE Transactions on Industry Applications, Vol. 60, 2024, No. 3, pp. 4832–4842, doi: 10.1109/TIA.2024.3372943.

[13] Liu, Y.—Chen, X.—Li, B.—Li, H.—Ye, Y.: WKS-Type Distributionally Robust Optimisation for Optimal Sub-Hourly Look-Ahead Economic Dispatch. IET Generation Transmission & Distribution, Vol. 14, 2020, No. 12, pp. 2237–2246, doi: 10.1049/iet-gtd.2019.1344.

[14] Liu, J.—Wang, M.—Guo, J.—Ge, J.: Multiple Time Resolution Dispatching Model for Distribution Network Considering Various Active Management Measures. 2023 5th Asia Energy and Electrical Engineering Symposium (AEEES), 2023, pp. 803–809, doi: 10.1109/AEEES56888.2023.10114254.

[15] Liu, Y.—Nair, N. K. C.: A Two-Stage Stochastic Dynamic Economic Dispatch Model Considering Wind Uncertainty. IEEE Transactions on Sustainable Energy, Vol. 7, 2016, No. 2, pp. 819–829, doi: 10.1109/TSTE.2015.2498614.

[16] Soares, J.—Canizes, B.—Ghazvini, M. A. F.—Vale, Z.—Venayagamoorthy, G. K.: Two-Stage Stochastic Model Using Benders' Decomposition for Large-Scale Energy Resource Management in Smart Grids. IEEE Transactions on Industry Applications, Vol. 53, 2017, No. 6, pp. 5905–5914, doi: 10.1109/TIA.2017.2723339.

[17] Liu, Y.—Ye, Y.—Chen, X.—Li, H.—Huang, Y.: Robust Day-Ahead Dispatch for Integrated Power-Heat-Gas Microgrid Considering Wind Power Uncertainty. Mathematical Problems in Engineering, Vol. 2020, 2020, Art. No. 4215906, doi: 10.1155/2020/4215906.

[18] Chen, X.—Liu, Y.—Li, B.: Adjustable Robust Optimization in Enabling Optimal Day-Ahead Economic Dispatch of CCHP-MG Considering Uncertainties of Wind-Solar Power and Electric Vehicle. Journal of Industrial and Management Optimization, Vol. 17, 2021, No. 4, pp. 1639–1661, doi: 10.3934/jimo.2020038.

[19] Wu, J.—Liu, Y.—Chen, X.—Wang, C.—Li, W.: Data-Driven Adjustable Robust Day-Ahead Economic Dispatch Strategy Considering Uncertainties of Wind Power Generation and Electric Vehicles. International Journal of Electrical Power and Energy Systems, Vol. 138, 2022, Art. No. 107898, doi: 10.1016/j.ijepes.2021.107898.

[20] Saberi, H.—Zhang, C.—Dong, Z. Y.: Data-Driven Distributionally Robust Hierarchical Coordination for Home Energy Management. IEEE Transactions on Smart Grid, Vol. 12, 2021, No. 5, pp. 4090–4101, doi: 10.1109/TSG.2021.3088433.

[21] Zou, C.—Wang, K.—Xiahou, T.—Cao, D.—Liu, Y.: Two-Stage Distributionally Robust Optimization for Infrastructure Resilience Enhancement: A Case Study of 220 kV Power Substations Under Earthquake Disasters. IEEE Transactions on Reliability, Vol. 74, 2025, No. 2, pp. 2576–2590, doi: 10.1109/TR.2024.3450107.

[22] Barua, S.—Islam, M. M.—Yao, X.—Murase, K.: MWMOTE–Majority Weighted Minority Oversampling Technique for Imbalanced Data Set Learning. IEEE Transactions on Knowledge and Data Engineering, Vol. 26, 2014, No. 2, pp. 405–425, doi: 10.1109/TKDE.2012.232.

[23] Wang, K.—Govindarasu, M.: FGSM-Based Synthetic Data Generation Technique and Application to Anomaly Detection in Smart Grid. IEEE Power & Energy Society General Meeting (PESGM), 2024, pp. 1–5, doi: 10.1109/PESGM51994.2024.10688539.

[24] Dey, I.—Pratap, V.: A Comparative Study of SMOTE, Borderline-SMOTE, and ADASYN Oversampling Techniques Using Different Classifiers. 2023 3rd International Conference on Smart Data Intelligence (ICSMDI), 2023, pp. 294–302, doi: 10.1109/ICSMDI57622.2023.00060.

[25] Huang, Y.—Liu, Y.—Xu, L.—Guo, H.: A Residential User Classification Ap-

proach Based on the Graded Portrait with Considering the IDR Adaptability and Potential. Frontiers in Energy Research, Vol. 10, 2022, Art. No. 1012721, doi: 10.3389/fenrg.2022.1012721.

[26] Fard, M. M.—Thonet, T.—Gaussier, E.: Deep k-Means: Jointly Clustering with k-Means and Learning Representations. Pattern Recognition Letters, Vol. 138, 2020, pp. 185–192, doi: 10.1016/j.patrec.2020.07.028.

[27] Chen, X.—Ma, Y.—Chen, X.—Liu, Y.: Data-Driven Adjustable Robust Optimization of Day-Ahead Economic Dispatch of Integrated Energy System with Combined Cool, Heat and Power System. Electric Power Construction, Vol. 42, 2021, No. 1, pp. 28–40, doi: 10.12204/j.issn.1000-7229.2021.01.004 (in Chinese).

[28] Keras. `https://keras.io/` (Last Accessed: 09/11/2025).

[29] Wu, S.—Sheng, W.—Shang, Y.—Liu, K.—Wang, S.—Ma, H.: Exploration of the Evolution of New Distribution System Forms: A Percolation Theory Perspective. Chinese Journal of Electrical Engineering, Vol. 46, 2026, No. 11, pp. 4365–4377, doi: 10.13334/j.0258-8013.pcsee.252441.



Yunjia Wang received her Master's degree from the North China Electric Power University (Baoding) in 2019. She is working with the Economic Research Institute, State Grid Hebei Electric Power Co., Ltd. Her research interests focus on energy and power economics research.



Ning Pang received his Master's degree from the University of Manchester in 2013. He is working with the Economic Research Institute, State Grid Hebei Electric Power Co., Ltd. His research interests focus on power system engineering.



Hongguang Yu received his Master's degree from Hebei Normal University in 2020. He is working with the Economic Research Institute, State Grid Hebei Electric Power Co., Ltd. His research interests focus on big data analysis in the electrical power system.



Yan Zhang received her Master's degree from Tianjin University in 2020. She is working with the Economic Research Institute, State Grid Hebei Electric Power Co., Ltd. Her research interests focus on load demand forecasting and electrical power economics research.



Guozhen Ma received his Master's degree from the North China Electric Power University (Baoding) in 2010. He is working with the Economic Research Institute, State Grid Hebei Electric Power Co., Ltd. His research interests focus on energy and the power economics research.



Shiyao Hu received his Master's degree from the North China Electric Power University in 2018. He is working with the Economic Research Institute, State Grid Hebei Electric Power Co., Ltd. His research interest is focusing on large-scale integration and coordinated control of clusters for distributed renewable energy.



Yang Liu received his doctoral degree from Brunel University in 2011. He is working with the College of Electrical Engineering, Sichuan University. His research interests focus on power system planning and operation, machine learning technologies and big data analysis in electrical power systems.