IET Generation, Transmission & Distribution

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

Data-driven chance-constrained dispatch for integrated power and natural gas systems considering wind power prediction errors

Yuehao Zhao[1] | **Zhiyi Li**[1] | **Ping Ju**[1,2] | **Yue Zhou**[3]

[1]College of Electrical Engineering, Zhejiang University, Hangzhou, China

[2]College of Energy and Electrical Engineering, Hohai University, Nanjing, China

[3]School of Engineering, Cardiff University, Cardiff, UK

Correspondence
Zhiyi Li, College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China.
Email: zhiyi@zju.edu.cn

Abstract

Stochastic wind power prediction errors hurt the normal operation of integrated power and natural gas systems (IPGS). First, the data-driven stochastic chance-constrained programming method is applied to deal with wind power prediction errors, and its probability distribution is accurately fitted by variational Bayesian Gaussian mixture model with massive historical data. In addition, the data-driven chance constraints of tie-line power and reserve capacity of gas turbine are built. Next, to utilize wind power more reasonably, the operational characteristics and optimal commitment of power-to-hydrogen devices are considered and modelled in proposed strategy to reflect the actual situation of IPGS. Then, the original complicated dispatch problem is converted into a tractable second-order cone programming problem via convex relaxation and quantile-based analytical reformulation techniques. Finally, the effectiveness of the proposed strategy is validated by numerical experiments based on a modified IEEE 33-bus system integrated with a 10-node natural gas system and a micro hydrogen system.

1 | INTRODUCTION

To combat global warming and energy shortage, renewable energy is being vigorously developed worldwide [1] to provide people with clean and green electricity. Renewable energy output is stochastic and volatile due to uncertain weather conditions. The electricity generated by renewable energy cannot be fully utilized due to the limited flexibility and reserve capacity of power system, which is often largely curtailed, especially in China.

The utilization rate of renewable energy is increased by the coordinated dispatch and coupled conversion of integrated power and natural gas systems (IPGS) [2], and the curtailed renewable energy can be converted to hydrogen by P2H (power-to-hydrogen) devices. Besides, the optimal dispatch for IPGS will also reduce the total operation costs [3, 4] and increase operational security [5, 6]. Therefore, researching the optimal dispatch strategy for IPGS is gaining increasing interest today. It is crucial to study the optimal dispatch of IPGS, considering wind power uncertainties and the operational characteristics of P2H devices.

The stochastic output power and wind power prediction errors (WPPE) often lead to the frequent and significant fluctuation of tie-line power, affecting the economic operation of the external power grid [7].Meanwhile, an adequate reserve capacity of gas turbine (GT) is required to cope with those stochastic uncertainties of wind power and load [8, 9]. Therefore, the question of how to reduce the fluctuations of tie-line power and determine an appropriate reserve capacity considering wind power uncertainty needs to be further investigated. Several references study WPPE, but the stochastic dispatch for IPGS has not yet been adequately addressed [8, 9]. The chance-constrained stochastic programming (CCSP) is a commonly used method to deal with WPPE. The CCSP method fully exploits the information efficiency of massive historical data, which has been widely used in the operation of IPGS [8]. The economics and the security level are coordinated with an appropriate confidence value in CCSP that indicates the operator's tolerance level for stochastic risks.

The sample-average approximation (SAA) is a common method in CCSP, which is close to the actual situation when its sample number is large enough [10]. However, each

sample in SAA introduces a binary variable. As the sample size increases, the number of binary variables grows rapidly, making this mixed-integer programming problem difficult to solve [10]. Therefore, the SAA is widely believed as time-consuming, inefficient, and even intractable [11, 12] with large samples. It is computationally challenging to solve SAA with multiple integer or discrete variables [10]. Worse still, the accuracy and performance of this method are unsatisfactory when its sample number is not sufficient or finite [13]. In contrast, another widely used method in CCSP is the quantile-based analytical reformulation. Its calculation speed is fast but it usually needs an accurate priori probability distribution of WPPE [8]. However, many references [9, 14, 15] assume that WPPE follows a conventional priori distribution, such as Gaussian distribution. For example, WPPE in IPGS are assumed to follow Gaussian distribution. The CCSP was used to solve this stochastic programming problem in power systems [14] and micro multi-energy systems [16]. This assumption that the probability distribution of WPPE follows Gaussian distribution is too simple and perfect, and it may result in inaccuracy in probability distribution fitting and solution [17]. Gaussian distribution is simple, unimodal, and symmetrical, but WPPE usually shows multi-peak and asymmetry characteristics [18]. As a result, the fitting accuracy of WPPE using Gaussian distribution is unsatisfactory, and it is usually arduous to implement in a real power system [19]. In addition, it should recognize that renewable energy prediction errors usually have unique distribution characteristics in some geographical regions. More-over, a simple distribution cannot fit the real distribution characteristics.

With the wide application of advanced measurement infrastructure in IPGS, vast amounts of data have been accumulated, including historical prediction and actual measurement data of renewable energy. Unfortunately, the information from these vast amounts of data has not been adequately used to fit WPPE at present. Gaussian mixture model (GMM) has been used incrementally in recent years to fit the actual probability distribution of WPPE. Then this fitted probability distribution is introduced as an input parameter for the CCSP [8, 17]. The fitting accuracy strongly depends on the number of Gaussian components, and relatively large fitting errors occur when the number of Gaussian components is inappropriate. Never-the-less, the number of Gaussian components is usually selected by manual observation [8], which easily leads to over-fitting or under-fitting problems. The probabilistic statistical information of historical WPPE data is not fully and appropriately utilized by the traditional GMM method. For this reason, how to accurately fit the probability density and automatically select the reasonable number of Gaussian components in GMM is extremely important. At present, the CCSP with GMM method is mainly used in the probabilistic power flow calculation [18, 20], economic dispatch [17], and unit commitment [8] in traditional power systems. The calculation speed of optimal dispatch with GMM method is fast [8, 17, 18, 20]. Still, it is seldom used in the operation of IPGS. More importantly, its upgrade version-the variational Bayesian Gaussian mixture model (VBGMM)

method is only adopted in the optimal power flow [19], and it is rarer in the operation of IPGS. Compared to traditional power systems, IPGS is a more complex integrated multiple energy systems [3], and its operation considering stochastic WPPE with the data-driven CCSP method, VBGMM, needs more attention.

Another fact that needs to be addressed is that, the operational characteristics and actual operation constraints of P2H device in IPGS needs to get more attention. Refs. [9, 21] neglect operational characteristics and ramping constraints of P2H devices. The models in the above two references are relatively simple, not reflecting their actual situation. There are two types of advanced P2H devices, including PEMEC (proton exchange membrane electrolysis cell) and SOEC (solid oxide electrolysis cell) [3, 22]. The former's start-up/shut-down and ramp speed is rapid, and it has a wide adjustable power range, but its conversion efficiency is relatively low. In contrast, the conversion efficiency of the latter is high, but it has a slow response speed and a narrow adjustable power range. In addition, most research papers do not consider the operational characteristics, optimal commitment, and power allocation of the above two P2H devices. The reliable and economical operation of IPGS will be further ensured if these three factors could be taken into account and modelled reasonably.

Generally, up until now, there are two obvious deficiencies in the existing dispatch strategies for IPGS: (1) Stochastic WPPE has not been effectively addressed, and its potential information of historical data has not been fully utilized, affecting the economic operation of IPGS; (2) The operational characteristics and optimal commitment of P2H devices have been hardly considered and modelled, not reflecting the actual operating situation.

To solve the above two problems, a stochastic data-driven chance-constrained day-ahead dispatch strategy for IPGS considering the operational characteristics of P2H devices is proposed. The main contributions are presented as follows:

1. The data-driven CCSP is proposed to cope with stochastic WPPE, and VBGMM is used to fit the probability distribution of WPPE [23]. Specifically, the tie-line power and GT reserve capacity are modelled as data-driven chance constraints with VBGMM.
2. The operational characteristics and optimal commitment of the two types of P2H devices are also considered and modelled, which can ensure the reliable and economic operation of IPGS.
3. Scheduled wind power (SWP), wind power consumed by P2H devices, purchased power from the external power grid, and the output of GTs are optimally coordinated to support the economic operation of IPGS effectively.

The remainder of this paper is organized as follows. The stochastic optimization for IPGS dispatch is introduced in Section 2. The solution method is presented in Section 3. The case study is shown in Section 4. Then the conclusions are summarized in Section 5.

FIGURE 1 IPGS structure.
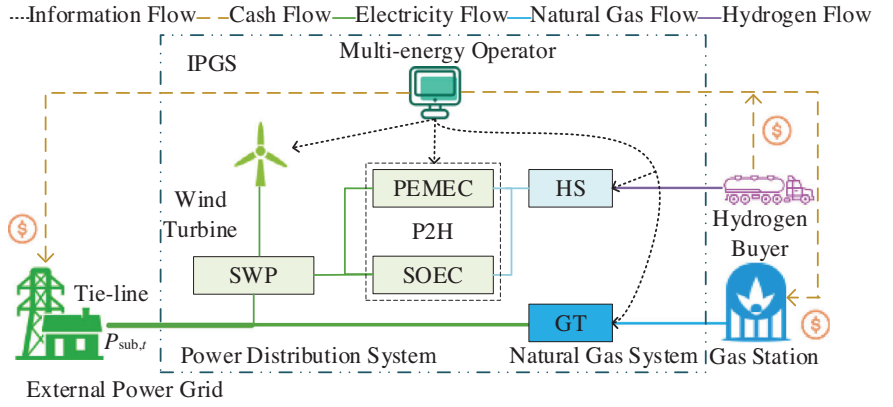
2 | STOCHASTIC OPTIMIZATION FOR IPGS DISPATCH

2.1 | Structure of IPGS

The IPGS here is mainly composed of two subsystems: A natural gas system and a power distribution system that integrates a micro hydrogen system. The joint operation and management of power system and natural gas system in some countries, such as the U.K., is common [24], so IPGS is assumed to be dispatched and managed as a whole by the multi-energy operator here.

There are many conversion facilities to meet the multi-energy demands in IPGS. Specifically, the power demand is met by the output power of GT, the wind turbine, and purchased power from the external power grid. The wind turbine output contains two parts: directly scheduled wind power and wind energy consumed by P2H device. Here, the produced hydrogen is stored in hydrogen storage (HS). The hydrogen stored in HS is profitably sold to an external hydrogen customer. Besides that, the external gas station supplies the consumed natural gas of IPGS. The structure of IPGS is shown in Figure 1.

In this paper: "operational characteristics" refers to the startup/ start-down and ramp speed and operating range of P2H devices. The typical process in P2H is also shown in Figures 1 and 2. Water is electrolyzed to hydrogen and oxygen by PEMEC or SOEC consuming renewable energy.

As shown in Figure 2, high-temperature conditions are essential in SOEC, so it usually takes several hours to startup [3]. Specifically, the water is electrolyzed by the solid oxide at high temperatures, about 800°C, and the conversion efficiency of SOEC improves due to the increased thermodynamic energy and kinetic energy. Therefore, it takes several hours to heat the electrolyzer and accumulate heat to reach such a high temperature. For example, the start-up time of SOEC is about 2 h [25]. Compared to SOEC, the operating and reaction temperature of PEMEC is relatively low, about 80°C, so its start-up time is short and can be neglected. Due this large thermal inertia, SOEC usually has low start-up and shut-down speed and low ramp rate, so it is unreasonable to ignore its operational characteristics. The operational characteristics model of P2H devices is then discussed in more detail in Section 2.3 below.
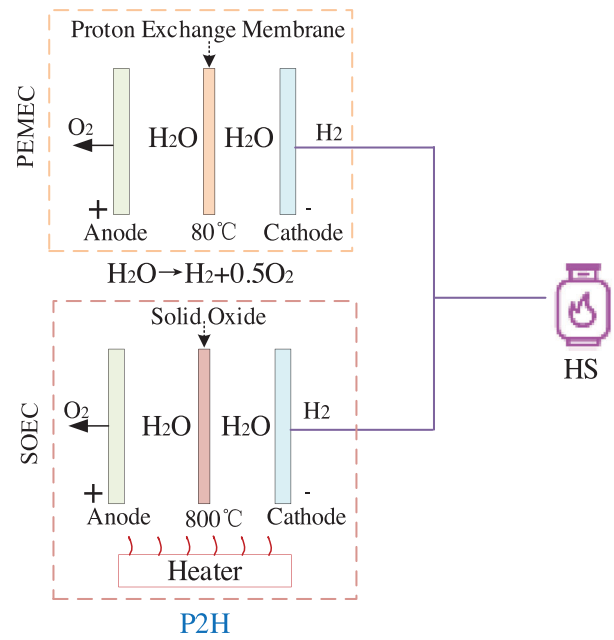
FIGURE 2 P2H structure.

2.2 | Data-driven chance constraints

With the increasing installed capacity of renewable energy in power distribution system, the tie-line power between power distribution system and external power grid often fluctuates randomly due to the stochastic output of distributed renewable energy. This fluctuation affects the security and economic operation of the external power system greatly [7]. To reduce the negative impact of power fluctuation caused by intermittent distributed wind power on the external power grid, the data-driven CCSP is built. It consists of two key elements: Stochastic chance constraints and the data-driven probability distribution fitting method of WPPE.

2.2.1 | Chance constraints of the tie-line and GT

To relieve the power regulation burden of external power grid, the power fluctuations of the tie-line must be limited. Besides, to

ensure users' power security in IPGS, adequate reserve capacity of GT is necessary.

The actual output of distributed wind power is stochastic and can be regarded as the predicted wind power plus stochastic prediction errors, as shown in Eq. (1). Without loss of generality, assume all power/ gas load is deterministic, and the influence of stochastic WPPE on the change of network losses is negligible. Due to the stochastic uncertainty of wind power, the tie-line power is also stochastic, as shown in Eq. (2).

$$\tilde{W}^{a}_{w,t} = W^{f}_{w,t} + \tilde{e}_{w,t} \tag{1}$$

$$\tilde{P}_{\text{sub},t} = P_{\text{sub},t} - \tilde{e}_{w,t} \tag{2}$$

To reduce the adverse impact of the stochastic power fluctuation of the tie-line, its maximum power is modelled as chance constraints, as described in Eq. (3). Specifically, the probability that the actual power of the tie-line $\tilde{P}_{\text{sub},t}$ is less than or equal to maximum permissible value $P^{\max}_{\text{sub}}$ is calculated as the equation on the left-hand side of Eq. (3). The threshold value of the probability of the tie-line is defined by the parameters on the right-hand side of the Eq. (3), which is a predefine constant. Eq. (3) denotes the possibility that the actual power of the tie-line with stochastic WPPE is less than or equal to the maximum permissible value should exceed a high-probability threshold value $1 - \alpha^{\max}_{\text{sub}}$.

$$\Pr(\tilde{P}_{\text{sub},t} \leq P^{\max}_{\text{sub}}) \geq 1 - \alpha^{\max}_{\text{sub}} \tag{3}$$

The directly scheduled wind power by the power distribution systems equals the predicted wind power minus the wind energy consumed by P2H device, as shown in (4).

$$W^{s}_{w,t} = W^{f}_{w,t} - W^{H}_{w,t} \tag{4}$$

where $W^{S}_{w,t}$ means the total wind power that is directly utilized by the power distribution systems, $W^{f}_{w,t}$ is the predicted wind power, and $W^{H}_{w,t}$ is the wind power consumed by P2H device.

To prevent the risk of stochastic WPPE and potential power load fluctuation, a reasonable reserve capacity of GT is needed to ensure the power balance in IPGS, calculated as (5) and (6).

$$\Pr\left(\sum_{g=1}^{N_G} P^{\text{UR}}_{g,t} \geq \sum_{w=1}^{N_W} W^{s}_{w,t} - \sum_{w=1}^{N_W} (\tilde{W}^{a}_{w,t} - W^{H}_{w,t}) + P^{\text{UR,L}}\right) \geq 1 - \alpha_{\text{UR}} \tag{5}$$

$$\Pr\left(\sum_{g=1}^{N_G} P^{\text{DR}}_{g,t} \geq -\sum_{w=1}^{N_W} W^{s}_{w,t} + \sum_{w=1}^{N_W} (\tilde{W}^{a}_{w,t} - W^{H}_{w,t}) + P^{\text{DR,L}}\right) \geq 1 - \alpha_{\text{DR}} \tag{6}$$

The demand for reserve capacity consists of two parts, one for stochastic WPPE and the other for potential load fluctuation $P^{\text{UR,L}}$/ $P^{\text{DR,L}}$.

In Eq. (5), its third variable (total actual stochastic wind power output) minus the fourth variable (the wind power for P2H), $\sum_{w=1}^{N_W} (\tilde{W}^{a}_{w,t} - W^{H}_{w,t})$, denotes the actual stochastic value of wind power directly utilized by the power distribution

systems. Then the second variable $\sum_{w=1}^{N_W} W^{s}_{w,t}$ minus the difference between the third and fourth variables, $\sum_{w=1}^{N_W} W^{s}_{w,t} - \sum_{w=1}^{N_W} (\tilde{W}^{a}_{w,t} - W^{H}_{w,t})$, is the opposite number of stochastic WPPE ($-\tilde{e}_{w,t}$). Similar to Eq. (5), the opposite number of the second variable plus the difference between the third and fourth variable in (6), $-\sum_{w=1}^{N_W} W^{s}_{w,t} + \sum_{w=1}^{N_W} (\tilde{W}^{a}_{w,t} - W^{H}_{w,t})$, equals stochastic WPPE ($\tilde{e}_{w,t}$).

Besides, the reserve capacity for the latter one is predefined as a fixed constant value (3% of the maximum value of load). The threshold value of the probability of the upward/downward reserve capacity is defined by the parameters on the right-hand side of chance constraints (4) and (5). Then the following two equations, (7) and (8), are obtained by substituting (3) to (5) and (6), which are presented as follows:

$$\Pr(\sum_{g=1}^{N_G} P^{\text{UR}}_{g,t} \geq -\tilde{e}_{w,t} + P^{\text{UR,L}}) \geq 1 - \alpha_{\text{UR}} \tag{7}$$

$$\Pr(\sum_{g=1}^{N_G} P^{\text{DR}}_{g,t} \geq \tilde{e}_{w,t} + P^{\text{DR,L}}) \geq 1 - \alpha_{\text{DR}} \tag{8}$$

To eliminate potential adverse impacts from stochastic WPPE and load fluctuation, then Equations (5), (6) and (7), (8) denote the probability that dispatched upward/downward reserve capacity of GT $\sum_{g=1}^{N_G} P^{\text{UR}}_{g,t}$/ $\sum_{g=1}^{N_G} P^{\text{DR}}_{g,t}$ exceeds its demand for reserve capacity should exceed a high-probability threshold value $1 - \alpha_{\text{UR}}$/ $1 - \alpha_{\text{DR}}$.

2.2.2 | Data-driven fitting method for WPPE

As mentioned earlier, massive historical data in IPGS, such as prediction power data and measured generated power data of wind power, has been accumulated. However, the information of those massive data has yet to be sufficiently utilized at present. With the increasing penetration of wind power in power distribution system, the power fluctuations in IPGS also increase. Therefore, it is critical to study the distribution characteristics of WPPE, and an accurate probability distribution of WPPE is an indispensable precondition for the economical and secure operation of IPGS.

Here, the data-driven CCSP based on VBGMM for the operation of IPGS is proposed, increasing the data utilization ratio of wind power with high feasibility and practical value. More importantly, the fitting accuracy of VBGMM for WPPE is higher than Gaussian distribution and traditional GMM method [19]. In theory, an arbitrary probability distribution can be fitted relatively accurately by GMM if it adjusts its parameters, such as the number of components, weight, means, and covariance [8]. It's fitting accuracy depends on the number of Gaussian components involved heavily [19]. However, the traditional GMM method relies on manual observation or prior knowledge set by experts [19] to choose the number of Gaussian components. It is hard to select reasonable parameters [19]. Besides, VBGMM
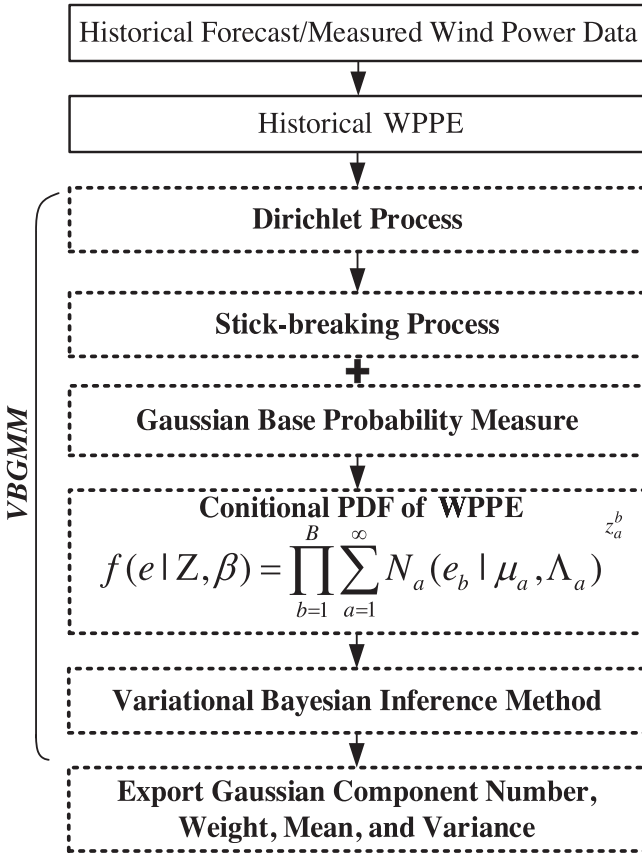
FIGURE 3 Flowchart of VBGMM method.

method is seldom adopted to fit the probability distribution of WPPE in IPGS.

Compared to the above traditional GMM, the number of components of VBGMM can be acquired automatically without prior setting [19]. Therefore, it's fitting accuracy is also better than the traditional GMM, and the symmetrical characteristic of WPPE can also be described [8, 17, 19]. VBGMM is a non-parametric Bayesian model based on Dirichlet process, fully utilizing the probabilistic statistical information of historical prediction data and measured data of wind power [23]. VBGMM can infer the number of Gaussian components of historical wind power data, so the predefined number of Gaussian components by manual observation is no longer needed [26]. Besides, some other vital parameters of VBGMM can also be inferred, including the weight, mean, and variance of each Gaussian component, the number of Gaussian components, and so on. Moreover, the probability density function (PDF) and cumulative distribution function (CDF) of WPPE can be obtained automatically. Then, the fitted PDF and CDF are introduced as the initial input parameter of the CCSP. The flowchart of VBGMM method is presented in Figure 3.

For finite partitions, $\theta_1, \ldots, \theta_i$ in measure space Θ, the distribution G satisfies the following property (9), and then G follows the Dirichlet process, which can be denoted by Eq. (10) [26]. G has two core elements, the concentration parameter ϕ and the base probability measure G_0 [27]. Specifically, Gaussian distri-

bution is adopted as the base probability measure in VBGMM method.

$$G((\theta_1), \ldots, (\theta_i)) \sim Dir(\phi G_0(\theta_1), \ldots, \phi G_0(\theta_i)) \tag{9}$$

$$G \sim DP(\phi, G_0) \tag{10}$$

An infinite number of Gaussian components can be provided to analyze the probability distribution of WPPE in VBGMM [26]. The stick-breaking process is a common method for constructing Dirichlet process. Eqs. (11)–(15) are used to build the stick-breaking process. The categorization weight can be calculated by the random sampling of the base probability measure, and the calculation formula of G is presented in (14). The hidden variables are described in (15) [19].

$$v_a \sim Beta(v_a|1, \phi) \tag{11}$$

$$\omega_a = v_a \prod_{s=1}^{k-1} (1 - v_s) \tag{12}$$

$$\sum_{a=1}^{\infty} \omega_a = 1 \tag{13}$$

$$G = \sum_{a=1}^{\infty} \omega_a \delta_{\theta_a} \tag{14}$$

$$\mho = \{v_a\}_{a=1}^{\infty}, Y = \{\omega_a\}_{a=1}^{\infty} \tag{15}$$

WPPE is regarded as a set of stochastic variables $E = \{e_b\}_{b=1}^{B}$, and the joint PDF of WPPE using VBGMM method is denoted as (16).

$$f(e|\omega, \beta) = \prod_{b=1}^{B} \sum_{a=1}^{\infty} \omega_a N_a(e_b|\mu_a, \Lambda_a) \tag{16}$$

where $\beta = \{\mu_a, \Lambda_a\}_{a=1}^{\infty}$.

The binary indicator variable z_a^b is introduced in (17) to indicate the category to which stochastic variables in the historical data belong. The conditional PDF of WPPE with indicator variable z_a^b is presented as Eq. (17).

$$f(e|Z, \beta) = \prod_{b=1}^{B} \sum_{a=1}^{\infty} N_a(e_b|\mu_a, \Lambda_a)^{z_a^b} \tag{17}$$

where $Z = \{z_a^b\}_{a=1, b=1}^{a=\infty, b=B}$.

Then the variational Bayesian inference method is used to estimate the posterior probability distribution of the hidden variable $\Xi = \{Z, \omega, \mho, Y\}$ [19]. The parameter estimation and fitting of VBGMM with the variational Bayesian inference are implemented based on sklearn. mixture package [23]. The inferred CDF and PDF of WPPE are as follows [17]:

$$f_E(\tilde{e}_t) = \sum_{e=1}^{N_E} \omega_e N_e(\tilde{e}_t|\mu_e, \sum_e) \tag{18}$$

$$F_E(\tilde{e}_t) = \sum_{i=1}^{N_E} \omega_i \Phi_i \left(\tilde{e}_t | \mu_i, \sum_i \right) \tag{19}$$

$$\sum_{i=1}^{N_E} \omega_i = 1, \omega_i \geq 0 \tag{20}$$

In Eqs. (18) and (19), the PDF and CDF of WPPE are composed of the linear combination of the multiple Gaussian components, respectively; In Eq. (20), The sum of the weights of each Gaussian component is 1.

2.3 | Operational characteristics of P2H devices

1. Minimum up time constraints of SOEC

$$\sum_{t=1}^{Q_{S,l}} \left[1 - u_{l,t}^S \right] = 0 , \ \forall t = 1, \dots, Q_{S,l} \tag{21}$$

$$\sum_{k=t}^{t+T_l^{S,U}-1} u_{l,k}^S \geq T_l^{S,U} \left[u_{l,t}^S - u_{l,t-1}^S \right],$$

$$\forall t = Q_{S,l} + 1, \dots, T - T_l^{S,U} + 1 \tag{22}$$

$$\sum_{k=t}^{T} \{ u_{k,l}^S - [u_{l,t}^S - u_{l,t-1}^S] \} \geq 0 ,$$

$$\forall t = T - T_l^{S,U} + 2, \dots, T - 1 \tag{23}$$

$$Q_{S,l} = \min \left\{ T, \left[T_l^{S,U} - U_{l,0}^S \right] u_{l,0}^S \right\} \tag{24}$$

where (21) is used to limit SOEC must be on at the initial time; (22) describes a typical minimum up time constraint at the intermediate time; (23) enforces SOEC must be on at the remaining time of the day; (24) calculates the minimum up time at the initial time.

2. Minimum down time constraints of SOEC

$$\sum_{t=1}^{L_{S,l}} \left[u_{l,t}^S \right] = 0 , \ \forall t = 1, \dots, L_{S,l} \tag{25}$$

$$\sum_{k=t}^{t+T_l^{S,D}-1} u_{l,k}^{\{.\}} \geq T_l^{S,D} \left[u_{l,t-1}^S - u_{l,t}^S \right]$$

$$\forall t = L_{S,l} + 1, \dots, T - T_l^{S,D} + 1 \tag{26}$$

$$\sum_{k=t}^{T} \{ 1 - u_{k,l}^S - [u_{l,t-1}^S - u_{l,t}^S] \} \geq 0 ,$$

$$\forall t = T - T_l^{S,D} + 2, \dots, T - 1 \tag{27}$$

$$L_{S,l} = \min \left\{ T, \left[T_l^{S,D} - V_{l,0}^S \right] [1 - u_{l,0}^S] \right\} \tag{28}$$

where (25) is used to limit SOEC must be off at the initial time; (26) describes a typical minimum down time constraint at the intermediate time; (27) enforces SOEC must be off at the remaining time of the day; (28) calculates the minimum off time at the initial time.

3. Ramp constraints of SOEC

$$P_{l,t}^S - P_{l,t-1}^S \leq P_{l,\text{up}}^S \tag{29}$$

$$P_{l,t-1}^S - P_{l,t}^S \leq P_{l,\text{down}}^S \tag{30}$$

where (29) and (30) are the ramp up/down constraints of the SOEC, respectively.

4. Operation constraints of SOEC

$$u_{l,t}^S P_{l,\text{min}}^S \leq P_{l,t}^S \leq u_{l,t}^S P_{l,\text{max}}^S \tag{31}$$

$$u_{p,t}^P P_{p,\text{min}}^P \leq P_{p,t}^P \leq u_{p,t}^P P_{p,\text{max}}^P \tag{32}$$

where (31) and (32) state the operating constraints of SOEC and PEMEC, respectively.

5. Conversion efficiency of P2H

$$H_{h,t}^{\{.\}} = \eta_h^{\{.\}} P_{h,t}^{\{.\}}, \{.\} = \{P, S\} \tag{33}$$

where (33) is the efficiency of SOEC and PEMEC.

2.4 | Operation constraints of subsystems in IPGS

2.4.1 | Power distribution system constraints

a. Wind power allocation constraints

$$W_{w,t}^H = \sum_{h \in \Omega_{PE}(w)} P_{h,t}^P + \sum_{h \in \Omega_{SO}(w)} P_{h,t}^S \tag{34}$$

where (34) denotes the total consumed power of P2H by SOEC and PEMEC.

b. Branch power flow constraints

$$P_{ij,t} - r_{ij} I_{ij,t} + P_{j,t}^N = \sum_{k \in c(j)} P_{jk,t} \tag{35}$$

$$Q_{ij,t} - x_{ij} I_{ij,t} + Q_{j,t}^N = \sum_{k \in c(j)} Q_{jk,t} \tag{36}$$

$$V_{j,t} = V_{i,t} - 2 \left(P_{ij,t} r_{ij} + Q_{ij,t} x_{ij} \right) + \left(r_{ij,t}^2 + x_{ij,t}^2 \right) I_{ij,t} \tag{37}$$

$$V_{i,t} I_{ij,t} \geq P_{ij,t}^2 + Q_{ij,t}^2 \tag{38}$$

where (35)–(38) describe the branch power flow in the power distribution system, also called DistFlow.

c. Nodal power balance constraints

$$P_{j,t}^N = \sum_{g \in \boldsymbol{\Omega}_{\mathrm{GT}}(j)} P_{g,t}^{\mathrm{GT}} + P_{\mathrm{sub},t} + \sum_{w \in \boldsymbol{\Omega}_{\mathrm{WT}}(j)} W_{w,t}^{\mathrm{s}} - \sum_{d \in \boldsymbol{\Omega}_{\mathrm{EL}}(j)} P_{d,t}^{\mathrm{EL}} \tag{39}$$

$$Q_{j,t}^N = \sum_{g \in \boldsymbol{\Omega}_{\mathrm{GT}}(j)} Q_{g,t}^{\mathrm{GT}} + Q_{\mathrm{sub},t} + \sum_{q \in \boldsymbol{\Omega}_{\mathrm{SVC}}(j)} Q_{q,t}^{\mathrm{SVC}} - \sum_{d \in \boldsymbol{\Omega}_{\mathrm{EL}}(j)} Q_{d,t}^{\mathrm{EL}} \tag{40}$$

where (39) and (40) represent the nodal active and reactive power balance constraints, respectively. The nodal power equals the sum of the output of GT and wind power, substation (if any) minus the load.

d. Running state of GT

$$U_{g,t}^{\mathrm{On}} - U_{g,t}^{\mathrm{Off}} = U_{g,t}^{\mathrm{GT}} - U_{g,t-1}^{\mathrm{GT}} \tag{41}$$

$$0 \leq U_{g,t}^{\mathrm{On}} \leq U_{g,t}^{\mathrm{GT}} \tag{42}$$

$$0 \leq U_{g,t-1}^{\mathrm{Off}} \leq 1 - U_{g,t}^{\mathrm{GT}} \tag{43}$$

where (41)–(43) show the relationship among GT's running state, start-up, and shut-down indicators, respectively.

e. Ramp constraints of GT

$$P_{g,t}^{\mathrm{GT}} - P_{g,t-1}^{\mathrm{GT}} \leq P_{g,\mathrm{up}}^{\mathrm{GT}} \tag{44}$$

$$P_{g,t-1}^{\mathrm{GT}} - P_{g,t}^{\mathrm{GT}} \leq P_{g,\mathrm{down}}^{\mathrm{GT}} \tag{45}$$

where (44) and (45) state GT's ramp up/down constraints.

f. Operation constraints

$$P_{g,t}^{\mathrm{GT}} + P_{g,t}^{\mathrm{UR}} \leq U_{g,t}^{\mathrm{GT}} \bar{P}_g^{\mathrm{GT}} \tag{46}$$

$$U_{g,t}^{\mathrm{GT}} \underline{P}_g^{\mathrm{GT}} \leq P_{g,t}^{\mathrm{GT}} - P_{g,t}^{\mathrm{DR}} \tag{47}$$

$$\underline{Q}_q^{\mathrm{SVC}} \leq Q_{q,t}^{\mathrm{SVC}} \leq \bar{Q}_q^{\mathrm{SVC}} \tag{48}$$

$$\underline{I}_{ij} \leq I_{ij,t} \leq \bar{I}_{ij} \tag{49}$$

$$\underline{V} \leq V_{i,t} \leq \overline{V} \tag{50}$$

where (46) and (47) denote the operating constraints considering the reserve capacity of GT; (48) states the output of the reactive power compensation device, static var compenstor (SVC) should not exceed its capacity; (49) and (50) denote the branch current of the line and bus voltage cannot exceed their limits.

2.4.2 | Natural gas system constraints

a. Pipeline natural gas flow constraints

The Weymouth equation is usually adopted to describe the relationship between the natural gas flow of pipeline and the gas pressure of node, as shown in Eq. (51) [28].

$$f_{mn,t}^2 = C_{mn}^2 (\pi_{m,t}^2 - \pi_{n,t}^2) \tag{51}$$

b. Nodal natural gas flow balance constraints

$$\sum_{m \in \boldsymbol{\Omega}_{\mathrm{P,in}(m)}} f_{gm,t} - \sum_{m \in \boldsymbol{\Omega}_{\mathrm{P,out}(m)}} f_{mn,t} + \sum_{y \in \boldsymbol{\Omega}_{\mathrm{GW}(m)}} f_{y,t}^{\mathrm{W}}$$

$$= \sum_{z \in \boldsymbol{\Omega}_{\mathrm{GL}(m)}} f_{z,t}^{\mathrm{L}} + \sum_{g \in \boldsymbol{\Omega}_{\mathrm{GT}(m)}} f_{g,t}^{\mathrm{GT}} \tag{52}$$

where (52) represents the natural gas flow balance at each node.

c. Operation constraints

$$\underline{\pi} \leq \pi_t \leq \overline{\pi} \tag{53}$$

$$\underline{f}_{mn} \leq f_{mn,t} \leq \bar{f}_{mn} \tag{54}$$

$$\underline{f}^W \leq f_{y,t}^W \leq \bar{f}^W \tag{55}$$

where (53) and (54) are the operating constraints of gas node pressure and pipeline gas flow, respectively; (55) denotes the operating constraint of the natural gas station.

d. Natural gas consumption of GT

The natural gas consumption of GT is calculated by the linear model in most cases, which is denoted as following (56):

$$f_{g,t}^{\mathrm{GT}} = P_{g,t}^{\mathrm{GT}} \eta_{g,t}^{\mathrm{GT}} \tag{56}$$

e. Hydrogen balance constraints

$$\sum_{h \in \boldsymbol{\Omega}_{\mathrm{PE}(i)}} H_{h,t}^{\mathrm{P}} + \sum_{h \in \boldsymbol{\Omega}_{\mathrm{SO}(i)}} H_{h,t}^{\mathrm{S}} = \sum_{s \in \boldsymbol{\Omega}_{\mathrm{SH}(i)}} (H_{s,t}^{\mathrm{in}} - H_{s,t}^{\mathrm{out}}) \tag{57}$$

In (57), the produced hydrogen of PEMEC and SOEC should equal the consumed hydrogen of HS.

f. Hydrogen storage constraints

$$S_{s,t}^{\mathrm{H}} = S_{s,t-1}^{\mathrm{H}} + H_{s,t}^{\mathrm{in}} \eta_s^{\mathrm{H,in}} - H_{s,t}^{\mathrm{out}} / \eta_s^{\mathrm{H,out}} \tag{58}$$

$$B_{s,t}^{\mathrm{H,in}} \underline{H}_s^{\mathrm{in}} \leq H_{s,t}^{\mathrm{in}} \leq \bar{H}_s^{\mathrm{in}} B_{s,t}^{\mathrm{H,in}} \tag{59}$$

$$B_{s,t}^{\mathrm{H,out}} \underline{H}_s^{\mathrm{out}} \leq H_{s,t}^{\mathrm{out}} \leq \bar{H}_s^{\mathrm{out}} B_{s,t}^{\mathrm{H,out}} \tag{60}$$

$$\underline{S}_{s,t}^{\mathrm{H}} \leq S_{s,t}^{\mathrm{H}} \leq \bar{S}_{s,t}^{\mathrm{H}} \tag{61}$$

$$B_{s,t}^{\mathrm{H,in}} + B_{s,t}^{\mathrm{H,out}} \leq 1 \tag{62}$$

where (58) denotes the hydrogen balance constraint of HS in two consecutive time intervals. (59) and (60) state the range of charging and discharging of HS. (61) enforces the residual hydrogen to be within its capacity limit. The binary variable is introduced in (62) to prevent charging and discharging simultaneously.

g. Hydrogen supply for hydrogen buyer constraints

$$S_{s,T}^{\mathrm{H}} = H_D \tag{63}$$

Be noted that, here, it is assumed that stored hydrogen in HS is regularly sold to the hydrogen buyer at 24:00 to meet their hydrogen demand.

2.5 | Objective function

The dispatch objective of the multi-energy operator is to minimize the total operating costs of IPGS based on day-ahead prediction data, consisting of the power and natural gas purchase costs, reserve capacity costs, start-up costs, and profits by selling hydrogen, which is denoted in (64).

$$\min C_{\mathrm{op}} = \sum_{t=1}^{T} \left(\begin{array}{l} \sum_{g=1}^{N_g} \left(p_g^{\mathrm{UR}} P_{g,t}^{\mathrm{UR}} + p_g^{\mathrm{DR}} P_{g,t}^{\mathrm{DR}} + p_g^{\mathrm{S}} U_{g,t}^{On} \right) \\ + p_{\mathrm{e},t} P_{\mathrm{sub},t} + \sum_{y=1}^{N_w} c_y f_{y,t}^{\mathrm{W}} \end{array} \right) - c_{\mathrm{H}} H_D \tag{64}$$

3 | SOLUTION METHOD

The day-ahead data-driven chance-constrained dispatch strategy for IPGS can be summarized as follows:

$$\begin{cases} \min C_{\mathrm{op}} \\ \mathrm{s.t.}(1)--(8),(18)--(63) \end{cases} \tag{65}$$

where (1)–(8) are the chance constraints, (18)–(20) are the fitted CDF and PDF by VBGMM, (21)–(33) are the operational characteristics model of P2H devices, and (34)–(63) are operation constraints of three subsystems in IPGS.

There are some non-linear constraints such as (3), (5)–(8) and (51), so the original dispatch strategy for IPGS is a stochastic mixed-integer non-linear programming (MINLP) problem. The stochastic MINLP problem is NP-hard and difficult to solve. The quantile-based analytical reformulation and second-order

cone programming (SOCP) technique are adopted to solve the above MINLP problem.

The chance constraints of the tie-line and GT in (3), (5)–(8) are usually regarded as non-convex, stochastic, and intractable , and solving this CCSP is often a non-trivial task [29]. For example, it is very hard to get the closed form of CCSP in many cases. The chance constraints cannot be solved by the commercial solvers directly. Consequently, the quantile-based analytical reformulation technique is adopted to solve this problem, then the CCSP problem is converted to a deterministic problem, described in (66)–(68).

$$P_{\mathrm{sub},t} \leq P_{\mathrm{sub}}^{\max} + Q_U(\alpha_{\mathrm{sub}}^{\max}|\tilde{e}_{w,t}) \tag{66}$$

$$\sum_{g=1}^{N_G} P_{g,t}^{\mathrm{UR}} \geq P^{\mathrm{UR,L}} - Q_U(\alpha_{\mathrm{UR}}|\tilde{e}_{w,t}) \tag{67}$$

$$\sum_{g=1}^{N_G} P_{g,t}^{\mathrm{DR}} \geq P^{\mathrm{DR,L}} + Q_U(1 - \alpha_{\mathrm{DR}}|\tilde{e}_{w,t}) \tag{68}$$

As the first parameter on the right-hand side of (66)–(68) is a predefined constant, then the core and difficulty of (66)–(68) is getting the second parameter, the quantile value of stochastic WPPE on the right-hand side. As VBGMM method is used to fit the probability distribution of WPPE, the CDF of WPPE is a linear combination of the CDF of Gaussian distribution, essentially a non-linear equation. As a result, the quantile value in Eqs. (64)–(66) can be regarded as the root of the CDF in the corresponding confidence level, which can be calculated by a common root-finding method for the non-linear equation. The fzero function in Matlab is used to find the root fast with high accuracy [30].

The SOCP handles the non-linear Weymouth constraint (51), presented below. However, this cone relaxation technique is not our contribution and focus, and more detailed descriptions are discussed in [31].

$$f_{mn,ave,t}^2 + C_{mn}^2 \pi_{n,t}^2 \leq C_{mn}^2 \pi_{m,t}^2 \tag{69}$$

The relaxation error may occur because of the cone relaxation technique, and the following equation is used to calculate the maximum relaxation gap of pipeline *mn* [32]. And the solution in relaxed SOCP model can be regarded as exact if the following relaxation gap is small enough.

$$gap = \max \left\{ \frac{C_{mn}^2 \pi_{m,t}^2 - C_{mn}^2 \pi_{n,t}^2 - f_{mn,t}^2}{f_{mn,t}^2}, \forall t \in T, \forall mn \right\} \tag{70}$$

where the *gap* is the maximum relaxation gap.

The original complex stochastic MINLP problem is converted to an easy mixed-integer SOCP (MISOCP) problem with the above cone relaxation and analytical reformulation technique. Then, many common commercial solvers can solve the mixed-integer SOCP problem quickly. The flowchart of the dispatch strategy of IPGS is shown in Figure 4.
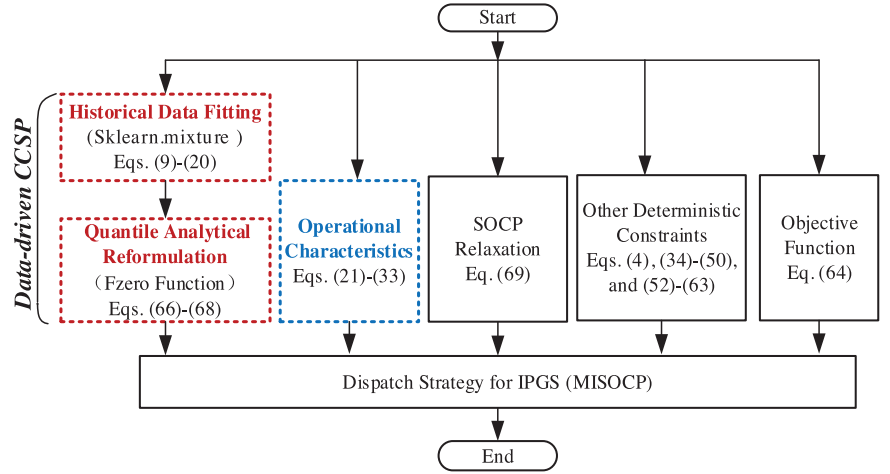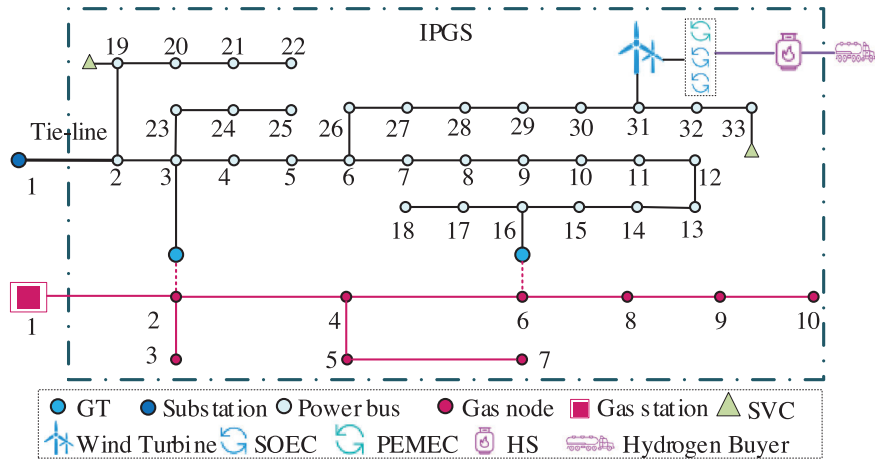
FIGURE 4 Flowchart of dispatch strategy.

FIGURE 5 Test system.

4 | CASE STUDY

In this section, the proposed data-driven chance-constrained dispatch strategy for IPGS is coded in Julia/JuMP [33] environment on a laptop with AMD R7-4800U (1.8 GHz) and 16 GB RAM. This problem is solved by the commercial solver Gurobi 9.0.3.

The modified IEEE 33-bus system integrated with a 10-node natural gas system and a micro hydrogen system is presented in Figure 5. The substation is located at bus 1, supplying electricity to IPGS through the tie-line. The wind power farm (3 MW) is located at bus 31, and one PEMEC (0.1 MW), two SOECs (0.5 MW*2), and HS is also installed at this bus. Furthermore, two GTs are located at bus 3 and 16, respectively. There is a city gas station located in gas node 1, and gas node 2 and node 6 are responsible for supplying natural gas to GTs. The base value of power is 1 MVA, and voltage magnitude is 4.16 kV. The maximum permissible value of the active power of the tie-line is 5.8 MW. The tolerance level for stochastic risks is 5%. The actual prediction errors and measured data of wind power from April 5, 2021, to April 5, 2022, of the Elia are considered as initial data [34].

There are six different cases to study the impact of WPPE and operational characteristics on the dispatch for IPGS, which

TABLE 1 Six different cases

Case	SAA	GAUSSIAN	GMM	VBGMM	Operational characteristics and optimal commitment
1	√	×	×	×	√
2	×	√	×	×	√
3	×	×	√	×	√
4	×	×	×	√	√
5	×	×	×	×	√
6	×	×	×	×	×

are described as follows: The six different cases are summarized in Table 1. Moreover, the dispatch results in six cases are presented in Table 2 and Figure 6–9, and the comparison between the six cases is discussed below. The results of the maximum relaxation gap of pipeline are 1.7×10^{-5}, and its value is small enough, so the relaxed MISOCP model is regarded as exact.

In case 1, WPPE and the operational characteristics and optimal commitment of P2H devices are considered, and the SAA method is used to deal with the chance constraints;
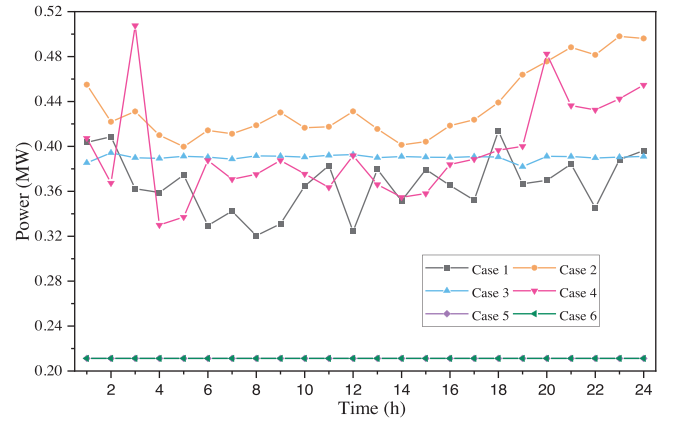
TABLE 2   Comparison of dispatch results in six cases

Case	Dispatch objective value ($)	Produced hydrogen (kg)	Computing time (s)	MIP gap	Log-likelihood Function (Bigger is Better)
1	27703.830	100	11,133.236	0.2%	–
2	27734.493	100	45.983	0	0.042
3	27744.028	100	56.799	0	0.095
4	27767.133	100	61.545	0	0.099
5	27327.228	100	48.807	0	–
6	27510.482	100	49.343	0	–
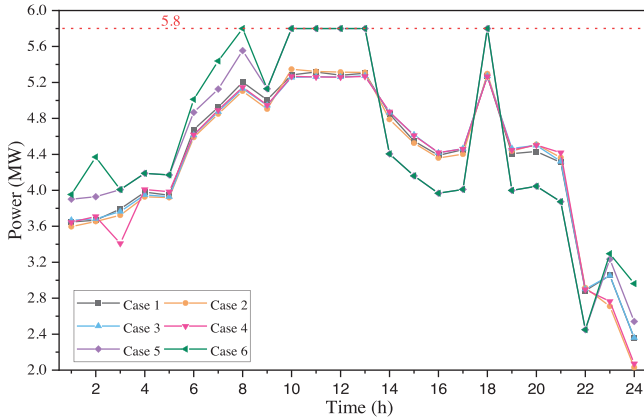
FIGURE 6   Dispatch results of the tie-line power in cases 1–6.
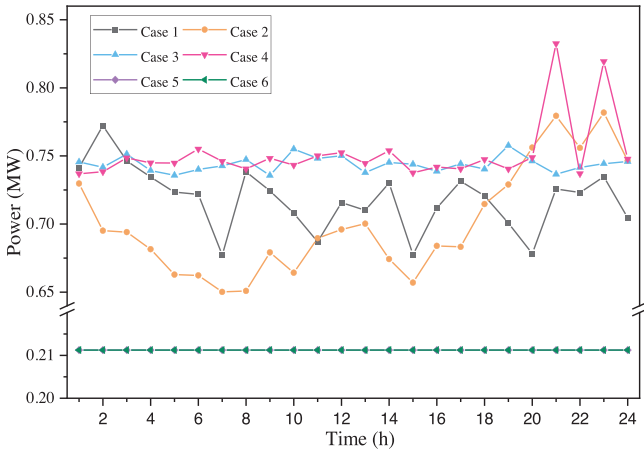
FIGURE 7   Dispatch results of the upward reserve in cases 1–6.

FIGURE 8   Dispatch results of the downward reserve in cases 1–6.
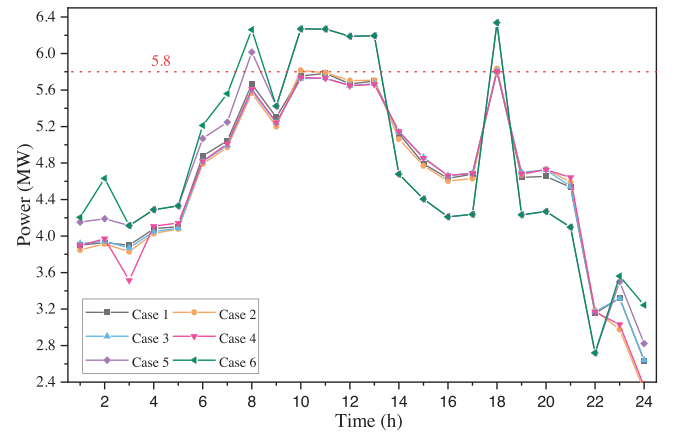
FIGURE 9   Comparison of actual tie-line power in a typical day.

In case 2, WPPE and the operational characteristics and optimal commitment of P2H devices are considered, and Gaussian distribution is used to fit the probability distribution of the WPPE;

In case 3, WPPE and the operational characteristics and optimal commitment of P2H devices are considered, and the traditional GMM is used to fit the probability distribution of the WPPE;

In case 4, WPPE and the operational characteristics and optimal commitment of P2H devices are considered, and VBGMM is adopted to fit the probability distribution of the WPPE;

In case 5, the operational characteristics and optimal commitment of P2H devices are considered, and WPPE is not considered.

In case 6, WPPE, the operational characteristics, and the optimal commitment of P2H devices are not considered, and only one traditional PEMEC (1.1 MW) is installed in this case, which is also called the traditional deterministic model.

1. Comparison of fitting effect under different probability distributions: Many references assume that WPPE follows Gaussian distribution, not conforming to reality in some geographical regions. The accuracy of the fitted probability distribution of WPPE has a great impact on the economical and secure operation of IPGS. As presented in Figure 10, the above conclusion can be obtained intuitively. The historical data of WPPE is non-Gaussian, obviously, so the fitting effect in case 1 with Gaussian distribution is the worst. Then the traditional three-component GMM method in case 3 is used to fit the probability distribution of WPPE, and the fitting effect of this method is better than in case 2. It is easy
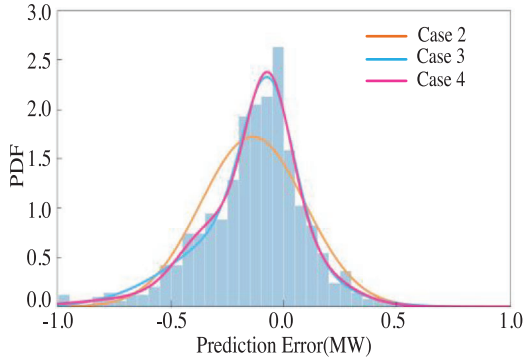
FIGURE 10 Comparison of the fitting effect in cases 2–4.

TABLE 3 Comparison of total cost with the adjustment in six cases

Case	Budgeted objective value ($)	Adjustment costs ($)	Total costs ($)	Overload times
1	27703.830	953.287	28,657.117	1
2	27734.493	957.839	28,692.332	1
3	27744.028	953.210	28,697.238	1
4	27767.133	951.943	28,719.076	1
5	27327.228	1422.426	28,749.654	5
6	27510.482	1459.455	28,969.937	6

TABLE 4 Comparison of violation rate in six cases

Case	Maximum violation rate			Average violation rate (Lower is Better)
	Overload	Insufficient upward reserve	Insufficient downward reserve	
1	6.6%	10.9%	9.5%	4.38%
2	7.2%	7.7%	3.7%	3.62%
3	5%	9.3%	8%	3.57%
4	4.8%	7.6%	5.4%	3.42%
5	81.4%	81.4%	29.7%	39.06%
6	81.4%	81.4%	29.7%	40.00%

to find that the fitting effect in case 4 with VBGMM is the best, accurately describing the characteristics of actual historical data. Besides, the fitting effect in cases 2–4 gets better, in turn, verified again by the statistical index in Table 2. Specifically, the value of the statistical index in case 4, the log-likelihood function, is the best.

2. Comparison of times of overload and adjustment cost in a typical day: As can be seen from Table 2, the dispatch objective value in cases 5 and 6 are the lowest. The dispatch decision making in case 5 and case 6 is, essentially, a deterministic optimization problem without considering WPPE, and its robustness for stochastic WPPE is weakest. To clearly show the advantage of our proposed dispatch strategy, the dispatch and adjustment results are compared on a typical day between cases 1 and 6.

As seen from Figure 6, the peak power of the tie-line in the 8th, 10th, 11th, 12th, and 18th h in cases 5 and 6 is the largest and close to the maximum permissible value of the active power of the tie-line. Moreover, according to actual historical data analysis of the Elia, the actual wind power output is usually lower than its predicted value, which can also be drawn from Figure 10. Therefore, it is necessary to purchase more power from the external power grid again to compensate for actual insufficient wind power. Then the actual power of the tie-line easily exceeds the maximum permissible value of the active power of the tie-line. The economic operation may be affected due to the adverse fluctuations of the tie-line power. By contrast, the dispatch results in cases 1–4 consider stochastic and adverse WPPE, so the peak power of the tie-line in cases 1–4 is lower, and its actual power of the tie-line does not usually exceed the maximum permissible value.

In addition, as shown in Table 3, the times of overload in case 5 are five, and the times of overload in case 6 are six. To compensate for the insufficient actual wind power output, the multi-energy operator needs to purchase more power from the external power grid again. As shown in Figure 9, the actual power of the tie-line on this typical day exceeds the maximum permissible value of the active power of tie-line in the 8th, 10th, 11th, 12th, and 18th hours in cases 5 and 6. Hence, the multi-energy operator needs to pay more power adjustment costs. In this work, the following assumptions about the adjustment price

are made, the adjustment price is time of use price if the actual power of the tie-line does not exceed the maximum permissible value. Otherwise, the adjustment price is 130% peak load price (230$/MWh). The adjustment costs in cases 5 and 6 are the largest, then their total costs are also the largest. By contrast, the adjustment costs and total costs in cases 1–4 are lower with considering stochastic WPPE. Besides, the adjustment costs in case 4 are the lowest, illustrating the benefits of the proposed data-driven CCSP method.

3. Comparison of violation rate: An actual simulation test is implemented to measure the actual violation rate of the chance constraints in the first four cases. One thousand actual WPPE samples are regarded as the input data in each chance constraint, and the violation rate is counted and analyzed. The comparison of violation rates in six cases is presented in Table 4.

As can be seen from Table 4, both the maximum violation rate and average violation rate of the deterministic optimization (cases 5 and 6) are the largest. More specifically, the maximum violation rate (overload and insufficient upward reserve violation rate) in the last two cases is very large, 81.4%, and the average violation rate of each chance constraint at every hour is about 40%. The dispatch strategy in the last two cases is a deterministic optimization problem without considering WPPE, so their violation rate is too large. IPGS using the dispatch results in cases 5 and 6 cannot operate safely with high probability.

In contrast, both the maximum and average violation rates in cases 1–4 are lower because stochastic WPPE is considered, and CCSP is used to alleviate its adverse effect. Specifically, the value of the maximum violation rate of the insufficient upward/downward reserve and average violation rate in case 1 (SAA method) is the largest among the first four cases considering WPPE, indicating its performance for resisting the uncertainty of stochastic WPPE is poor. The accuracy and feasibility of the SAA method heavily depend on the number of samples. An accurate SAA method usually requires a sufficiently large number of samples, which results in massive binary variables. Then the MISOCP with massive binary variables is difficult to solve and time-consuming. In case 1, the number of samples of the SAA method is 500 and relatively small, but it is hard to solve yet. As a result, the violation rate of the SAA is high.

Besides, the maximum violation rate of overload in case 2 (Gaussian distribution) is the largest among the first four cases, and its average violation rate is also large, indicating its performance against stochastic WPPE is better than the first two methods.

And most remarkably, the maximum violation rate of overload and insufficient upward reserve and the average violation rate in case 4 (VBGMM method) are the lowest. Hence, its robustness for stochastic prediction errors is the strongest in all cases, which also indicates the dispatch results of the tie-line and reserve capacity are more reasonable. In addition, the lowest violation also means the fitting effect of VBGMM is the best.

4. Comparison of Computing time: As shown in Table 2, the SAA with 500 samples is time-consuming and takes 11,133.236 s to get dispatch results with a large mixed-integer programming (MIP) gap (0.2%). As a result, the SAA may not be effective for practical application.

By contrast, the computing time using the quantile-based a analytical reformulation among cases 2–6 is short, about 45–60 s, and the MIP gap is 0, so this data-driven CCSP will be convenient for practical application.

5. Comparison with/without the operational characteristics and optimal commitment of P2H devices: As shown in Figure 11, the operating power of P2H device (single PEMEC) in case 6 varies greatly between the 1st and 2nd hour. By contrast, the SOECs in case 5 operate smoothly considering the operational characteristics, which is more suitable for its actual operation. In the actual operation cases, operating power of SOEC cannot vary greatly and rapidly, so it is necessary to consider the operational characteristics of P2H devices in the dispatch strategy.

Besides, as presented in Table 2, the operation cost in case 5 is lower than that in case 6 because the optimal commitment and power allocation of the two types of P2H devices (SOECs PEMEC) are considered in case 5. The conversion efficiency of compared to PEMEC, so most of the wind power is consumed by SOEC to produce hydrogen, and the remaining wind power is consumed by PEMEC. By contrast, there is only one type of P2H device, PEMEC, and all wind power is consumed by this
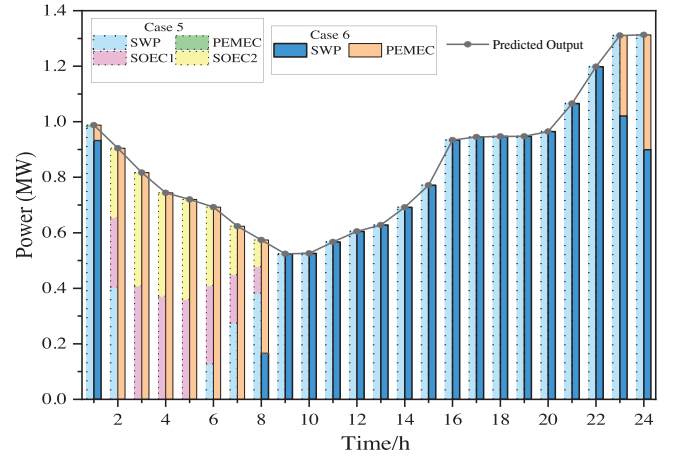

FIGURE 11 Allocation results of wind power in the last two cases.

low-conversion-efficiency PEMEC in case 6, so the operation costs in this case are larger.

In addition, as shown in Figure 11, P2H devices prefer to produce hydrogen at night due to the lower electricity price. In the actual energy market, the price of hydrogen is usually higher than natural gas, so it is wiser to produce the hydrogen and sell it to external hydrogen buyers to make more profits.

5 | CONCLUSIONS

This paper proposes a data-driven chance-constrained dispatch strategy for IPGS considering P2H operational characteristics, and some conclusions can be summarised. First, the probability distribution of WPPE can be accurately fitted by VBGMM method, whose fitting effect is better than the traditional GMM method. Next, the proposed data-driven CCSP is used to deal with stochastic WPPE and reduce its negative impact on external power grid and IPGS. Moreover, the dispatch objective value (budgeted costs) in the first four cases considering stochastic WPPE is higher, but their actual total costs are lower, illustrating the importance of considering stochastic WPPE. The actual total costs of the proposed dispatch strategy are close to that of the first three cases considering stochastic WPPE, but its maximum violation rate and average violation rate are the lowest, indicating its robustness for stochastic WPPE is the strongest. Finally, it is critical to consider the operational characteristics and optimal commitment of P2H devices in the optimal dispatch strategy for IPGS. The operation costs decrease with the optimal commitment of P2H device.

NOMENCLATURE

Sets and Indices

$d \in \boldsymbol{\Omega}_{EL}(j)$	Set of power load connected to bus j.
$g \in \boldsymbol{\Omega}_{GT}(j)$	Set of GT (gas turbine) connected to bus j.
$g \in \boldsymbol{\Omega}_{GT}(m)$	Set of GT connected to gas node m.
$i \in \boldsymbol{\Omega}_{GL}(m)$	Set of gas load connected to gas node m.
$j \in \boldsymbol{\Omega}_{GW}(m)$	Set of gas station connected to gas node m.

$m \in \mathbf{\Omega}_{\mathrm{SVC}}(j)$ Set of reactive power compensation devices connected to bus j.

$m \in \mathbf{\Omega}_{\mathrm{p,in}}(m)$ Set of pipeline ingoing gas flow.

$m \in \mathbf{\Omega}_{\mathrm{p,out}}(m)$ Set of pipeline outgoing gas flow.

$o \in \mathbf{\Omega}_{\mathrm{SO}}(i)$ Set of SOEC.

$w \in \mathbf{\Omega}_{\mathrm{PE}}(i)$ Set of PEMEC.

$w \in \mathbf{\Omega}_{\mathrm{HS}}(i)$ Set of HS (hydrogen storage) connected to bus i.

$w \in \mathbf{\Omega}_{\mathrm{WT}}(j)$ Set of wind turbine connected to bus j.

d, q Index for power demand and the reactive power compensation device.

l, g, w Index for SOEC, GT, and wind turbine.

m, n Index for gas node m/ n.

S, P   Sets of SOEC (solid oxide electrolysis cell)/ PEMEC (proton exchange membrane electrolysis cell)

y, z, t Index for natural gas station/ gas demand/ time.

Variables

$\tilde{e}_{w,t}$ Prediction errors of the wind turbine w at time t.

$f_{j,t}^{\mathrm{W}}$ Gas production of natural gas station j.

$f_{mn,t}$ Natural gas flow of the pipeline mn at time t.

$H_{h,t}^{\{.\}}$ Produced hydrogen power by PEMEC or SOEC h.

$I_{ij,t}$ Square of current of branch ij at time t.

$P_{p,t}^{\mathrm{P}}$ Operating power of PEMEC p at time t.

$Q_{q,t}^{\mathrm{SVC}}$ The output power of the reactive power compensation device q at time t.

$u_{p,t}^{\mathrm{P}}$ PEMEC p is ON at time t/$t-1$, and 0 otherwise.

$U_{i,t}^{\mathrm{GT}}$ Running state, binary variable (1 if GT g is ON at time t, and 0 otherwise).

$H_{s,t}^{\mathrm{in}}, H_{s,t}^{\mathrm{out}}$ Dis/charging power of HS s at time t.

$P_{l,t}^{\mathrm{S}}, P_{l,t-1}^{\mathrm{S}}$ Operating power of SOEC at time t/$t-1$.

$S_{s,t}^{\mathrm{H}}, S_{s,t-1}^{\mathrm{H}}$ Residual hydrogen capacity of HS s.

$\pi_{m,t}, \pi_{n,t}$ Natural gas pressure in the node m/n at time t.

$B_{s,t}^{\mathrm{H,in}}, B_{s,t}^{\mathrm{H,out}}$ The binary variable that is equal to 1 if HS s is dis/charging at time t, and 0 otherwise.

f_E, F_E PDF (probability density function)/CDF (cumulative distribution function).

$f_{z,t}^{\mathrm{L}}, f_{g,t}^{\mathrm{GT}}$ Gas demand of load z/ GTg at time t.

$P_{g,t}^{\mathrm{UR}}, P_{g,t}^{\mathrm{DR}}$ Upward/downward reserve of GT g.

$P_{g,t}^{\mathrm{GT}}, P_{g,t-1}^{\mathrm{GT}}$ Output active power of GT at time t/$t-1$.

$\tilde{P}_{\mathrm{sub},t}, P_{\mathrm{sub},t}$ Actual/Scheduled power of the tie-line

$P_{w,t}^{\mathrm{P}}, P_{w,t}^{\mathrm{S}}$ Consumed wind power w by PEMEC/ SOEC at time t.

$P_{d,t}^{\mathrm{EL}}, Q_{d,t}^{\mathrm{EL}}$ Active/ reactive power of electrical load d at time t.

$P_{j,t}^{\mathrm{N}}, Q_{j,t}^{\mathrm{N}}$ Injection active/reactive power in node j at time t.

$P_{ij,t}, Q_{ij,t}$ Active/ reactive power of branch ij.

$Q_{g,t}^{\mathrm{GT}}, Q_{g,t-1}^{\mathrm{GT}}$ Output reactive power of GT at time t/$t-1$.

$u_{l,t}^{\mathrm{S}}, u_{l,t-1}^{\mathrm{S}}$ Binary variable (1 if SOEC l is ON at time t/$t-1$, and 0 otherwise).

$U_{g,t}^{\mathrm{On}}, U_{g,t}^{\mathrm{Off}}$ Start-up/ shut-down state of GT at time t

$V_{i,t}, V_{j,t}$ Square of the voltage of bus i/j at time t.

$\tilde{W}_{w,t}^{a}, W_{w,t}^{f}$ Actual/Predicted output of the wind turbine w

$W_{w,t}^{\mathrm{H}}, W_{w,t}^{\mathrm{C}}, W_{w,t}^{\mathrm{s}}$ Consumed wind power by P2H (power-to-hydrogen), curtailed/scheduled wind power at time t.

$\omega_i, \mu_i, \sum_i$ Weight/mean/variance of component i.

C_{op} Total operation cost

Parameters

C_{mn} The constant of the Weymouth equation.

N_w, N_g, N_E Total number of wind turbine/GT/ Gaussian components.

$p_{\mathrm{e},t}$ The electricity price at the time t.

$P_{\mathrm{sub}}^{\mathrm{max}}$ The maximum permissible value of the active power of the tie-line.

$u_{l,0}^{\mathrm{S}}$ Initial commitment state of SOEC l (1 if it is online, 0 otherwise).

$\alpha_{\mathrm{sub}}^{\mathrm{max}}$ The tolerance level for overload of the tie-line

c_y, c_{H} The price of natural gas/hydrogen.

$\bar{f}_{mn}, \underline{f}_{mn}$ Max/min permissible gas flow of the pipeline mn.

$\bar{f}^W, \underline{f}^W$ Max/min permissible gas production of the city gas station.

$\bar{H}_s^{\mathrm{in}}, \underline{H}_s^{\mathrm{in}}$ Max/min charging power of HS s.

$\bar{H}_s^{\mathrm{out}}, \underline{H}_s^{\mathrm{out}}$ Max/min discharging power of HS s.

$\bar{I}_{ij}, \underline{I}_{ij}$ Max/min current square of branch ij.

$P_{p,\mathrm{max}}^{\mathrm{P}}, P_{p,\mathrm{min}}^{\mathrm{P}}$ Max/min operating power of PEMEC p.

$P_{l,\mathrm{up}}^{\mathrm{S}}, P_{l,\mathrm{down}}^{\mathrm{S}}$ Max upward/downward ramping ability of SOEC l at time t.

$P_{l,\mathrm{max}}^{\mathrm{S}}, P_{l,\mathrm{min}}^{\mathrm{S}}$ Max/min operating power of SOEC l.

$P_{p,\mathrm{max}}^{\mathrm{P}}, P_{p,\mathrm{min}}^{\mathrm{P}}$ Max/min operating power of PEMEC p.

$P^{\mathrm{UR,L}}, P^{\mathrm{DR,L}}$ Extra upward/downward reserve for power load.

$P_{g,\mathrm{up}}^{\mathrm{GT}}, P_{g,\mathrm{down}}^{\mathrm{GT}}$ Max upward/downward ramping power of GT g at time t.

$\bar{P}_g^{\mathrm{GT}}, \underline{P}_g^{\mathrm{GT}}$ Maxi/min output power of GT g.

$Q_{\mathrm{S},l}, L_{\mathrm{S},l}$ Min uptime/down time of SOEC device$l$in the initial time.

$U_{l,0}^{\mathrm{S}}, V_{l,0}^{\mathrm{S}}$ Accumulated time of SOEC l has been online/offline in the initial time.

$\bar{Q}_q^{\mathrm{SVC}}, \underline{Q}_q^{\mathrm{SVC}}$ Max/min output power of the reactive power compensation device q.

r_{ij}, x_{ij} Resistance/ reactance of branch ij.

$\bar{S}_{s,t}^{\mathrm{H}}, \underline{S}_{s,t}^{\mathrm{H}}$ Max/min capacity of HS s.

$T_l^{\mathrm{S,U}}, T_l^{\mathrm{S,D}}$ Min uptime/downtime of SOEC l.

$U_{l,0}^{\mathrm{S}}, V_{l,0}^{\mathrm{S}}$ Accumulated up/ down time of the SOEC l in the initial time.

$\overline{V}, \underline{V}$	Max/min voltage square of the bus.
α_{UR}, α_{DR}	The tolerance level for insufficient upward/downward reserve.
$\eta_h^{\{.\}}, \eta_{g,t}^{GT}$	The conversion efficiency of PEMEC/ SOEC, or the GT g at time t.
$\overline{\pi}, \underline{\pi}$	Max/min gas pressure of the node.
$p_g^{UR}, p_g^{DR}, p_g^{S}$	Cost coefficients of upward/downward reserve and startup of GT g.
Q_U, Pr	Quantile/Probability

Other notations are defined in the text

T Dispatch period, 24 h.

AUTHOR CONTRIBUTIONS

Yuehao Zhao: Data curation, Formal analysis, Investigation, Writing—original draft. Zhiyi Li: Methodology, Supervision, Review, and Editing. Ping Ju: Conceptualization, Funding acquisition, Supervision, Validation, Writing—review and editing. Yue Zhou: Writing—review and editing.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

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

ORCID

Zhiyi Li https://orcid.org/0000-0001-7498-8005

REFERENCES

1. Davidson, D.J.: Exnovating for a renewable energy transition. Nat. Energy 4(4), 254–256 (2019)
2. Wang, J., Zhong, H., Ma, Z., et al.: Review and prospect of integrated demand response in the multi-energy system. Appl. Energy 202, 772–782 (2017)
3. Raheli, E., Wu, Q., Zhang, M., et al.: Optimal coordinated operation of integrated natural gas and electric power systems: A review of modeling and solution methods. Renewable Sustainable Energy Rev. 145, 111134 (2021)
4. Fang, J., Zeng, Q., Ai, X., et al.: Dynamic optimal energy flow in the integrated natural gas and electrical power systems. IEEE Trans. Sustainable Energy 9(1), 188–198 (2018)
5. Sawas, A.M., Khani, H., Farag, H.E.Z.: On the resiliency of power and gas integration resources against cyber attacks. IEEE Trans. Ind. Inf. 17(5), 3099–3110 (2021)
6. Zhang, M., Zhang, N., Guan, D., et al.: Optimal design and operation of regional multi-energy systems with high renewable penetration considering reliability constraints. IEEE Access 8, 205307–205315 (2020)
7. Jiang, Y., Xu, J., Sun, Y., et al.: Coordinated operation of gas-electricity integrated distribution system with multi-CCHP and distributed renewable energy sources. Appl. Energy 211, 237–248 (2018)
8. Yang, Y., Wu, W., Wang, B., et al.: Analytical reformulation for stochastic unit commitment considering wind power uncertainty with Gaussian mixture model. IEEE Trans. Power Syst. 35(4), 2769–2782 (2020)
9. Qi, F., Shahidehpour, M., Li, Z., et al.: A chance-constrained decentralized operation of multi-area integrated electricity–natural gas systems with variable wind and solar energy. IEEE Trans. Sustainable Energy 11(4), 2230–2240 (2020)
10. Qiu, F., Wang, J.: Chance-Constrained Transmission Switching With Guaranteed Wind Power Utilization. IEEE Trans. Power Syst. 30(3), 1270–1278 (2015)
11. Habibi, M.K.K., Battaïa, O., Cung, V.-D., Dolgui, A., Tiwari, M.K.: Sample average approximation for multi-vehicle collection–disassembly problem under uncertainty. Int. J. Prod. Res. 57(8), 2409–2428 (2019). https://doi.org/10.1080/00207543.2018.1519262
12. Patil, G.R., Ukkusuri, S.V.: Sample average approximation technique for flexible network design problem. J. Comput. Civ. Eng. 25(3), 254–262 (2011), https://doi.org/10.1061/(ASCE)CP.1943-5487.0000086
13. Bertsimas, D., Gupta, V., Kallus, N.: Robust sample average approximation. Math. Program. 171(1–2), 217–282 (2018). https://doi.org/10.1007/s10107-017-1174-z
14. Jiang, Y., Wan, C., Wang, J., et al.: Stochastic receding horizon control of active distribution networks with distributed renewables. IEEE Trans. Power Syst. 34(2), 1325–1341 (2019)
15. Wu, H., Shahidehpour, M., Li, Z., et al.: Chance-constrained day-ahead scheduling in stochastic power system operation. IEEE Trans. Power Syst. 29(4), 1583–1591 (2014)
16. Wang, Y., Zheng, Y., Yang, Q.: Nash bargaining based collaborative energy management for regional integrated energy systems in uncertain electricity markets. Energy 269, 126725 (2023). https://doi.org/10.1016/j.energy.2023.126725
17. Wang, Z., Shen, C., Liu, F., et al.: Chance-constrained economic dispatch with non-Gaussian correlate wind power uncertainty. IEEE Trans. Power Syst. 32(6), 4880–4893 (2017)
18. Ye, L., Zhang, Y., Zhang, C., Lu, P., Zhao, Y., He, B.: Combined Gaussian mixture model and cumulants for probabilistic power flow calculation of integrated wind power network. Communist Chin. Sci. Abstr. 74, 117–129 (2019). https://doi.org/10.1016/j.compeleceng.2019.01.010
19. Sun, W., Zamani, M., Hesamzadeh, M.R., et al.: Data-driven probabilistic optimal power flow with nonparametric Bayesian modeling and inference. IEEE Trans. Smart Grid 11(2), 1077–1090, (2020)
20. Wang, Z., Shen, C., Liu, F., Gao, F.: Analytical expressions for joint distributions in probabilistic load flow. IEEE Trans. Power Appar. Syst. 32(3), 2473–2474 (2017). https://doi.org/10.1109/TPWRS.2016.2612881
21. Qi, F., Shahidehpour, M., Wen, F., et al.: Decentralized privacy-preserving operation of multi-area integrated electricity and natural gas systems with renewable energy resources. IEEE Trans. Sustainable Energy 11(3), 1785–1796 (2020)
22. Xing, X., Lin, J., Song, Y., et al.: Modeling and operation of the power-to-gas system for renewables integration: A review. CSEE J. Power Energy Syst. 4(2), 168–178 (2018).
23. Scikit-learn. Variational Bayesian Gaussian Mixture. https://scikit-learn.org/sTable/modules/mixture.htm-l#variational-bayesian-gaussian-mixture. Accessed July 07th 2022.
24. National Grid. Introduction of National Grid. https://www.nationalgrid.com/about-us (2022). Accessed July 07th 2022.
25. Li, J., Jin, J., Xing, X., et al.: Technology portfolio selection and optimal planning of power-to-hydrogen modules in active distribution network. Proc. CSEE 41(12), 4021–4032 (2021)
26. Li, Y., Schofield, E., Gönen, M.: A tutorial on Dirichlet process mixture modeling. J. Math. Psychol. 91, 128–144 (2019)
27. Blei, D.M., Jordan, M.I.: Variational inference for Dirichlet process mixtures. Bayesian Anal. 1(1), 121–143 (2006)
28. Correa-Posada, C.M., Sanchez-Martin, P.: Integrated power and natural gas model for energy adequacy in short-term operation. IEEE Trans. Power Syst. 30(6), 3347–3355 (2015)

29. Hu, Z., Sun, W., Zhu, S.: Chance constrained programs with Gaussian mixture models. IISE Trans. 54(12), 1117–1130 (2022). https://doi.org/10.1080/24725854.2021.2001608

30. MATALB. Fzero function. https://ww2.mathworks.cn/help/matlab/ref/fzero.html (2022). Accessed July 07th 2022.

31. Wen, Y., Qu, X., Li, W., et al.: Synergistic operation of electricity and natural gas networks via ADMM. IEEE Trans. Smart Grid 9(5), 4555–4565 (2018)

32. Li, Y., Li, Z., Wen, F., Shahidehpour, M.: Minimax-regret robust co-optimization for enhancing the resilience of integrated power distribution and natural gas systems. IEEE Trans. Sustainable Energy 11(1), 61–71 (2020). https://doi.org/10.1109/TSTE.2018.2883718

33. Dunning, I., Huchette, J., Lubin, M.: JuMP: A modeling language for mathematical optimization. SIAM Rev. 59(2), 295–320 (2017)

34. Elia. Wind power generation. https://www.elia.be/en/grid-data/power-generation/wind-power-generation (2021). Accessed May 07th 2022.