Energy 243 (2022) 123113

Contents lists available at ScienceDirect

Energy

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

A data-enhanced distributionally robust optimization method for economic dispatch of integrated electricity and natural gas systems with wind uncertainty

Baining Zhao, Tong Qian*, Wenhu Tang, Qiheng Liang

School of Electric Power Engineering, South China University of Technology, Guangzhou, 510 641, China

ARTICLE INFO

Article history:
Received 27 June 2021
Received in revised form
22 November 2021
Accepted 2 January 2022
Available online 5 January 2022

Keywords:
Integrated electricity and natural gas systems
Economic dispatch
Distributionally robust optimization
Generative adversarial networks
Wind uncertainty

ABSTRACT

With growing penetrations of wind power in electricity systems, the coordinated dispatch of integrated electricity and natural gas systems is becoming a popular research topic. Distributionally robust optimization can cope with the wind uncertainty of integrated electricity and natural gas systems by providing optimal solutions for the worst-case probability distribution. However, limited historical wind data hinder the estimation of worst-case probability distribution. As a breakthrough in artificial intelligence, generative adversarial networks can be established to approximate a complex uncertain probability distribution from raw data and generate realistic data subject to the identical distribution. This paper proposes a data-driven optimization method for economic dispatch of integrated electricity and natural gas systems with wind uncertainty, whose probability distribution is free. Based on limited historical data, the data-driven generative adversarial network generates artificial wind power data, which helps to improve the estimation of worst-case probability distribution in distributionally robust optimization. Moreover, the robustness of optimization solutions can be adjusted cost-effectively by controlling the auxiliary data number. In a case study, optimization solutions of the proposed method are shown to achieve a lower probability of chance constraint violation at a nearly negligible cost increase compared with those from four typical optimization methods.

1. Introduction

In recent years, the proliferation of renewable energy reaps increasing concern and application for energy security, environmental protection and climate change in all countries of the world [1]. Wind power gradually plays an important role in power generation. However, it is known that intermittency and fluctuation of wind power makes it difficult to increase the proportion of wind energy consumption [2]. The integration of wind power augments the uncertainties in electricity systems and significantly affect the stability and safety of system operation [3]. In order to promote wind power consumption, the natural gas units (NGUs) are utilized to deal with wind uncertainty because of their flexibility and fast response capability [4]. With the number of NGU installations growing dramatically, the electricity system and gas system are linked tightly. The uncertainty of wind generation output connected to electricity system affects the operation of the gas system directly, thus the two systems should be analyzed simultaneously to achieve greater economic benefits [5].

Economic dispatch is one of the most important problems of integrated electricity and natural gas systems (IEGS) with wind uncertainty [6]. Many optimization methods including stochastic optimization, robust optimization, distributionally robust optimization (DRO) aim to produce solutions that are immunized against wind uncertainty. In Ref. [7], a new stochastic decomposition method was presented, which was well-suited to deal with the unit commitment problems. In Ref. [8], a two-stage method and a multistage method based on stochastic programming were built to analyze the impact of wind uncertainty on IEGS. However, stochastic optimization is sensitive to the perturbation in the probability distribution of uncertain data. The optimal solutions could have poor optimization performances if the designed distribution is far away from the real distribution of uncertain data. In Ref. [9], the robust optimization was used to make optimal power flow achieve

* Corresponding author.
E-mail addresses: epzhaobn@mail.scut.edu.cn (B. Zhao), qian.tong@mail.scut.edu.cn (T. Qian), wenhutang@scut.edu.cn (W. Tang), epnhlqsh@mail.scut.edu.cn (Q. Liang).

https://doi.org/10.1016/j.energy.2022.123113
0360-5442/

Nomenclature

Variables

P	Output of thermal generation units, including that of natural gas units (MW)
G	Production of gas wells (kcf/h)
Y_P	Participation factor of thermal generation units' power output
Y_G	Participation factor of gas wells' gas output
$C_P, \mathcal{C}_P$	Dispatch cost of electricity system ($)
$C_G, \mathcal{C}_G$	Dispatch cost of natural gas system ($)
X	Gas consumption of natural gas units (kcf/h)
Y_X	Participation factor of natural gas units' gas consumption
H^{Gen}	Generator network in generative adversarial network
H^{Dis}	Discriminator network in generative adversarial network
L_G	Loss term of the generator
L_D	Loss term of the discriminator
L_D	Loss term of the discriminator
L_{GAN}	Adversarial loss of the original generative adversarial network
L_{WGAN}	Adversarial loss of the Wasserstein generative adversarial network
C_{emp}	Empirical dispatch cost ($)
C_I	Dispatch cost of simulations on test sets ($)
V	Violation probability for test samples
V_{mean}	Empirical violation probability
σ	Empirical standard deviation of the violation probability
$\mathbb{I}$	Indicator whether the chance constraints are violated under test sets

Parameters

$\tilde{W}$	Wind power output (MW)
W	Forecast output of wind power (MW)
ξ	Forecast error of wind power (MW)
$c_{P,0}$, $c_{P,1}$, $c_{P,0}$	Power consumption coefficients of thermal generation units
$c_{G,0}$, $c_{G,1}$, $c_{G,0}$	Gas consumption coefficients of gas wells
P^{min}/P^{max}	Power generation upper and lower limits of thermal generation units (MW)
G^{min}/G^{max}	Gas production upper and lower limits of gas wells (kcf/h)
I_P	Incidence matrix for mapping thermal generation units to the electricity network
I_W	Incidence matrix for mapping wind generation units to the electricity network
I_{PD}	Incidence matrix for mapping electricity demand to the electricity network
D_P	Electricity power demand (MW)
D_G	Natural gas demand (kcf/h)
F_P	Upper limit of active power transmitted by power lines
F_G	Upper limit of natural gas flow through pipelines
I_G	Incidence matrix for mapping gas wells to the gas network
I_{GD}	Incidence matrix for mapping gas demand to the gas network
I_{GP}	Incidence matrix for mapping natural gas units to the gas network
ϵ	Risk level
Φ	Power conversion factor
A	Part related to uncertain data
B	Certain part in chance constraints
C	Constant difference between the real probability distribution function of wind power output and that of wind power output
$\mathcal{D}$	Ambiguity set, including a family of possible candidate distribution built on historical data
$\hat{\mu}$	Empirical mean value of wind power
μ	Real mean value of wind power
$\hat{\Sigma}$	Empirical covariance matrix of wind power
$\mathcal{K}_\epsilon$	Coefficient to control the robustness of chance constraints
ψ	Cumulative distribution function of the Gaussian distribution
z	Random noise vectors
$F(\xi)$	Empirical probability distribution function of forecast errors
$p_{\tilde{\xi}}(\xi)$	Real probability distribution function of forecast errors
$p_{\tilde{W}}(\tilde{W})$	Real probability distribution function of wind power output
$p_{\tilde{Z}}(z)$	Simple and known probability distribution of random noise vectors z
N	Number of historical data in one training set.
N'	Number of historical data in one test set.
N''	Number of generated data in auxiliary training set.
Υ'_N	Training set.
$\Theta_{N'}$	Test set.
$\hat{\Upsilon}_{N''}$	Auxiliary training set.
M	Number of simulations
N_I	Number of chance constraints in integrated electricity and natural gas systems

Subscripts and Superscripts

q	Wind power generation unit
i	Thermal generation unit
j	Gas well
v	Power demand.
r	Natural gas unit
k	Gas demand.
l	Power line.
s	Pipeline.
h	Chance constraints
p	Historical data
a	Chance constraint
m	Data set of simulation

minimized generation cost in uncertain operating states. To relieve the conservatism of robust optimization solutions, the power system operating decisions in Ref. [10] are divided into two processes: pre-dispatch considering the worst-case scenario in robust optimization and re-dispatch based on the nominal scenario. However, the results of robust optimization are still too conservative, which brings expensive cost for IEGS. In Ref. [11], a DRO model was proposed to solve unit commitment problems considering volatile wind power generation. For the problem of the day-ahead operation of power systems, the DRO was utilized to build a decision-

making framework and handle a proposed two-stage gas contracting mechanism in Ref. [12]. The DRO is good at the dispatch of IEGS when the probability distribution of wind power can not be inferred by limited historical data.

The DRO is a kind of data-driven stochastic robust program. The ambiguity set designed in DRO consists of a set of possible candidates, the worst case of which is considered to solve the system dispatch problem [13]. Compared with insufficient historical data, relatively more data may improve the reliable estimation of distributions in the ambiguity set. Then the worst-case probability distribution can be considered robustly, resulting in better optimization solutions [14].

A related method of wind data generation is scenario generation, widely used in scenario analysis of stochastic programming [15]. Scenario generation method can generate a set of discrete scenarios subject to the distribution of historical data, which transforms the stochastic optimization problem into a deterministic problem [16]. In Ref. [17], the wind power forecast error distribution was assumed to follow the joint normal distribution. After the distribution parameters were obtained, scenarios of wind power forecast error could be obtained by sampling. In Ref. [18], forecast errors were characterized via empirical distributions of a set of forecast bins and wind power fluctuations were assumed to follow the location-scale distribution, by which wind power data were generated. Existing literature mainly focused on model-based methods to generate wind power data, which highly depends on the supposed probability model. As for data-driven methods, a generative adversarial network (GAN) [19] is a popular generative model, which can automatically learn the implicit data probability distribution and sample new data easily. GANs have been successfully utilized to generate realistic data in many fields including image, series data generation, and scenario generation. Different kinds of GANs were utilized to generate data of wind power successfully in Ref. [20]. A wind power data generation framework proposed in Ref. [21] was based on the conditional Wasserstein generative adversarial network (WGAN). The generated data of wind power are similar to real data in the statistical analysis. However, data generated by GANs-based methods have not been verified to be effective in the specific stochastic optimization problem, e.g., unit commitment, economic dispatch, etc. Additionally, the GAN training in the above-mentioned literature requires a lot of historical data because of the deep structures of GANs. However, enough historical data may not be available in actual industrial applications.

The required historical data size to estimate the ambiguity set could not be too small for DRO problems with a huge number of decision variables and a small value of risk level. In many cases, the available amount of uncertainty data might not be enough for the purpose of proper estimation of the worst-case probability distribution [22]. However, the number of uncertainty data can still be sufficient for training a rough generative model [23]. It is feasible that the generative model generates realistic uncertainty data to meet the data requirement of the ambiguity set estimation in DRO. In this paper, the distributionally robust optimization method is enhanced by generated data using a GAN, which is applied in the economic dispatch problem of IEGS with wind uncertainty. A GAN is trained to learn the probability distribution of uncertain wind power implicitly and generate uncertainty data subject to the distribution. More historical data enable the worst-case probability distribution to comprehensively cover the real distribution of uncertain wind power. With the improved ambiguity set, the data-enhanced DRO can provide a cost-effective solution to enhance the robustness of the IEGS economic dispatch.

The main contributions of this paper are as follows.

1) A novel data-enhanced distributionally robust optimization method is proposed for the IEGS economic dispatch. By using a GAN, realistic historical data can be generated, which is beneficial to establishment of a more robust ambiguity set in DRO. To authors' knowledge, this is the first work applying generative adversarial networks to improve the performance of distributionally robust optimization.
2) The proposed data-enhanced DRO is data-driven. In both historical data generation and dispatch optimization process, the probability distribution of uncertain parameters is free and only limited wind power data are required.
3) By simplifying the GAN structure, the required historical data size of the rough GAN training is smaller than that of ambiguity set establishment.
4) Using the proposed optimization method, the violation probability of the IEGS dispatch is reduced significantly at a slight increase of dispatch cost. The robustness of the optimization solutions can be controlled through the number of auxiliary data.

The rest of this paper is organized as follows. The model of IEGS with wind uncertainty is built in Section 2. Section 3 proposes the historical wind power data augmentation using the GAN for a more reliable estimation of the worst-case probability distribution in DRO. Section 4 performs the case study on the IEEE 24-bus electricity system coupled with a 12-node natural gas system and analyzes the dispatch results of different optimization methods. Finally, conclusions are drawn in Section 5.

2. Modeling of integrated electricity and natural gas systems with wind uncertainty

As shown in Fig. 1, the model of integrated electricity and natural gas systems with wind generation uncertainty is built. First, the wind generation uncertainty in IEGS is illustrated. Then, the mathematical formulation of the IEGS with wind uncertainty problem is expressed by utilizing chance constraints. It consists of system economic cost and constraints of the electricity system, gas system, and their linkage. Finally, three distribution assumptions for wind generation uncertainty are used to convert the objective function and chance constraints to an easily solvable reformulation.

2.1. Wind generation uncertainty

The wind uncertainty directly affects wind generation outputs. For each wind generation unit, the wind generation output $\tilde{W}_q$ is the uncertain data in DRO. $\tilde{W}_q$ can be divided into certain and

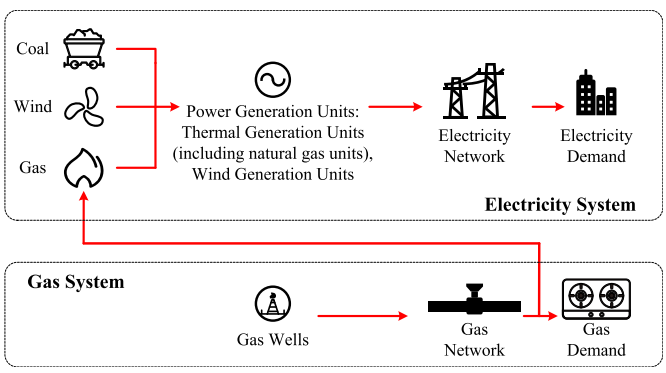

Fig. 1. Integrated electricity and natural gas systems considering wind power and natural gas units.

uncertain parts. The certain part is recognized as the forecast output W_q and another is the forecast error ξ_q.

$$\tilde{W}_q = W_q + \xi_q \tag{1}$$

2.2. Integrated electricity and natural gas systems model considering chance constraints

The integrated electricity and natural gas systems with wind uncertainty is modelled by an objective function, constraints of the electricity system, constraints of the gas system, and the linkage between the two systems.

2.2.1. Objective function

A chance-constrained optimal model of integrated power and gas flow in the steady state is built to investigate the economic dispatch problem [24]. All decision variables are here-and-now type. The objective of the IEGS considering chance constraints is to minimize the operational cost of electricity and gas systems, as shown in (2).

$$\begin{aligned}
&\min_{P,Y_P,G,Y_G} \max_{F(\xi)\in\mathcal{D}} E^{F(\xi)}\left(C_P(P,Y_P,\xi) + C_G(G,Y_G,\xi)\right)\\
&C_P(P,Y_P,\xi) = \sum_i \left[c_{P,2}\left(P_i + \sum_q Y_{P,i,q}\xi_q\right)^2 + c_{P,1}\left(P_i + \sum_q Y_{P,i,q}\xi_q\right) + c_{P,0}\right]\\
&C_G(G,Y_G,\xi) = \sum_j \left[c_{G,2}\left(G_i + \sum_q Y_{G,j,q}\xi_q\right)^2 + c_{G,1}\left(G_i + \sum_q Y_{G,j,q}\xi_q\right) + c_{G,0}\right]
\end{aligned} \tag{2}$$

where $\mathcal{P}$ is a family of possible candidate distribution built on historical data. The objective function is to minimize the maximum system risk cost caused by any distribution $F(\xi)$ in $\mathcal{P}$. Additionally, the thermal generation units and gas well operate according to day-ahead scheduling in the first stage of the decision making. They maintain power balance for the forecast error of wind generation through the participation factors Y_P and Y_G in the real-time scheduling stage. The optimization problem is subject to constraints of electricity and gas systems as follows.

2.2.2. Electricity system

In order to reduce the difficulty of solution, only linear direct-current flows without loss is taken into account. The matrices of power transfer distribution factor I_P, I_W, I_{PD} are utilized to build the incidence matrices from the thermal generation units, wind generation units, and electricity demand to the electricity network [3]. Then the constraints of electricity system are written as

$$P_i^{\min} \le P_i \le P_i^{\max} \tag{3}$$

$$\sum_i \left(P_i + \sum_q Y_{P,i,q}\xi_q\right) + \sum_q \tilde{W}_q = \sum_v D_{P,v} \tag{4}$$

$$\Pr\left(P_i + \sum_q Y_{P,i,q}\xi_q \le P_i^{\max}\right) \ge 1-\epsilon \tag{5}$$

$$\Pr\left(-P_i - \sum_q Y_{P,i,q}\xi_q \le -P_i^{\min}\right) \ge 1-\epsilon \tag{6}$$

$$\Pr\left(\sum_i I_{P,l,i}\left(P_i + \sum_q Y_{P,i,q}\xi_q\right) + \sum_q I_{W,l,q}\tilde{W}_q - \sum_v I_{PD,l,v}D_{P,v} \le F_{P,l}\right) \ge 1-\epsilon \tag{7}$$

$$\Pr\left(-\sum_i I_{P,l,i}\left(P_i + \sum_q Y_{P,i,q}\xi_q\right) - \sum_q I_{W,l,q}\tilde{W}_q + \sum_v I_{PD,l,v}D_{P,v} \le F_{P,l}\right) \ge 1-\epsilon \tag{8}$$

The power generation limits of thermal units are given in (3). (4) is the power balance equation in the first and second stages. The chance constraints on generation and transmission limits are presented in (5), (6), (7), (8). Without loss of generality, suppose the risk level of ϵ in every constraint is identical.

2.2.3. Gas system

The steady-state gas system is built for computation reduction. And the influence of the temperature, flow rate and pipe friction on node gas pressure and filling volume is ignored. Similar to the electricity system, the matrices I_G, I_{GD}, I_{GP} represent the incidence matrices from the gas wells, gas demand, and NGUs to the gas network. Thus, the constraints of the gas system are formulated as below.

$$G_j^{\min} \le G_j \le G_j^{\max} \tag{9}$$

$$\sum_j \left(G_j + \sum_q Y_{G,j,q}\xi_q \right) = \sum_r \left(X_r + \sum_q Y_{X,r,q}\xi_q \right) + \sum_k D_{G,k} \tag{10}$$

$$\Pr\left(G_j + \sum_q Y_{G,j,q}\xi_q \le G_j^{\max} \right) \ge 1-\epsilon \tag{11}$$

$$\Pr\left(-G_j - \sum_q Y_{G,j,q}\xi_q \le -G_j^{\min} \right) \ge 1-\epsilon \tag{12}$$

$$\Pr\left(\begin{array}{c} \sum_j I_{G,s,j}\left(G_j + \sum_q Y_{G,j,q}\xi_q \right) \\ -\sum_r I_{GP,s,r}\left(X_r + \sum_q Y_{X,r,q}\xi_q \right) - \sum_k I_{GD,s,k} D_{G,k} \le F_{G,s} \end{array} \right) \ge 1-\epsilon \tag{13}$$

$$\Pr\left(\begin{array}{c} -\sum_j I_{G,s,j}\left(G_j + \sum_q Y_{G,j,q}\xi_q \right) \\ +\sum_r I_{GP,s,r}\left(X_r + \sum_q Y_{X,r,q}\xi_q \right) + \sum_k I_{GD,s,k} D_{G,k} \le F_{G,s} \end{array} \right) \ge 1-\epsilon \tag{14}$$

The supply limits of gas wells are given in (9). The balance of total gas supply and demand is satisfied to (10). The chance constraints of gas generation and pipeline flows are shown in (11), (12), (13), (14).

2.2.4. Linkage between electricity system and gas system

The GNUs link the electricity and gas systems. The relationship between the gas flow consumed and electrical power generated is expressed as

$$\sum_i \Phi_{r,i} P_i = X_r \tag{15}$$

$$\sum_i \Phi_{r,i} Y_{P,i,q} = Y_{X,r,q} \tag{16}$$

In (15), the energy conversion from NGUs gas demand to electrical power output is expressed through Φ. In the same way, the conversion from gas consumption participation factor to power output participation factor is given in (16).

2.3. Solution approach

The model built in Section 2.2 is reformulated to a solvable form according to different distribution assumptions of the forecast errors ξ.

2.3.1. Reformulation of chance constraints

The chance constraints (5), (6), (7) (8) and (11), (12), (13) (14) can be expressed as a general form:

$$\Pr\left(\sum_q A_{h,q}\xi_q \le B_h \right) \ge 1-\epsilon \tag{17}$$

where $A_{h,q}$ is the part related to uncertain data ξ_q and B_h denotes the determinate part in the hth chance constraint. For different distribution assumptions, this section gives the corresponding form after reformulation of chance constraints.

1) The chance constraints which are robust to all possible distributions are analyzed. The inaccessible distribution of forecast error ξ is included in an empirical probability distribution function which is denoted as $F(\xi)$. According to limited historical data, the empirical mean value $\hat{\mu}$ and empirical covariance matrix $\hat{\Sigma}$ are obtained. There is a critical presumption that the empirical mean value and covariance matrix are identical to the real mean value and covariance matrix of ξ. Then the ambiguity set can be designed as [25]:

$$\mathcal{D} = \left\{ F(\xi) : \begin{array}{c} \int_{\xi} F(\xi)\mathrm{d}\xi = 1 \\ \mathbb{E}[\xi] - \hat{\mu} = 0 \\ \mathbb{E}[(\xi-\hat{\mu})(\xi-\hat{\mu})^{\mathrm{T}}] = \hat{\Sigma} \end{array} \right\} \tag{18}$$

The distributionally robust chance-constrained programming problem requires that (17) should be satisfied under all possible distribution in $\mathcal{D}$. Therefore, (17) is converted into

$$\min_{F(\xi)\in\mathcal{D}} \Pr\left(\sum_q A_{h,q}\xi_q \le B_h \right) \ge 1-\epsilon \tag{19}$$

Further, the chance constraints of the IEGS considering wind uncertainty is equivalent to second-order conic constraints [26]. For $\forall \epsilon \in (0,1)$, the chance constraints are formulated as:

$$\mathcal{K}_\varepsilon \sqrt{A_h \hat{\Sigma} A_h^{\mathrm{T}}} \leqslant \left(B_h - \hat{\mu}^{\mathrm{T}} A_h^{\mathrm{T}} \right) \tag{20}$$

where $\mathcal{K}_\varepsilon$ is determined by:

$$\mathcal{K}_\varepsilon = \sqrt{\frac{1-\epsilon}{\epsilon}} \tag{21}$$

2) If the distribution $F(\xi)$ is symmetrical, $\mathcal{K}_\varepsilon$ can be computed by Ref. [27]:

$$\mathcal{K}_\varepsilon = \sqrt{\frac{1}{2\epsilon}} \tag{22}$$

3) If the forecast error of wind generation is subject to the Gaussian distribution, the value of $\mathcal{K}_\varepsilon$ is given by Ref. [28]:

$$\mathcal{K}_\varepsilon = \psi^{-1}(1-\varepsilon) \tag{23}$$

where ψ is the cumulative distribution function of the Gaussian distribution.

For a specific ε, $\mathcal{K}_\varepsilon$ for Gaussian distribution is the highest. And the value of $\mathcal{K}_\varepsilon$ in the distributionally-robust case is the lowest, which means that optimization solutions from DRO are the most

robust and leads to the highest economic cost.

2.3.2. Reformulation of objective function

In this paper, the worst-case expected cost under all distribution in $\mathcal{D}$ is minimized. The original objective (2) is equivalent to

$$\begin{aligned}
&\min_{P,Y_P,G,Y_G}\max_{F(\xi)\in\mathcal{D}} E^{F(\xi)}(C_P(P,Y_P,\xi)+C_G(G,Y_G,\xi))\\
&\quad=\min_{P,Y_P,G,Y_G}\mathcal{C}_P(P,Y_P)+\mathcal{C}_G(G,Y_G)\\
&\mathcal{C}_P(P,Y_P)=tr\left(\begin{bmatrix}P\\Y_P\end{bmatrix}diag(c_{P,2})[P\quad Y_P]\begin{bmatrix}1&0\\0&\hat{\Sigma}\end{bmatrix}\right)\\
&\quad+{c_{P,1}}^{\mathrm{T}}[P\quad Y_P]\begin{bmatrix}1\\\hat{\mu}\end{bmatrix}+\sum c_{P,0}\\
&\mathcal{C}_G(G,Y_G)=tr\left(\begin{bmatrix}G\\Y_G\end{bmatrix}diag(c_{G,2})[G\quad Y_G]\begin{bmatrix}1&0\\0&\hat{\Sigma}\end{bmatrix}\right)\\
&\quad+{c_{G,1}}^{\mathrm{T}}[G\quad Y_G]\begin{bmatrix}1\\\hat{\mu}\end{bmatrix}+\sum c_{G,0}
\end{aligned}\tag{24}$$

$$\begin{aligned}
\Sigma&=\int_{\xi}\xi\xi^{\mathrm{T}}F(\xi)\mathrm{d}\xi-\left(\int_{\xi}\xi F(\xi)\mathrm{d}\xi\right)\left(\int_{\xi}\xi F(\xi)\mathrm{d}\xi\right)^{\mathrm{T}}\\
&=\lim_{n_W\to\infty}\frac{1}{n_W}\sum_q^{n_W}\xi_q{\xi_q}^{\mathrm{T}}-\left(\lim_{n_W\to\infty}\frac{1}{n_W}\sum_q^{n_W}\xi_q\right)\left(\lim_{n_W\to\infty}\frac{1}{n_W}\sum_q^{n_W}\xi_q\right)^{\mathrm{T}}\\
&\approx\frac{1}{n_W}\sum_q^{n_W}\xi_q{\xi_q}^{\mathrm{T}}-\left(\frac{1}{n_W}\sum_q^{n_W}\xi_q\right)\left(\frac{1}{n_W}\sum_q^{n_W}\xi_q\right)^{\mathrm{T}}=\hat{\Sigma}
\end{aligned}\tag{27}$$

where $\mathcal{C}_P(P,Y_P)$ is the maximum risk cost of the electricity system and $\mathcal{C}_G(G,Y_G)$ is the maximum risk cost of the gas system. $tr(*)$ means the trace of a matrix $*$ and $diag(*)$ denotes a diagonal matrix with elements $*$.

3. Historical data generation based on generative adversarial networks

The feasibility of data enhancement for the DRO is analyzed first. Then a popular and powerful data-driven data generation method namely Wasserstein generative adversarial network is introduced to wind power data generation.

3.1. Estimation improvement of worst-case probability distribution

Historical data of wind generation output greatly affected the decision of DRO [29]. The real probability distribution functions of the wind generation output and the forecast error are denoted as $p_{\tilde{W}}(\tilde{W})$ and $p_{\tilde{\xi}}(\xi)$, respectively. In fact, the forecast output of wind generation is determined in IEGS day-ahead scheduling. Thus, the relationship between the probability distribution of $\tilde{W}$ and that of ξ can be formulated as:

$$p_{\tilde{\xi}}(\xi)=p_{\tilde{W}}(\tilde{W})+C\tag{25}$$

There is only one constant difference C between these two distributions.

The real probability distribution functions $p_{\tilde{W}}(\tilde{W})$ and $p_{\tilde{\xi}}(\xi)$ are uncertain. In the ambiguity set of DRO, $F(\xi)$ should consider all probability distribution functions with attributes of the identical $\hat{\mu}$ and $\hat{\Sigma}$, which includes $p_{\tilde{\xi}}(\xi)$. However, the ambiguity set designed in (18) requires that the real mean value and covariance matrix of ξ can be estimated by the empirical mean value and covariance matrix. Otherwise, the worst-case probability distribution is figured inaccurately and fail to be robust to the real probability distribution $p_{\tilde{\xi}}(\xi)$. Theoretically, the mean value and covariance matrix can be calculated by

$$\mu=\int_{\xi}\xi F(\xi)\mathrm{d}\xi=\lim_{n_W\to\infty}\frac{1}{n_W}\sum_q^{n_W}\xi_q\approx\frac{1}{n_W}\sum_q^{n_W}\xi_q=\hat{\mu}\tag{26}$$

The more effective forecast error data, the closer the empirical mean value $\hat{\mu}$ is to μ. In the same way, the empirical covariance matrix $\hat{\Sigma}$ is accurate when sufficient historical data are acquired. The estimation of worst-case probability distribution depends on reliable $\hat{\mu}$ and $\hat{\Sigma}$.

However, the economic dispatch problem of IEGS with the wind uncertainty have the small-sample attribute, i.e., only limited historical wind output data are available [30]. The forecast error data ξ are derived from output data of wind generation $\tilde{W}$. According to (1), limited output data of wind generation are equivalent to insufficient forecast error data. In this case, the distribution solved by DRO are probably not the worst when data corresponding to extreme events are missing. Besides, the robustness of DRO solutions is fixed at a given confidence level. In some cases, a more robust optimization result is required with an acceptable dispatch cost.

Data generation is able to provide realistic data for DRO programming. GAN is a kind of data-driven generative model, which learns the distribution of the input data without expert knowledge and generates new data easily [19]. Specifically, consider a set of historical data for a group of wind generation at n_W sites. Each sample ξ is a vector of historical data indexed by wind generation units. Obviously, there is spatial correlation among different outputs of wind generation. The objective of data generation is to train

a generative model based on GANs by using historical wind generation forecast error $\{\tilde{W}^p\}_{p=1}^N$ as the training set. The generative model is able to learn the implicit expression of $p_{\tilde{W}}(\tilde{W})$ automatically. Moreover, $p_{\tilde{\xi}}(\tilde{\xi})$ can be inferred from $p_{\tilde{W}}(\tilde{W})$ through (25). Then artificial data of forecast error are sampled from the trained generative model.

Data generation by GAN can enrich the diversity of wind power data for estimation of $\hat{\mu}$ and $\hat{\Sigma}$. Derived from the trained GAN, the artificial data of wind power are subject to the real probability distribution $p_{\tilde{W}}(\tilde{W})$. They include missing data of critical events and extreme but reasonable samples of wind power. 1) Through more wind power samples, the estimated $\hat{\mu}$ and $\hat{\Sigma}$ will be closer to μ and Σ in (26), (27). The effect of generated data is illustrated in Fig. 2. More reliable statistical indicators $\hat{\mu}$ and $\hat{\Sigma}$ are beneficial to the estimation of the worst-case probability distribution of DRO. 2) Moreover, lack of historical data means that the diversity of wind power data is insufficient and the empirical covariance matrix $\hat{\Sigma}$ is relatively small. However, the wind power data in real wind farms are various and possess great randomness. With effective artificial wind power data, the proportion of data that corresponds to bad-case scenarios rises to the level in the real world, which leads to a more conservative and robust estimation of $\hat{\Sigma}$. Overall, given the more reliable empirical mean value and covariance matrix, the worst-case probability distribution is improved and robustness of DRO solutions is further enhanced.

Additionally, the satisfaction of chance constraints in DRO under a low-risk level requires a great number of samples to guarantee that a solution of the approximate problem is feasible in the original chance constrained program. The required number of wind power is mainly dependent on the number of decision variables, the risk level, and the number of support constraints [31]. When large-scale IEGS are considered, relative constraints and variables increase exponentially [32]. The number of accessible historical data usually fails to match the required number of wind power data. The proposed data-enhanced distributionally robust optimization is able to easily expand the number of wind power samples and improve the estimation of the ambiguity set, as shown in Fig. 3. By generating sufficient realistic samples, the chance constraints of a fixed risk level in DRO can be satisqfied with great confidence, which brings about a more robust economic dispatch of IEGS.

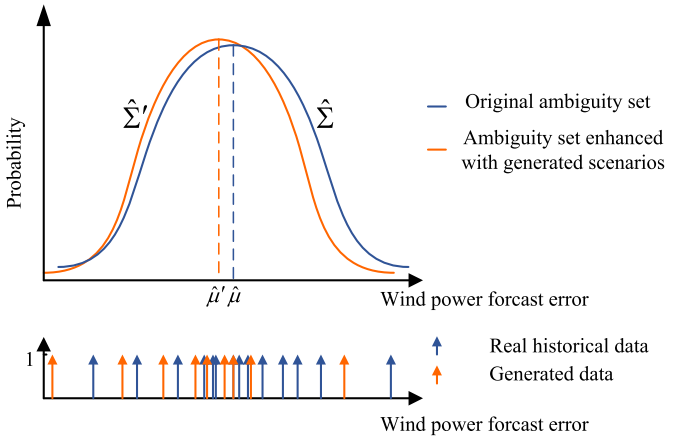

Fig. 2. The worst-case probability distribution estimated on limited historical data and that with generated data added.

3.2. Wasserstein generative adversarial networks

Based on the binomial zero-sum game theory, the generative adversarial networks are powerful for data generation tasks [33]. The basic structure consists of two neural networks: a generator H^{Gen} and a discriminator H^{Dis}, as shown in Fig. 3. Built on the neural networks, GAN is capable of approximate almost any probability distribution function and easily sampled, which is suitable to grasp the connection among wind farms for data generation in DRO programming.

Suppose a simple probability distribution $z \sim p_z$ is known and random noise vectors $\{z\}$ are easily accessed. The generator tends to generate data that look similar to original data from random noise vectors $\{z\}$. In other words, it builds a mapping from the prior distribution $p_z(z)$ to the wind power data distribution $p_{\tilde{W}}(\tilde{W})$. The loss term of the generator H^{Gen} can be expressed as:

$$L_{H^{Gen}} = \mathbb{E}_{z\sim p_z(z)}[\log\left(1 - H^{Dis}(H^{Gen}(z))\right] \tag{28}$$

On the contrary, the discriminator takes both original data $\{\tilde{W}\}$ and generated data $\{H^{Gen}(z)\}$ as input and derives probabilities that the data are recognized as real data. Specifically, the loss function of the discriminator H^{Dis} is derived from cross entropy, which is utilized to measure the similarity between the real wind power data distribution $p_{\tilde{W}}(\tilde{W})$ and the implicit distribution learned by H^{Gen}. The discriminator is expected to recognize historical wind power data $\tilde{W}$ and generated samples $H^{Gen}(z)$ as real and false samples respectively, which corresponds to the cross entropy of binary classification. The loss term of the discriminator H^{Dis} is written as followed:

$$L_{H^{Dis}} = -\mathbb{E}_{\tilde{W}\sim p_{\tilde{W}}(\tilde{W})}[\log H^{Dis}(\tilde{W})] - \mathbb{E}_{z\sim p_z(z)}[\log\left(1 - H^{Dis}(H^{Gen}(z))\right] \tag{29}$$

The competitive idea leads to better optimization of the H^{Gen} and H^{Dis} and forces the generated data to be indistinguishable from original data. The adversarial mechanism is reflected by the adversarial loss of H^{Gen} and H^{Dis}, which is defined as:

$$\begin{aligned}\min_{H^{Gen}} \max_{H^{Dis}} L_{GAN}(H^{Gen}, H^{Dis}) = \mathbb{E}_{\tilde{W}\sim p_{\tilde{W}}(\tilde{W})}[\log H^{Dis}(\tilde{W})] \\ + \mathbb{E}_{z\sim p_z(z)}[\log\left(1 - H^{Dis}(H^{Gen}(z))\right]\end{aligned} \tag{30}$$

where $\mathbb{E}$ denotes the expectation of data, $\sim$ is the symbol of data subject to a distribution, $H^{Dis}(\tilde{W})$ and $H^{Dis}(H^{Gen}(z))$ represent the probability that $\tilde{W}$ and generated data $H^{Gen}(z)$ are regarded as real data by the discriminator, respectively. In the beginning, the data generated by H^{Gen} is distinguishable from original data, which is recognized easily by H^{Dis}. With the increase of training steps, the generator is capable of generating data that the discriminator recognizes as true. Meanwhile, H^{Dis} develops itself with training by distinguishing these newly generated samples from H^{Gen}. In the end, the generator can generate data that are subject to $p_{\tilde{W}}(\tilde{W})$, while the discriminator fails to distinguish $H^{Gen}(z)$ from $p_{\tilde{W}}(\tilde{W})$.

The optimization of the original GAN in (30) is equivalent to minimizing an unreasonable distance measure, which leads to two problems: one is gradient instability, the other is collapse mode, i.e., lack of diversity [34]. Besides, the original GAN has a problem of convergence, because it is unable to know whether the training has finished. Therefore, the Wasserstein distance is introduced into GAN to solve the above problems [35]. The objective is expressed by:

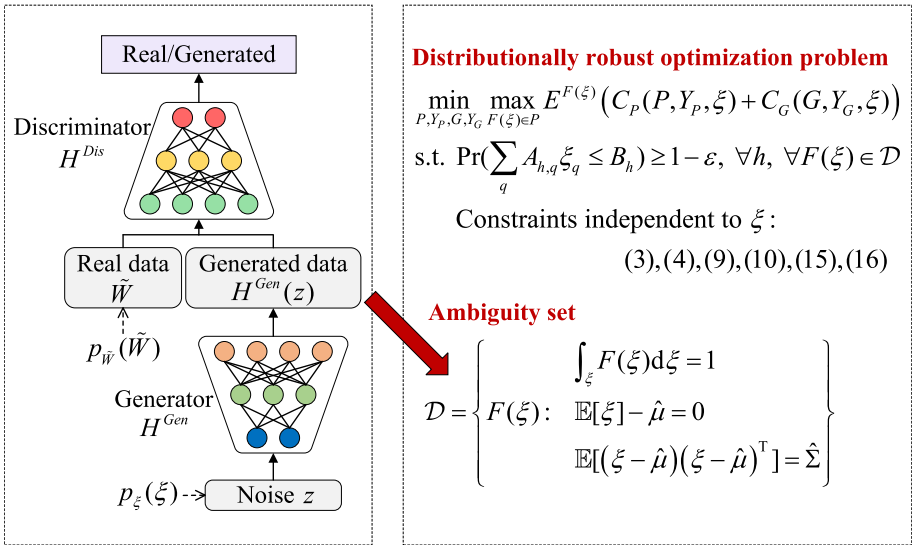

Fig. 3. Architecture of the proposed data-enhanced distributionally robust optimization, where GAN is composed of the generator network H^{Gen} and discriminator network H^{Dis}. The wind power data generated by GAN and original historical data are used to build the ambiguity set in DRO.

$$\min_{H^{Gen}} \max_{H^{Dis}} L_{WGAN}(H^{Gen},H^{Dis}) = \mathbb{E}_{\tilde{W}\sim p_{\tilde{W}}(\tilde{W})}[H^{Dis}(\tilde{W})] + \mathbb{E}_{z\sim p_z(z)}[H^{Dis}\left(H^{Gen}(z)\right)], \tag{31}$$

The exact kind of GAN used to generate wind power data is WGAN. The WGAN in (31) is better than original GAN in (30) at data generation. On the one hand, owing to its superior smoothing property compared with the Jensen-Shannon divergence in the original GAN, the Wasserstein distance solves the problem of unstable training and provides a reliable index of training process. On the other hand, the Wasserstein distance is able to measure the costs needed to build the mapping from the probability distribution $p_z(z)$ to $p_{\tilde{W}}(\tilde{W})$. The optimization objective of WGAN tends to directly minimize the distance between $H^{Gen}(z)$ and $p_{\tilde{W}}(\tilde{W})$. Comparatively, the original GAN with the Jensen-Shannon divergence prefer to guide the generator to generate a single pattern of power profile of highest probability. Therefore, WGAN has better performance in generating diverse extreme events that follows the probability distribution $p_{\tilde{W}}(\tilde{W})$. Data sampled from $p_{\tilde{W}}(\tilde{W})$ are utilized to improve the estimation of worst-case probability distribution.

In order to train the WGAN based on limited historical data, the generator network and the discriminator network use a shallow structure. The data amount may be not enough for the ambiguity set estimation but adequate for the WGAN training. The rough generator is effective enough for robustness enhancement of DRO. Generally, even though the explicit form of $p_{\xi}(\xi)$ is unknown, the result of DRO programming is robust to all possible probability distributions thus what exact $p_{\xi}(\xi)$ is makes no difference. Through WGAN, sampling from $p_{\tilde{W}}(\tilde{W})$ is replaced by sampling from a known p_z and going through the generator H^{Gen}. The distribution of wind generation output $p_{\tilde{W}}(\tilde{W})$ is learned by the generator, which means the implicit distribution of forecast error $p_{\xi}(\xi)$ is inferred. More forecast error data are exploited without expert experience. Then the generated data can be used to further enhance the robustness of DRO programming, as presented in Fig. 3.

4. Case study

First, the basis of the case study including the data and concrete processes is descripted. Then, artificial wind power data generated by WGAN are evaluated. Finally, the optimization performance of data-enhanced DRO is verified in the economic dispatch of the IEGS, including robustness, cost, and operation.

4.1. Data and process description

The proposed data-driven enhanced distributionally robust optimization method is verified on an IEGS consisting of the IEEE 24-bus electricity system and a 12-node natural gas system [36]. Six wind generation units are connected at Bus 1, Bus 2, Bus 11, Bus 12, Bus 12, and Bus 16 of the electricity network respectively. The IEGS are shown in Fig. 4. Wind farms with 250 MW installed capacity are assumed. Wind power data from the six real wind farms in the southeastern Australia, 2012–2013 are assigned to IEGS' wind farms [37]. The wind generation data are divided into total 50 data sets $\{\Upsilon_N^m, \Theta_{N'}^m\}_{m=1,\ldots,50}$. For each data set, Υ_N^m represents the training set composed of N independent historical data that are subject to the identical distribution $p_{\tilde{W}}(\tilde{W})$, while $\Theta_{N'}^m$ is the test set consisting of N' historical wind generation data. Without loss of generality, N is set to 200 to simulate the situation of small sample in DRO and $N' = 800$. Total 50 experiments of economic dispatch are conducted respectively corresponding to 50 independent data sets. There are two steps in each experiment.

First, the data generation based on WGAN is perform based on the training set Υ_N^m. The generator H^{Gen} consists of one fully connected multilayer perceptron layer and two 1-D de-convolutional layers to up-sample the noise z. The discriminator H^{Dis} is composed of a convolutional layer and a fully connected multilayer perceptron layer to output the probability whether the sample is real. Additionally, the WGAN is programmed and trained via a widely used platform TensorFlow in Python with two Nvidia RTX GPUs 2080Ti. The trained H^{Gen} can generate auxiliary training set $\hat{\Upsilon}_{N''}^m$ including N'' samples.

Second, the economic dispatch problems of the IEGS considering wind generation uncertainty subject to different distribution hypotheses are solved. The ambiguity set corresponding to data-

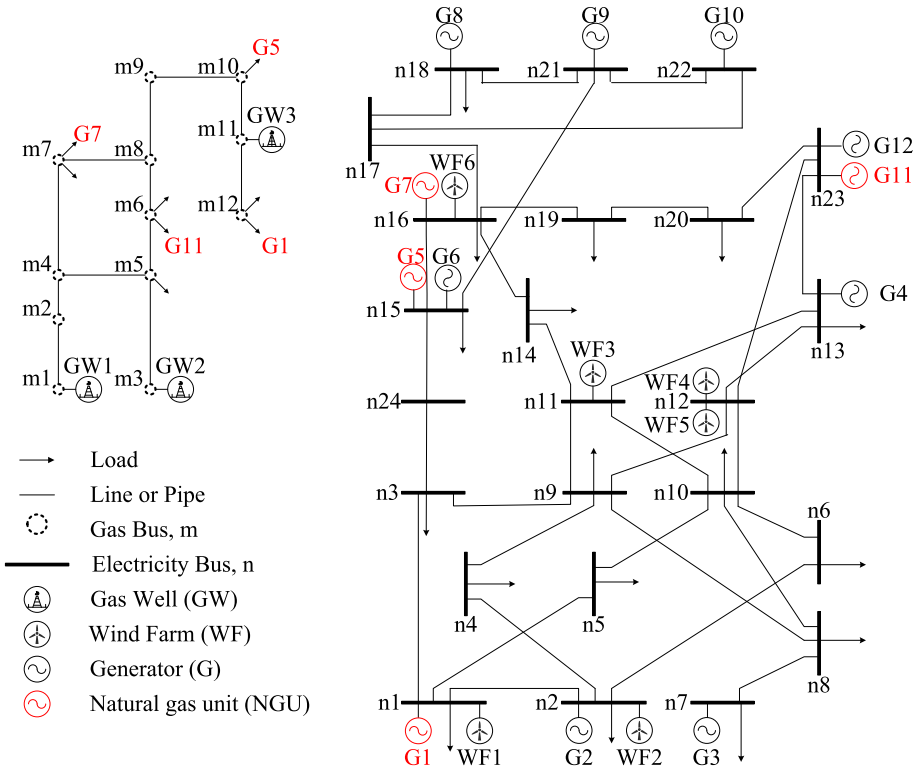

Fig. 4. IEGS of IEEE 24-bus system and 12-node natural gas system.

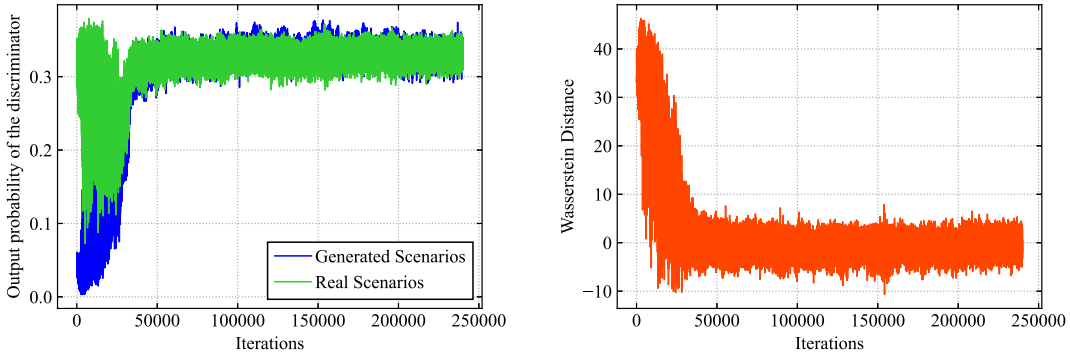

Fig. 5. Training evolution for WGAN. As training moves, the variation of (a) output probability of discriminator fed in generated and real data; (b) WGAN loss, i.e. Wasserstein distance of $H^{Gen}(z)$ to $p_{\tilde{W}}(\tilde{W})$.

driven enhanced DRO estimate the worst-case probability distribution using combined training set $\{\Upsilon'^{m}_{N'}, \hat{\Upsilon}'^{m}_{N''}\}$. Then, the day-ahead scheduling and participation factors are derived through data-driven enhanced DRO. Comparatively, three typical optimization methods are performed utilizing the original training set $\Upsilon'^{m}_{N'}$. Additionally, different confidence levels $1-\epsilon$ are considered. All optimization problems are solved via Gurobi 8.1.0. Finally, the optimization performances are analyzed and compared through the test set $\Theta^{m}_{N'}$, using deterministic optimization, stochastic optimization under Gaussian distribution, symmetrical distributionally robust optimization, distributionally robust optimization, and the proposed data-enhanced distributionally robust optimization.

4.2. Evaluation of generated data

Firstly, the generated data is evaluated by the Wasserstein distance in WGAN training process. Secondly, the generated data can capture the spatial correlation among different wind farms, which is illustrated by comparing the wind power data between the auxiliary training set and the test set. Finally, four common statistical indicators about similarity measurement between two distributions are introduced to evaluate the generated wind power data.

4.2.1. Training process of generative adversarial networks

As shown in Fig. 5, the WGAN reaches the optimal solution and learn the mapping from the known distribution $p_z(z)$ to the target distribution $p_{\tilde{W}}(\tilde{W})$ successfully. At first, the output probability of

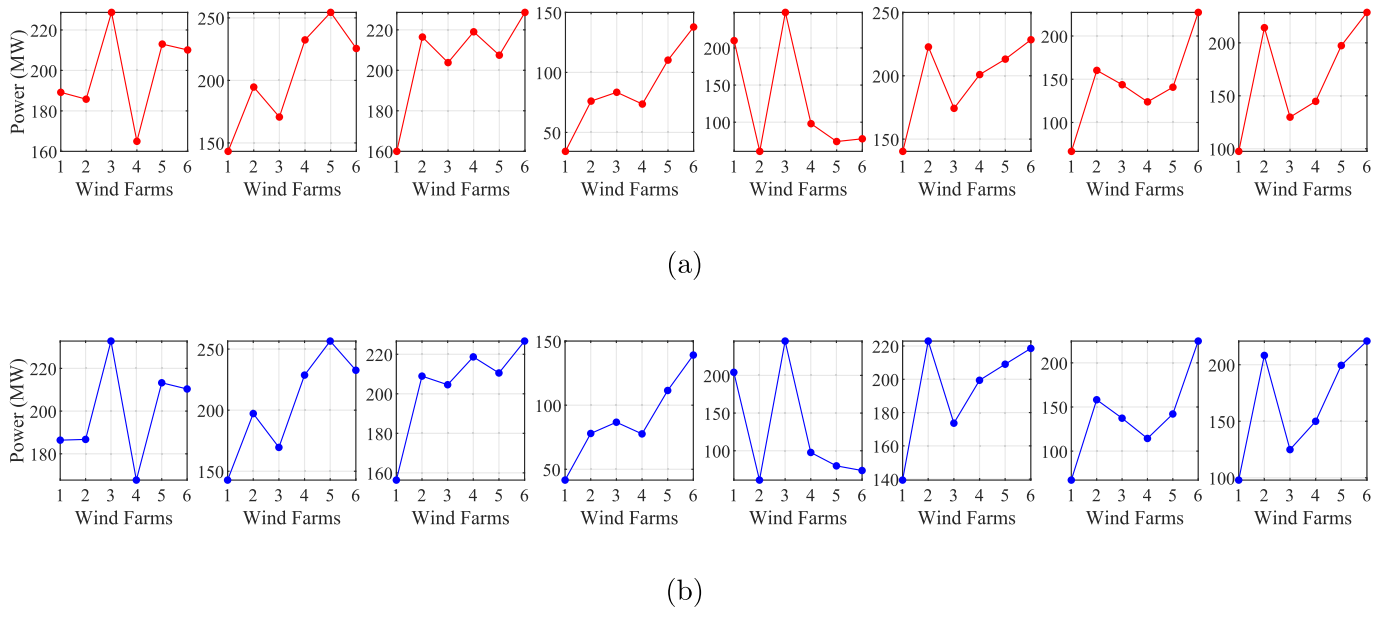

Fig. 6. Comparison of generated data and real data. (a) Generated samples from auxiliary training set $\hat{\Upsilon}^m_{N''}$. (b) Real samples from test set $\Theta^m_{N'}$.

Table 1
Four statistical indicators based on two of three data sets: generated data set $\hat{\Upsilon}^m_{N''}$, training set Υ^m_N, and test set $\Theta^m_{N'}$.

Indicator	$(\hat{\Upsilon}^m_{N''}, \Upsilon^m_N)$	$(\hat{\Upsilon}^m_{N''}, \Theta^m_{N'})$	$(\Theta^m_{N'}, \Upsilon^m_N)$
ED	0.672 4	0.688 5	0.675 6
K-LD	0.112 6	0.112 4	0.111 7
MMD	0.010 5	0.013 8	0.009 5
WD	0.028 8	0.025 0	0.026 2

generated samples $H^{Dis}(H^{Gen}(z))$ significantly lower than that of real data $H^{Dis}(\tilde{W})$, which indicates that the discriminator can easily distinguish the real and generated data. After 50 000 iterations, the training process is stable gradually. The generator is able to generate realistic historical data similar to real samples and the discriminator fail to distinguish two kinds of data. The WGAN objective value in (31) can measure the distance between the real distribution and the learned implicit distribution [35]. When the Wasserstein distance falls to zero, the generator learns the implicit expression of $p_{\tilde{W}}(\tilde{W})$ and the distribution of forecast error $p_{\tilde{\xi}}(\xi)$ can be inferred.

4.2.2. Spatial correlation

Fig. 6 shows the details of the generated data from the auxiliary training set $\hat{\Upsilon}^m_{N''}$ and real data from the test set $\Theta^m_{N'}$. The generated samples in Fig. 6(a) and the real samples in Fig. 6(b) are paired for detailed observation using Euclidean distance-based search. The real samples from the test set are inaccessible in the training of WGAN. However, the generated samples are highly similar to real samples, including the power output of corresponding wind farms and relative relationship of different wind farms. It means that the WGAN is able to capture the characteristics of wind farms and their spatial correlation. The designed simplified GAN is able to be trained based on limited wind power data. Learned by the generator, $p_{\tilde{W}}(\tilde{W})$ can sample reasonable training data to enhance the mode diversity of the training set. In this way, diverse data subject to $p_{\tilde{\xi}}(\xi)$ enable the worst-case probability distribution to consider relatively more extreme events and become more reliable. Then the robustness of DRO method can be further improved.

4.2.3. Statistical evaluation of generated wind power data

Valuable wind power data $\hat{\Upsilon}^m_{N''}$ generated by WGAN should be subject to the real probability distribution function $p_{\tilde{W}}(\tilde{W})$. The distribution $p_{\tilde{W}}(\tilde{W})$ is represented by the training set Υ^m_N and the test set $\Theta^m_{N'}$. Four typical statistical indicators are utilized to measure the similarity among $\hat{\Upsilon}^m_{N''}$, Υ^m_N, and $\Theta^m_{N'}$, including Euclidean distance (ED), Kullback–Leibler divergence (K–LD), maximum mean discrepancy (MMD), and Wasserstein distance (WD).

ED [38] is a common feature to describe the distance between two wind power samples and the average ED can be used to measure the distance between two data sets. Based on entropy theory, K–LD [39] is a frequently-used indicator to evaluate the discrepancy between two probability distributions. MMD [40] calculates two mean values of samples in respective data sets after being mapped by a specific function and then derives the maximum difference between two mean values by finding the optimal function in Reproducing Kernel Hilbert Space. The advantage of WD [35] over many other divergences is that Wasserstein distance can reflect the distance between two distributions properly even if the two distributions do not overlap.

The four statistical features are calculated between $\hat{\Upsilon}^m_{N''}$ and Υ^m_N, $\hat{\Upsilon}^m_{N''}$ and $\Theta^m_{N'}$, $\Theta^m_{N'}$ and Υ^m_N. As listed in Table 1, the discrepancies between generated wind power data sets and real data sets are close to that between training sets and test sets. The generated data are similar not only to wind power data in the training set but also to samples in the test set, which are inaccessible in the training process of WGAN. This verifies that the generated wind power data are reliable for IEGS purposes.

4.3. Economic dispatch

In this section, different optimization methods are carried out to address the economic dispatch problem of the IEGS. Specifically, the data-enhanced distributionally robust optimization (DEDRO) utilizes extra $N'' = 120$ generated data to improve the estimation of worst-case probability distribution in the ambiguity set. Then the proposed method is compared with following four methods. In addition, all parameters of the optimization are similar in

comparative experiments.

1) Deterministic optimization: The wind uncertainty is not considered in the optimization of the IEGS dispatch. The deterministic model commits dispatchable generation resources including power generation units and gas wells to operate at a lowest cost.
2) Stochastic optimization under Gaussian distribution: The forecast errors of wind generation are assumed to follow the Gaussian distribution. The dispatch of dispatchable power and gas resources are derived.
3) Symmetrical distributionally robust optimization (S-DRO): The dispatchable result of S-DRO method is robust to all symmetrical distribution of forecast errors.
4) Distributionally robust optimization (DRO): Considering the wind uncertainty subject to all possible distributions in a specific ambiguity set, the DRO programming hedges against the worst distribution, which are more conservative method than other three comparative methods. The DRO is the main method for comparison.

4.3.1. Empirical dispatch cost, empirical violation probability, and empirical standard deviation of violation probability

The optimization methods including the proposed data-enhanced DRO and other four comparative methods can be evaluated by three aspects: empirical dispatch cost, empirical violation probability, and empirical standard deviation of the violation probability. After the day-ahead scheduling acquired through the training set, the real output of the power generation units and gas wells are calculated through the test set. Empirical dispatch cost is the average generation cost of the whole IEGS when facing different test samples.

$$C_{emp} = \frac{1}{MN'} \sum_{m=1}^{M} \sum_{n'=1}^{N'} C_{I,m,n'} \tag{32}$$

where $C_{I,m,n'}$ is the IEGS cost of the test sample n' based on decision information in the mth simulation. The empirical violation probability is used to evaluate the robustness of optimization solutions. First, the violation frequency of chance constraints for each test sample is given as

$$V_{m,n'} = \frac{1}{N_I} \sum_{a=1}^{N_I} \mathbb{I}_{m,n',a} \tag{33}$$

where $\mathbb{I}_{m,n',a}$ indicates whether the ath chance constraint is violated under the test scenario n' in the mth simulation. Then the empirical violation probability is estimated by the mean value of the violation frequency of chance constraints under test sets.

$$V_{mean} = \frac{1}{MN'} \sum_{m=1}^{M} \sum_{n'=1}^{N'} V_{m,n'} \tag{34}$$

And it is another aspect of robustness whether the optimization results are centralized or not. The empirical standard deviation of the violation probability can measure the dispersion of optimization solutions, which is estimated by

$$\sigma = \sqrt{\frac{1}{MN'-1} \sum_{m=1}^{M} \sum_{n'=1}^{N'} \left(V_{m,n'} - \frac{1}{MN'} \sum_{m=1}^{M} \sum_{n'=1}^{N'} V_{m,n'} \right)^2} \tag{35}$$

Table 2
Empirical dispatch cost ($) of IEGS using different optimization methods with different confidence level $1-\epsilon$.

	0.5	0.6	0.7	0.8	0.9
Deterministic	60 972	60 972	60 972	60 972	60 972
Gaussian	60 852	61 025	61 082	61 180	61 395
S-DRO	61 251	61 311	61 400	61 559	61 946
DRO	61 251	61 369	61 532	61 803	63 346
DEDRO	**61 265**	**61 388**	**61 570**	**61 857**	**63 727**

Table 2 and Table 3 show the dispatch cost, empirical violation probability and the empirical standard deviation of the violation probability, respectively using different optimization methods. The conservatism of optimization results increases with the confidence level rising. When the confidence level $1-\epsilon$ is higher than 0.9, further reduction of the risk level will lead to a great increase in the operating cost of the IEGS, as presented in Table 2. Under the same confidence level, the proposed DEDRO is the most robust method compared with other four methods. The cost of the IEGS will increase when wind uncertainty is taken into account. The deterministic optimization maintains a 5.21% empirical violation probability and the wind uncertainty challenges the safe operation of the IEGS. Acquired from solutions of the optimization methods considering wind uncertainty, the violation probability reduces significantly at the expense of a slight cost increase. The optimization results of stochastic optimization under Gaussian distribution and S-DRO will be effective if the real wind power probability distribution is close to their distribution assumptions. However, the optimization solutions of stochastic optimization under Gaussian distribution are unstable and its empirical violation probability is up to 10.93% when confidence level is 0.5. The DRO and enhanced DRO are obviously more robust than other two optimization methods considering uncertainty when the confidence level is higher than 0.6.

As shown in Fig. 7, when the confidence level is 0.5, the DEDRO achieves a more than 6% decrease ratio of empirical violation probability at the expense of a less than 0.25% increase ratio in cost. With the growing confidence levels, the dispatch cost increases rapidly and robustness enhancement becomes more and more difficult. However, with the confidence level ranging from 0.5 to 0.8, the proposed DEDRO can maintain an improvement of the IEGS robustness by more than 5% at the cost increase of less than 0.1%, as presented in Fig. 7. The empirical violation probability reduces to approximately 0 when the confidence level is higher than 0.9. Moreover, the empirical standard deviation of the violation probability derived from the proposed DEDRO is lower than that from the DRO in the most confidence levels, as listed in Table 3. This indicates that the proposed DEDRO not only brings the lower violation risk but reduces the fluctuation of optimization results. The proposed DEDRO is able to produce a centralized optimization results with the lower violation probability. This validates that the auxiliary data generated by GAN are effective and the proposed method is a cost-effective way to improve the estimation of the worst-case probability distribution of DRO and enhance the robustness of optimization performance.

4.3.2. Dispatch and operation

The parameters of power generation units and gas wells and their average output in test sets using different optimization methods are listed in Table 4 and Table 5.

The power generation units G1, G5, G8, G9 and the gas well GW2 operate at their maximum capacity in all optimization solutions because of the lower generation costs of these plants. However, the optimization methods considering uncertainty set the corresponding participation factors to 0 to prevent the violation of the upper limits. For example, the boxplot reflects the dispatch results

Table 3
Empirical violation probability of IEGS using different optimization methods with different confidence level $1-\epsilon$.

		0.5	0.6	0.7	0.8	0.9
Deterministic	mean	5.21%	5.21%	5.21%	5.21%	5.21%
	standard deviation	0.025 3	0.025 3	0.025 3	0.025 3	0.025 3
Gaussian	mean	10.93%	1.40%	1.35%	1.42%	0.72%
	standard deviation	0.073 0	0.011 2	0.012 9	0.014 4	0.010 7
S-DRO	mean	1.17%	0.95%	0.70%	0.39%	0.06%
	standard deviation	0.013 4	0.012 2	0.010 6	0.008 0	0.003 1
DRO	mean	1.18%	0.78%	0.44%	0.13%	0.00%
	standard deviation	0.013 4	0.011 0	0.008 5	0.004 5	0.000 7
DEDRO	mean	**1.10%**	**0.74%**	**0.40%**	**0.12%**	**0.00%**
	standard deviation	**0.012 9**	**0.010 7**	**0.008 0**	**0.004 3**	**0.000 8**

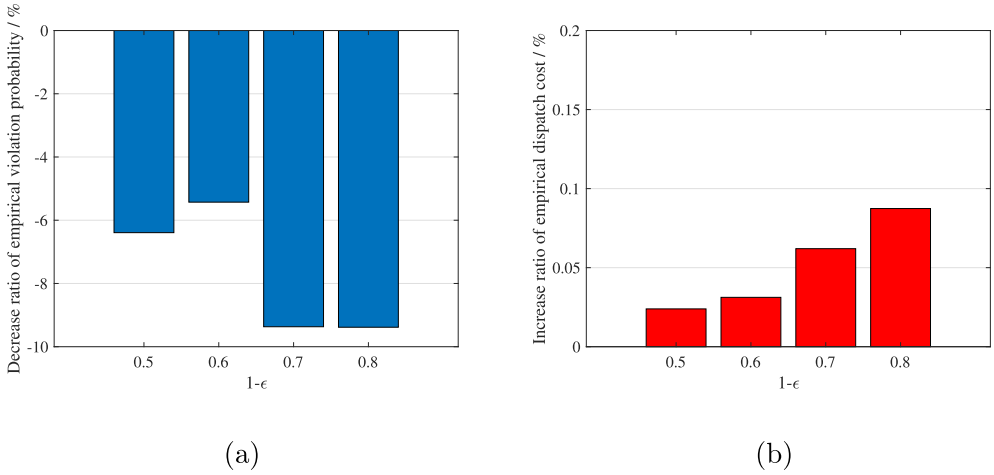

Fig. 7. Compared with DRO optimization results, (a) decrease ratio of DEDRO empirical violation probability and (b) increase ratio of DEDRO empirical dispatch cost.

Table 4
Parameters and average ouput of power generation units (MW) in total 40 000 simulations using different optimization methods when $1-\epsilon=0.7$. G1, G5, G7, G11 are natural gas units.

	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12
Upper limit	152	152	300	591	60	155	155	400	400	300	310	350
Lower limit	0	0	0	0	0	0	0	0	0	0	0	0
Deterministic	152.0	0.3	0.2	389.6	60.5	10.6	155.0	400.6	400.4	5.8	0.0	153.7
Gaussian	152	0.1	0.0	336.5	60.0	85.3	151.4	400.0	399.9	61.7	0.1	81.9
S-DRO	151.9	36	0.0	315.1	60.0	77.1	141.5	400.0	399.9	82.1	0.1	65.1
DRO	151.9	36.1	0.0	315.0	60.0	76.7	141.4	400.0	399.9	82.0	0.1	65.5
DEDRO	**151.9**	**38.9**	**0.0**	**311.5**	**59.9**	**77.8**	**141.3**	**400**	**399.9**	**83.3**	**0.1**	**63.9**

Table 5
Parameters and average ouput of three gas wells (kcf/h) in total 40 000 simulations using different optimization methods when $1-\epsilon=0.7$.

	GW1	GW2	GW3
Upper limit	3500.0	2500.0	7500.0
Lower limit	0	0	0
Deterministic	2001.0	2500.9	2948.9
Gaussian	1999.7	2499.6	2921.7
S-DRO	1999.7	2499.7	2848.2
DRO	1999.6	2499.6	2847.7
DEDRO	1999.6	2499.5	2847.2

of the gas well GW2 for test scenarios, as shown in Fig. 8. In deterministic optimization, the output of the gas well GW2 is usually out of the maximum gas generation limit 2500 kcf/h without considering the wind uncertainty. Other optimization methods considering the wind uncertainty are able to prevent the violation. The operation of power generation units G1, G5, G8, G9 is similar.

Other power and gas generation scheduling changes when considering the wind uncertainty. Using the proposed DEDRO, the energy exchange between the electricity system and gas system decreases slightly, as shown in the output of NGUs G5, G7 of Table 3. It can improve the tight limits of power flows. Power generation units G4, G7, G12 decrease their output considering the wind uncertainty. In DEDRO, the power energy produced by G4, G7, G12 is the least among the output from the other comparative methods, as listed in Table 4. As shown in Table 6, the times of violating the upper limit of G4, G7, G12 decrease about 5% than those in DRO. Meanwhile, the power outputs of G2, G6, G10 increase to satisfy the load, which leads to higher generation cost. Even though the generation costs of G2, G6, G10 are higher, the rise of their power generation helps to decrease the times of violating the lower limits, as listed in Table 6.

The boxplots of the power generation unit G7 with an upper

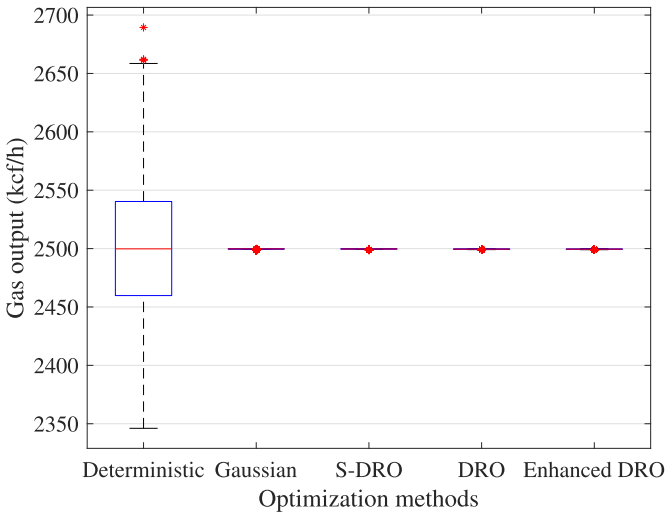

Fig. 8. Output of gas well GW2 when $1-\epsilon=0.7$.

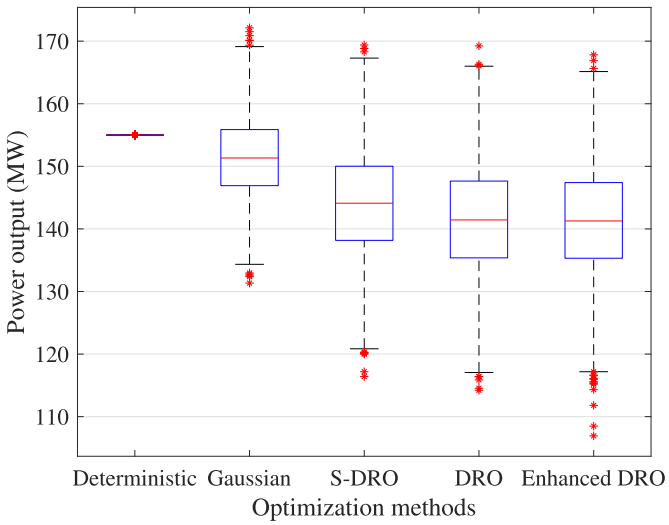

Fig. 9. Output of power generation unit G7 when $1-\epsilon=0.7$.

limit of 155 MW is presented in Fig. 9. And Fig. 10 reflects the power line operation of the IEGS. It can be observed that the number of exceeding limits using the DEDRO obviously decreases and the optimization results from the DEDRO are more centralized. In general, compared with the DRO, G2, G6, G10 further increase the power output to reserve some room to avoid the violation of the G7 generation limit and relative transmission limits of G4 and G12 using the data-enhanced DRO. Therefore, GAN can generate effective historical data of wind uncertainty, which enriches diversity of bad cases in optimization. Equipped with the more reliable mean value and covariance matrix, the proposed data-enhanced DRO can further avoid the violation of the chance constraints and maintain the IEGS security considering the wind uncertainty.

4.3.3. Robustness-controllable through different number of generated data

Different from robust optimization and DRO, the proposed enhanced DRO can control the robustness of optimization solutions by adding different amounts of data. In some cases, the goal is to achieve acceptable robustness of IEGS at an acceptable cost. The proposed method is able to provide a cost-effective method to adjust the robustness of the IEGS dispatch through generated data of the GAN.

In order to investigate the impact of the auxiliary data number in DEDRO, the combined training sets $\{\Upsilon_N^m, \hat{\Upsilon}_{N''}^m\}$ with different data number in auxiliary training set are used to evaluate the optimization performance. It is considered in experiments that the auxiliary data numbers are 0, 40, 80, and 120, respectively. Under different confidence levels, the impact of the auxiliary data number N'' on empirical violation probability and dispatch cost is shown in Fig. 11. At the identical confidence level, the violation probability of the IEGS decreases with the auxiliary data number adding. When the confidence level is lower than 0.9, the robustness can be enhanced effectively at an almost negligible cost increase. For example, compared with DRO without auxiliary data, the decrease ratios of empirical violation probability using 40, 80, 120 artificial data are 6.38%, 7.82%, and 9.37% respectively and the increase ratios of the dispatch cost are lower than 0.075% at the 0.7 confidence level, as shown in Fig. 7. When the confidence level is higher than 0.9, the violation probability nearly falls to 0. And the cost for improvement of system robustness and safety increase explosively, which encounters the marginal effect. In general, this validates that

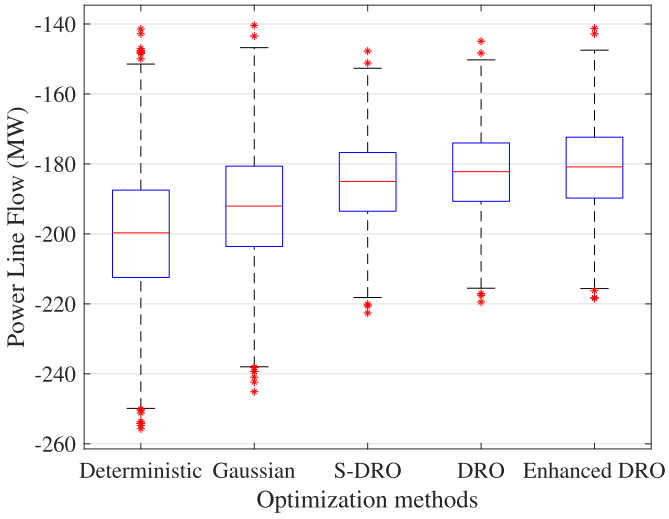

Fig. 10. Power flow of a line when $1-\epsilon=0.7$.

Table 6
Times of violating upper limit ($>P^{max}$) and lower limit ($<P^{min}$) of power generation units in total 40 000 simulations using DRO and DEDRO when $1-\epsilon=0.7$. G1, G5, G7, G11 are natural gas units.

		G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12
DRO	$>P^{max}$	81	0	0	126	0	2639	2601	0	0	0	0	0
	$<P^{min}$	0	2569	361	0	0	2488	0	0	0	2471	0	2349
DEDRO	$>P^{max}$	**76**	0	0	**120**	0	**2479**	**2434**	0	0	1	0	0
	$<P^{min}$	0	**1963**	**288**	0	0	**2200**	0	0	0	**2224**	0	**2225**

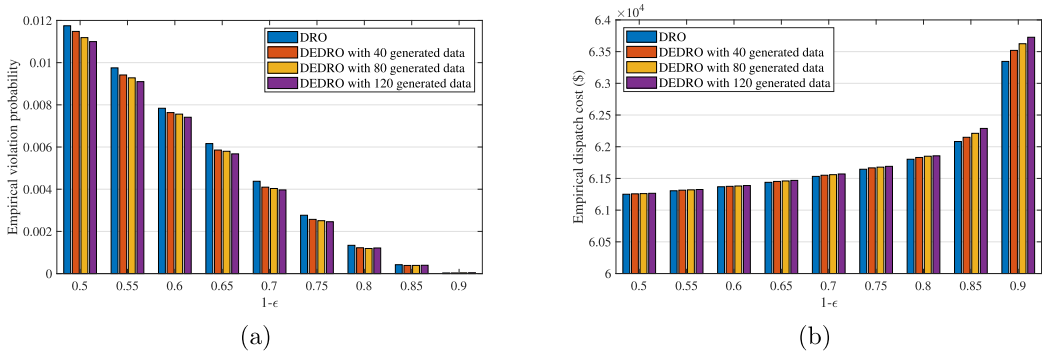

Fig. 11. Using different auxiliary data number N'', (a) the empirical violation probability and (b) empirical dispatch cost of optimization solutions.

the generated data follows the wind power data distribution $p_{\xi}(\xi)$ and possesses diversity of extreme events. And it is a reasonable way to control the robustness of optimization solutions at an acceptable cost through the number of auxiliary data generated by GAN.

5. Conclusion

By integrating the generative adversarial network and distributionally robust optimization, this paper proposed a fully data-driven optimization method to solve the economic dispatch problem of integrated electricity and natural gas systems with wind uncertainty. The method was totally distribution-free and designed to improve the estimation of the worst-case probability distribution when only limited historical data are accessed. Auxiliary data generated by the generative adversarial network can enhance the robustness of optimization solutions of distributionally robust optimization. From the simulation results, three main conclusions were drawn as follows.

1) The generative adversarial network was able to be trained by limited wind generation data automatically and generate data extremely similar to real data in test set.
2) The proposed data-enhanced distributionally robust optimization produced solutions with the minimum violation probability and a slight dispatch cost increase compared with deterministic optimization, stochastic optimization under Gaussian distribution, symmetrical distributionally robust optimization, and distributionally robust optimization.
3) The optimization solutions from the proposed method are more robust than those from distributionally robust optimization. The decline range of the empirical violation probability is even up to 9.38%. And the robustness of the optimization solutions can be controlled through auxiliary data number.

It is a feasible method that utilizing artificial data acquired from the generative adversarial network to further enhance the robustness of integrated electricity and natural gas systems optimization solutions from the distributionally robust optimization. In the future research work, the proposed method will be further compared with other data augmentation methods. And the spatial-temporal correlation of wind uncertainty will be considered in the IEGS economic dispatch.

Credit author statement

Baining Zhao: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - original draft. **Tong Qian:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing-review & editing, Visualization. **Wenhu Tang:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Writing—review& editing, Visualization. **Qiheng Liang:** Methodology, Software, Validation, Visualization, Investigation.

Declaration of competing interest

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

Acknowledge

This paper was supported by the National Key Research and Development Program of China (Grant No. 2018YFE0208400), the Science and Technology Project of State Grid Corporation of China (Key Technologies of Novel Integrated Energy System Considering Cross-border Interconnection).

References

[1] Qian T, Tang W, Wu Q. A fully decentralized dual consensus method for carbon trading power dispatch with wind power. Energy 2020;203:117634. https://doi.org/10.1016/j.energy.2020.117634.
[2] Farrokhifar M, Nie Y, Pozo D. Energy systems planning: a survey on models for integrated power and natural gas networks coordination. Appl Energy 2020;262:114567. https://doi.org/10.1016/j.apenergy.2020.114567.
[3] Jabr RA. Adjustable robust opf with renewable energy sources. IEEE Trans Power Syst 2013;28(4):4742–51. https://doi.org/10.1109/TPWRS.2013.2275013.
[4] He G, Yan H, Chen L, Tao W. Economic dispatch analysis of regional electricity–gas system integrated with distributed gas injection. Energy 2020;201:117512. https://doi.org/10.1016/j.energy.2020.117512.
[5] Li G, Zhang R, Jiang T, Chen H, Bai L, Li X. Security-constrained bi-level economic dispatch model for integrated natural gas and electricity systems considering wind power and power-to-gas process. Appl Energy 2017;194:696–704. https://doi.org/10.1016/j.apenergy.2016.07.077.
[6] Alabdulwahab A, Abusorrah A, Zhang X, Shahidehpour M. Coordination of interdependent natural gas and electricity infrastructures for firming the variability of wind energy in stochastic day-ahead scheduling. IEEE Trans Sustain Energy 2015;6(2):606–15. https://doi.org/10.1109/TSTE.2015.2399855.
[7] Carpentier P, Gohen G, Culioli J-C, Renaud A. Stochastic optimization of unit commitment: a new decomposition framework. IEEE Trans Power Syst 1996;11(2):1067–73. https://doi.org/10.1109/59.496196.
[8] Qadrdan M, Wu J, Jenkins N, Ekanayake J. Operating strategies for a gb integrated gas and electricity network considering the uncertainty in wind power forecasts. IEEE Trans Sustain Energy 2013;5(1):128–38. https://doi.org/10.1109/TSTE.2013.2274818.
[9] Yu H, Rosehart W. An optimal power flow algorithm to achieve robust operation considering load and renewable generation uncertainties. IEEE Trans Power Syst 2012;27(4):1808–17. https://doi.org/10.1109/TPWRS.2012.2194517.
[10] Wei W, Liu F, Mei S, Hou Y. Robust energy and reserve dispatch under variable renewable generation. IEEE Trans Smart Grid 2014;6(1):369–80.

[11] Xiong P, Jirutitijaroen P, Singh C. A distributionally robust optimization model for unit commitment considering uncertain wind power generation. IEEE Trans Power Syst 2016;32(1):39–49. https://doi.org/10.1109/tpwrs.2016.2544795.
[12] Sayed AR, Wang C, Chen S, Shang C, Bi T. Distributionally robust day-ahead operation of power systems with two-stage gas contracting. Energy 2021;231:120840. https://doi.org/10.1016/j.energy.2021.120840.
[13] Jiang R, Guan Y. Data-driven chance constrained stochastic program. Math Program 2016;158(1):291–327. https://doi.org/10.1007/s10107-015-0929-7.
[14] Zhao C, Guan Y. Data-driven stochastic unit commitment for integrating wind generation. IEEE Trans Power Syst 2015;31(4):2587–96. https://doi.org/10.1109/tpwrs.2015.2477311.
[15] Yuan R, Wang B, Mao Z, Watada J. Multi-objective wind power scenario forecasting based on pg-gan. Energy 2021;226:120379. https://doi.org/10.1016/j.energy.2021.120379.
[16] Campi MC, Garatti S, Prandini M. The scenario approach for systems and control design. Annu Rev Control 2009;33(2):149–57. https://doi.org/10.1016/j.arcontrol.2009.07.001.
[17] Duque ÁJ, Castronuovo ED, Sánchez I, Usaola J. Optimal operation of a pumped-storage hydro plant that compensates the imbalances of a wind power producer. Elec Power Syst Res 2011;81(9):1767–77. https://doi.org/10.1016/j.epsr.2011.04.008.
[18] Ma X-Y, Sun Y-Z, Fang H-L. Scenario generation of wind power based on statistical uncertainty and variability. IEEE Trans Sustain Energy 2013;4(4):894–904. https://doi.org/10.1109/tste.2013.2256807.
[19] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. In: Ghahramani Z, Welling M, Cortes C, Lawrence N, Weinberger KQ, editors. Advances in neural information processing systems, vol. 27. Curran Associates, Inc.; 2014.
[20] Chen Y, Wang Y, Kirschen D, Zhang B. Model-free renewable scenario generation using generative adversarial networks. IEEE Trans Power Syst 2018;33(3):3265–75. https://doi.org/10.1109/tpwrs.2018.2794541.
[21] Zhang Y, Ai Q, Xiao F, Hao R, Lu T. Typical wind power scenario generation for multiple wind farms using conditional improved wasserstein generative adversarial network. Int J Electr Power Energy Syst 2020;114:105388. https://doi.org/10.1016/j.ijepes.2019.105388.
[22] Alamo T, Tempo R, Luque A, Ramirez DR. Randomized methods for design of uncertain systems: sample complexity and sequential algorithms. Automatica 2015;52:160–72. https://doi.org/10.1016/j.automatica.2014.11.004.
[23] Ning C, You F. Optimization under uncertainty in the era of big data and deep learning: when machine learning meets mathematical programming. Comput Chem Eng 2019;125:434–48. https://doi.org/10.1016/j.compchemeng.2019.03.034.
[24] Ordoudis C, Pinson P, Kazempour J, Morales González J, Chatzivasileiadis S, Van Hentenryck P, Siddiqui A. Market-based approaches for the coordinated operation of electricity and natural gas systems. Technical University of Denmark; 2018. Ph.D. thesis, Ph. D. Thesis, 2018.
[25] Zhang Y, Shen S, Mathieu JL. Distributionally robust chance-constrained optimal power flow with uncertain renewables and uncertain reserves provided by loads. IEEE Trans Power Syst 2016;32(2):1378–88. https://doi.org/10.1109/tpwrs.2016.2572104.
[26] Wagner MR. Stochastic 0–1 linear programming under limited distributional information. Oper Res Lett 2008;36(2):150–6. https://doi.org/10.1016/j.orl.2007.07.003.
[27] Calafiore GC, El Ghaoui L. On distributionally robust chance-constrained linear programs. J Optim Theor Appl 2006;130(1):1–22. https://doi.org/10.1007/s10107-019-01445-5.
[28] Fang X, Hodge B-M, Jiang H, Zhang Y. Decentralized wind uncertainty management: alternating direction method of multipliers based distributionally-robust chance constrained optimal power flow. Appl Energy 2019;239:938–47. https://doi.org/10.1016/j.apenergy.2019.01.259.
[29] Zymler S, Kuhn D, Rustem B. Distributionally robust joint chance constraints with second-order moment information. Math Program 2013;137(1):167–98. https://doi.org/10.1007/s10107-011-0494-7.
[30] Qu K, Yu T, Pan Z, Zhang X. Point estimate-based stochastic robust dispatch for electricity-gas combined system under wind uncertainty using iterative convex optimization. Energy 2020;211:118986. https://doi.org/10.1016/j.energy.2020.118986.
[31] Calafiore GC, Campi MC. The scenario approach to robust control design. IEEE Trans Automat Control 2006;51(5):742–53. https://doi.org/10.1109/TAC.2006.875041.
[32] Schildbach G, Fagiano L, Morari M. Randomized solutions to convex programs with multiple chance constraints. SIAM J Optim 2013;23(4):2479–501. https://doi.org/10.1137/120878719.
[33] Liang J, Tang W. Sequence generative adversarial networks for wind power scenario generation. IEEE J Sel Area Commun 2019;38(1):110–8. https://doi.org/10.1109/jsac.2019.2952182.
[34] Arjovsky M, Bottou L. Towards principled methods for training generative adversarial networks. 2017. arXiv:1701.04862.
[35] Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks. In: International conference on machine learning. PMLR; 2017. p. 214–23.
[36] Ordoudis C, Pinson P, Morales JM. An integrated market for electricity and natural gas systems with stochastic power producers. Eur J Oper Res 2019;272(2):642–54. https://doi.org/10.1016/j.ejor.2018.06.036.
[37] Dowell J, Pinson P. Very-short-term probabilistic wind power forecasts by sparse vector autoregression. IEEE Trans Smart Grid 2015;7(2):763–70. https://doi.org/10.1109/tsg.2015.2424078.
[38] Liberti L, Lavor C, Maculan N, Mucherino A. Euclidean distance geometry and applications. SIAM Rev 2014;56(1):3–69. https://doi.org/10.1137/120875909.
[39] Basseville M. Divergence measures for statistical data processing—an annotated bibliography. Signal Process 2013;93(4):621–33. https://doi.org/10.1016/j.sigpro.2012.09.003.
[40] Gretton A, Borgwardt KM, Rasch MJ, Schölkopf B, Smola A. A kernel two-sample test. J Mach Learn Res 2012;13(1):723–73.