energies

MDPI

Article

Data Driven Robust Energy and Reserve Dispatch Based on a Nonparametric Dirichlet Process Gaussian Mixture Model

Li Dai *, Dahai You and Xianggen Yin

State Key Laboratory of Advanced Electromagnetic Engineering and Technology, School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, No.1037, Luoyu Road, Wuhan 430074, China; youdahaihust@163.com (D.Y.); xgyin@hust.edu.cn (X.Y.)

* Correspondence: daili2007@yeah.net; Tel.: +86-158-7175-7029

Received: 12 July 2020; Accepted: 3 September 2020; Published: 7 September 2020

Abstract: Traditional robust optimization methods use box uncertainty sets or gamma uncertainty sets to describe wind power uncertainty. However, these uncertainty sets fail to utilize wind forecast error probability information and assume that the wind forecast error is symmetrical and independent. This assumption is not reasonable and makes the optimization results conservative. To avoid such conservative results from traditional robust optimization methods, in this paper a novel data driven optimization method based on the nonparametric Dirichlet process Gaussian mixture model (DPGMM) was proposed to solve energy and reserve dispatch problems. First, we combined the DPGMM and variation inference algorithm to extract the GMM parameter information embedded within historical data. Based on the parameter information, a data driven polyhedral uncertainty set was proposed. After constructing the uncertainty set, we solved the robust energy and reserve problem. Finally, a column and constraint generation method was employed to solve the proposed data driven optimization method. We used real historical wind power forecast error data to test the performance of the proposed uncertainty set. The simulation results indicated that the proposed uncertainty set had a smaller volume than other data driven uncertainty sets with the same predefined coverage rate. Furthermore, the simulation was carried on PJM 5-bus and IEEE-118 bus systems to test the data driven optimization method. The simulation results demonstrated that the proposed optimization method was less conservative than traditional data driven robust optimization methods and distributionally robust optimization methods.

Keywords: energy and reserve dispatch; data driven robust optimization; polyhedral uncertainty set; nonparametric Dirichlet process Gaussian mixture model

1. Introduction

In recent years, the world has witnessed a dramatic increase in wind power integration into the power grid, as it diminishes fossil fuel consumption and environmental pollution. The U.S. Department of Energy forecasts that wind power will supply 20% of electricity generation by 2030 [1]. In Germany, system operators should see renewable energy, including wind power, as a priority [2]. In China, more than 90 GW of wind farms have been built, and the wind power penetration rate is expected to reach 11% by 2020 [3]. However, the large integration of wind power also poses enormous challenges in power system operations due to the variable and uncertain nature of wind power. In this cases, it is important to co-optimize the energy and reserve to ensure adequacy of the electricity supply.

Two main optimization methods, stochastic optimization and robust optimization, have been adopted to tackle the uncertainty of wind power in power system operations. For stochastic

optimization, the uncertainty of wind power is captured by a set of scenarios that are sampled from a certain deterministic distribution. However, it is hard to obtain the exact distribution of wind power, as the accuracy of the stochastic optimization relies on the sample generation technique [4,5]. A robust optimization puts the uncertainty information in an uncertainty set, which includes the worst case scenario and finds a solution that is optimal for any realization of the given uncertainty sets. Compared to stochastic optimization, a robust optimization tends to produce conservative solutions because it only optimizes the worst case scenario in the uncertainty set [6]. The conservativeness of the optimal solution is determined by the size of the uncertainty sets. However, it is not easy to determine the uncertainty set that can make a trade-off between the optimality and conservativeness. To overcome the stochastic optimization's specificity deficiency and the robust optimization's conservativeness deficiency, a distributionally robust optimization (DRO) assumes that the uncertainty probability distribution lies in the ambiguity set [7] and minimizes worst case expected cost. The moment-based approach and statistical distance-based approach are two typical ways to construct the ambiguity set. The moment-based approach considers the wind power probability distribution with a known mean and variance derived from historical data [8–10]. The statistical distance-based approach establishes a confidence set of the ambiguity set by introducing a statistical distance between two probability distributions. Ref. [11]. used historical data to build an ambiguity set and provide a confidence band estimation for the cumulative distribution function (CDF) based on a nonparametric estimation of the parameters. Ref. [12] proposed a two-stage data driven, distributionally robust reserve and energy scheduling model where the operational risk is obtained with a Wasserstein ball-based method. The aforementioned studies demonstrate that extracting reliable statistical information from available data is crucial for making a robust and less conservative dispatch decision. However, these methods convert the DRO model into a semidefinite programming or a second-order conic programming problem to improve numerical tractability, which may lead to a suboptimal solution.

The power industry provides a large amount of wind power forecast error historical data which provides valuable information to support power systems' optimization dispatch [13,14]. Recent work has proposed data driven robust uncertainty sets to overcome the overconservative nature of traditional robust uncertainty sets [15,16]. Ref. [17] proposed calibration and volume indices to assess the uncertainty set. Calibration shows the gap between the empirical coverage rate of an uncertainty set to its nominal coverage rate. Volume shows the size of the uncertainty for the required probability guarantee. In this paper, we aimed to propose a highly skilled uncertainty set with a predefined probability guarantee. This helps the operator to know in advance the probability of optimization that could reslt violate the constraint [18]. Traditional uncertainty sets are constructed based on a Gaussian distribution [19] for nodal load and [20] wind power. However, a Gaussian distribution is not adequate in terms of describing the wind power uncertainty [17]. Deping Ke [21] and Zhiwen Wang [22] used a Gaussian mixture model (GMM) to approximate the probability distribution of the wind power forecast error. However, the accuracy of the GMM depends on the number of Gaussian components.

The Dirichlet process Gaussian mixture model (DPGMM), as an infinite mixture model, is capable of classifying the historical data set without any prior knowledge of the number of mixture components. The parameters in the DPGMM are estimated by the variational inference algorithm. Based on these estimated DPGMM parameters, we developed a data driven polyhedral uncertainty set for the wind power forecast error. The conservativeness of the solution is controlled by the scale parameter. The scale parameter is set according to the predefined data coverage rate the operator prefers. We involved the uncertainty set in the robust energy and reserve dispatch model, relaxing the model into a master and subproblem framework. In this framework, the master problem determined the generator output, up and down the reserve, whereas the subproblem identified the worst case scenario that led to the largest imbalance. We used the column and constraint generation algorithm to solve this model. The IEEE-5 bus and IEEE-118 buses are used to test the effectiveness of our proposed model.

The major novelties of this paper are summarized as follows:

1. A novel data driven polyhedral uncertainty set for wind power forecast error based on the nonparametric Dirichlet process Gaussian mixture model was developed. By using the combined DPGMM framework and variational inference algorithm, we estimated the GMM parameter embedded within the historical data. To control the conservativeness of the solution, we developed a method to calculate the minimum scale parameter with a predefined data coverage rate. Furthermore, the performance of the proposed data driven polyhedral uncertainty set was compared with other data driven uncertainty sets. The results indicated that the proposed uncertainty set has smaller volume than other uncertainty sets with the same predefined data coverage rate.

2. A novel data driven robust energy and reserve dispatch optimization framework was proposed. In this framework, the first stage determined the generator output and reserve, while the second stage determined the reschedule output of generators after the wind power output was determined. To improve the solution efficiency and solution quality, we introduced the Binary Expansion algorithm to identify inactive constraints.

3. The proposed data driven optimization method was compared with other data driven robust optimization methods and the distributionally robust optimization methods on an IEEE-5 bus and IEEE-118 bus. The simulation results indicated that the proposed data driven optimization method was less conservative than traditional robust optimization methods and distributionally robust optimization methods.

The remainder of this paper is organized as follows. Section 2 presents a detailed formulation of data driven robust polyhedral uncertainty set based on the Dirichlet process Gaussian mixture model. Section 3 presents the data driven robust energy and reserve dispatch model. Section 4 presents the simplified model and model solution methodology. Section 5 presents the computational experiments. Section 6 presents the conclusion.

2. Data Driven Polyhedral Uncertainty Set Based on the Dirichlet Process Gaussian Mixture Model

The power system provides massive wind power forecast error historical data. To extract accurate distribution information from the data and build an uncertainty set based on this information, two steps were conducted. First, we built a Dirichlet process Gaussian mixture model (DPGMM) for wind power and used variational inference to estimate the relative distributional parameters. Then, a data driven polyhedral uncertainty set based on these estimated parameters was built.

2.1. Nonparametric Dirichlet Process Gaussian Mixture Model

In this paper, a power system provided N wind power forecast error data $w = \left\{w_1^D, \ldots, w_N^D\right\}$ and where D is the dimensional of data. The joint probability distribution of w can be expressed as a Gaussian mixture model (GMM):

$$p(w|\pi, \theta) = \sum_{m=1}^{M} \pi_m N(w|\mu_m, \Lambda_m^{-1}) \tag{1}$$

where $\theta = \{\mu_m, \Lambda_m\}_{m=1}^{M}$, π_m is the weight of the m component. $\pi_m \geq 0$, $\sum_{m=1}^{M} \pi_m = 1$. $N(\cdot)$ is the Gaussian distribution. μ_m, Λ_m denotes the mean and precision of mth Gaussian component. In GMM, the number of Gaussian components is unknown. If M is not set properly, there is a significant gap between the observed data distribution and the GMM distribution estimated from the data. The DPGMM is a nonparametric mixture model with the Dirichlet process as the prior distribution of the number of components in the model. The model is an infinite mixture model, which means that specifying the number of components in advance is no longer necessary.

According to the DPGMM mixture theory, each w_i^D is generated by first choosing a component indexed by z_i, which is distributed according to $\pi = [\pi_1, \pi_2, \ldots, \pi_M]$. Afterward, the w_i^D is generated from choosing the Gaussian component with the parameter θ_i. The parameter θ_i is generated from

a prior distribution G because different θ_i are exchangeable and may be of same values. Therefore, we assumed that G is generated from a random draw in the Dirichlet process $DP(\alpha, G_0)$ rather than a continuous function. We summarized the basic form of the DPGMM as follows:

$$\begin{aligned} G &\sim DP(\alpha, G_0) \\ \theta_i|G &\sim G \\ x_i|\theta_i &\sim N(\mu_m, \Lambda_m^{-1}) \end{aligned} \tag{2}$$

where α is a concentration parameter and G_0 is the base distribution. Because the mean μ_m and the precision Λ_m are unknown, we set the conjugate priors for base distribution G_0 as a Normal–Wishart distribution:

$$G_0 \sim N(\mu|\mu_0, (\beta_0\Lambda)^{-1})Wi(\Lambda|\Psi_0, v_0) \tag{3}$$

where $Wi(\cdot)$ denotes the Wishart distribution and the μ_0, β_0, Ψ_0 and v_0 are the hyperparameters of the base distribution G_0.

The unique values of $\theta_1, \ldots, \theta_n$ induce a partitioning of the data set w into clusters such that within each cluster m, the θ_i take on the same value θ_m^*. The distribution over partitions is called the Chinese restaurant process. In this process, a Chinese restaurant has an infinite number of tables, each of which can seat an infinite number of customers. A sequence of customers $\{\theta\}_1^{N-1} = \{\theta_1, \theta_2, \ldots, \theta_{N-1}\}$ comes into the restaurant and chooses a table to sit at. The Nth customer θ_N can either sit at an existing table or choose a new one. The distribution form is given below [23]:

$$\theta_N = \begin{cases} \frac{N_c}{\alpha+N-1}\delta(\theta_N = \theta_m^*)(m = 1, \ldots, M) & \textit{at an occupied table } m \\ \frac{\alpha}{\alpha+N-1}\theta, \theta \sim G_0 & \textit{at a new table} \end{cases} \tag{4}$$

where N_c is the total sample number with a value equal to θ_m^* and $\delta(\cdot)$ is the delta function. From the above equation, one can infer that the posterior shows interesting clustering effects since θ_N shall be assigned with a new cluster with probability $(\alpha/\alpha + N - 1)$ or an existing cluster with probability $(N_c/\alpha + N - 1)$. Note that for the nonparametric Bayesian method, this clustering behavior is used to allocate parameters and should be the cornerstone for model optimization. In other words, parameter sets can be automatically clustered and adjusted within several discrete clusters according to the potentially allocated samples. Thus, giving two infinite collections of random variables: $V_k \sim Beta(1, \alpha)$ and $\theta_k \sim G_0$. The stick breaking construction can be given as [24]

$$G = \sum_{k=1}^{\infty} \pi_k(\mathrm{V})\delta(\theta = \theta_k^*) \tag{5}$$

with weight $\pi_k(\mathrm{V}) = v_k \prod_{j=1}^{k-1}(1 - v_j)$. The construction of π can be understood as follows. Starting with a stick of length 1, we break it at v_1, assigning π_1 to be the length of stick we just broke off. Then we recursively break the other portion to obtain π_2, π_3, and so forth. $\delta(\theta = \theta_k^*)$ is an indicator function centered on θ_k^*. The whole structure of the DPGMM is shown in Figure 1.

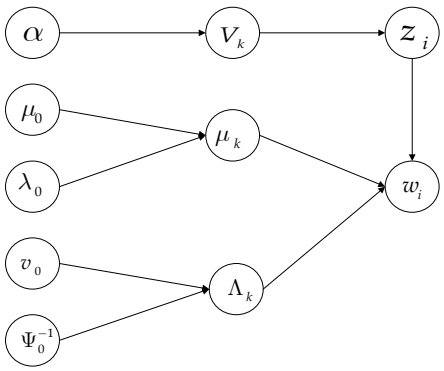


Figure 1. The Dirichlet process Gaussian mixture model.

2.2. Data Driven Polyhedral Uncertainty Set

The data driven polyhedral uncertainty set is constructed based on the posterior predictive distribution $\Pr(w_n|\mathrm{W}^e)$, where random vector w_n represents the future wind power forecast errors. Given the model and empirical parameters, the prediction distribution for a new sample is a mixture of a Student's t-distribution [25]:

$$p(w_n|M,\Theta) = \sum_{m=1}^{M} \pi_m St(w_n\big|\mu_m, S_m, v_m + 1 - D) \tag{6}$$

where π_m is the weight of the mth GMM mixture component. $\pi_m = v_m \prod_{j=1}^{M-1}(1-v_j)$, $v_m = \frac{\gamma_1}{\gamma_2}$. M is the truncation level. For each component, the expected value is μ_m and the covariance is $S_m = \left[\frac{(v_m+1-D)}{1+\beta_m}\beta_m\Psi_m\right]^{-1}$. Parameter γ_1, γ_2, μ_m, β_m, v_m, and Ψ_m are the inference results of the mth component using variational inference [26].

In order to build a flexible and easy to solve uncertainty set capable of capturing the correlation, asymmetry, and multimode of the wind forecast error, we proposed a data driven polyhedral uncertainty set based on the DPGMM. The DPGMM can automatically extract the number of GMM components and parameter information embedded in the wind power forecast error data. Based on the estimated GMM components and parameter information, the uncertainty set is defined as a union of several basic polyhedral uncertainty sets. Mathematically, the proposed data driven polyhedral uncertainty sets are cast as follows:

$$W_1^D = \underset{m:\pi_m \geq \pi^*}{\cup} \left\{w|\ \|\eta_m(w-\mu_m)\|_1 \leq \Gamma_m^\alpha\right\} \tag{7}$$

$$W_\infty^D = \underset{m:\pi_m \geq \pi^*}{\cup} \left\{w|\ \|\eta_m(w-\mu_m)\|_\infty \leq \Delta_m^\alpha\right\} \tag{8}$$

where π^* is the threshold. We set it as 2~5%. α is the predefined coverage rate of uncertainty set. Γ_m^α is called scale parameters. η_m is a upper triangular matrix and can be obtained by the Cholesky decomposition of S_m^{-1}.

The conservativeness of the uncertainty set is controlled by scale parameter Γ_m^α and Δ_m^α. However, for the W_1^D and W_∞^D polyhedra uncertainty set, there is no direct way to calculate the scale parameter. Therefore, we proposed a methodology to find the minimum scale parameter that provides the required data coverage rate. In the proposed method, for each component m, we input the mean value μ_m and η_m to the following equation and calculated ξ_i for each observed wind power forecast error data w_i.

$$\xi_i = \|\eta_m(w_i - \mu_m)\|_1 \tag{9}$$

For the desired data coverage rate α, Γ_m^α is considered as the N_1^{th} smallest ξ_i, where N_1 is as follows:

$$N_1 = round(N_m \times \alpha) \tag{10}$$

$round(x)$ is a function which returns the closest integer to x. N_m is the number of data for the mth component.

3. Data Driven Robust Energy and Reserve Dispatch Model

The data driven robust energy and reserve dispatch problem is described below. The predispatch variables $\left\{p_g, r_g^{up}, r_g^{dn}\right\}$ are the generator output and generator up reserve and down reserve, respectively. The redispatch variables $\left\{\Delta p_g\right\}$ are an adjustment of generators after $\{p_m^w\}$ is observed.

A. Objective function

$$\min_{p_g^f, r_g} F = \sum_{g=1}^{N_G} (a_g (p_g)^2 + b_g p_g + c_g + c_r^u r_g^{up} + c_r^d r_g^{dn}) \tag{11}$$

Objective function (11) minimizes the operating costs of generators, including the generation cost and up and down reserve cost. The quadratic term can be linearized by using the piecewise linear approximation technique [27].

B. Predispatch constraints

(1) Power balance constraint

$$\sum_{g=1}^{N_G} p_g + \sum_{m=1}^{N_W} p_m^{we} = \sum_{d=1}^{N_D} p_d \tag{12}$$

Equation (12) is the power balance constraint corresponding to the forecast wind power.

(2) Generation physical constraints

$$p_g + r_g^{up} \le p_g^u \ \forall g \tag{13}$$

$$P_g^l \le p_g - r_g^{dn} \ \forall g \tag{14}$$

Equation (13) ensures the combination of scheduled generator output p_g and up reserve r_g^{up} is below the maximum generation limits p_g^u. Equation (14) ensures the combination of the scheduled generator output p_g and down reserve r_g^{dn} is above the minimum generation limits P_g^l.

(3) Network power flow constraints

$$-F_l \le \sum_{g=1}^{N_G} \pi_{gl} p_g + \sum_{m=1}^{N_w} \pi_{ml} p_m^{we} - \sum_{d=1}^{N_D} \pi_{ql} p_d \le F_l \quad l = 1, 2, \ldots, L \tag{15}$$

Equation (15) is the network power flow constraint corresponding to the forecast wind power.

(4) Spinning reserve constraints

$$0 \le r_g^{up} \le R_g^+ \Delta t, \ 0 \le r_g^{dn} \le R_g^- \Delta t \tag{16}$$

Equation (16) ensures that generator g has enough ramp capability to provide up and down reserve capacity r_g^{up}, r_g^{dn} in the dispatch interval.

C. Redispatch constraints

The redispatch constraints must be satisfied for all possible realizations of wind generation in the set W^D. The redispatch constraints are given as follows:

$$-r_g^{dn} \le \Delta p_g \le r_g^{up} \ \forall g \tag{17}$$

$$p_g^c = p_g^f \Delta p_g \ \forall g \tag{18}$$

$$\sum_{g=1}^{N_G} p_g^c + \sum_{m=1}^{N_W} p_m^{we} + \sum_{m=1}^{N_W} w_m = \sum_{q=1}^{N_Q} p_d \quad \forall w \in W^D \tag{19}$$

$$-F_l \le \sum_{g=1}^{N_G} \pi_{gl} p_g^c + \sum_{m=1}^{N_W} \pi_{ml}(p_m^{we} + w_m) - \sum_{d=1}^{N_d} \pi_{dl} p_d \le F_l \ \forall l \ \ \forall w \in W^D \tag{20}$$

Equation (17) ensures that the generator redispatch action ΔP_g is constrained by the up and down reserve capacity determined in the predispatch. Equation (18) defines the generator corrective dispatch output. Equation (19) is the power balance constraint under the redispatch condition. Equation (20) is the network power flow constraints under the redispatch condition.

4. Simplified Model and Model Solution Methodology

4.1. Simplified Model

The proposed dispatch model can be equivalent to two stage robust optimization problems. The abstract form of the proposed dispatch model is provided as below:

$$\min c^T x, \tag{21}$$

$$s.t.\ Ax \le b - Cw^0, \tag{22}$$

$$Hx + Dy \ge 0, \tag{23}$$

$$\forall w \in W^D, \exists y : By \ge Ax + Cw - b. \tag{24}$$

In the above, x is the first stage dispatch decision variable, including the generator output p_g, as well as the generator up and down reserve r_g^{up}, r_g^{dn}. y is the second stage redispatch variable Δp_g. Equation (22) represents constraints (12)–(16). Equation (23) represents constraint (17). Equation (24) represents constraints (18)–(20).

4.2. Solution Methodology

As constraint (24) consists of many wind power forecast error scenarios, it is impossible to enumerate all these scenarios in the model. In this study, we used the column and constraint generation (C&CG) method to violate the worst case scenario only when necessary. The C&CG algorithm iterative solves the master problem and a set of subproblems in which the master problem is a relaxation of the original problem with finite constraints and generates the generator output p_g, as well as the up and down reserve r_g^{up}, r_g^{dn}. For a given master problem variable, the subproblem is then used to identify the worst case wind forecast error scenario. The column and constraint associated with the identified worst case violated scenario are then fed back to problem (MP) for the next iteration. The process is repeated until no violated scenarios are identified. The flowchart of the C&CG method is given in Figure 2. ε represents the tolerance of the relative optimality gap. For each iteration k, we generated a new column and constraint in the master problem, including additional decision variable $y^{(k)}$ and constraint $By^{(k)} \ge Ax + Cw^{(k)} - b, \ \forall k$.

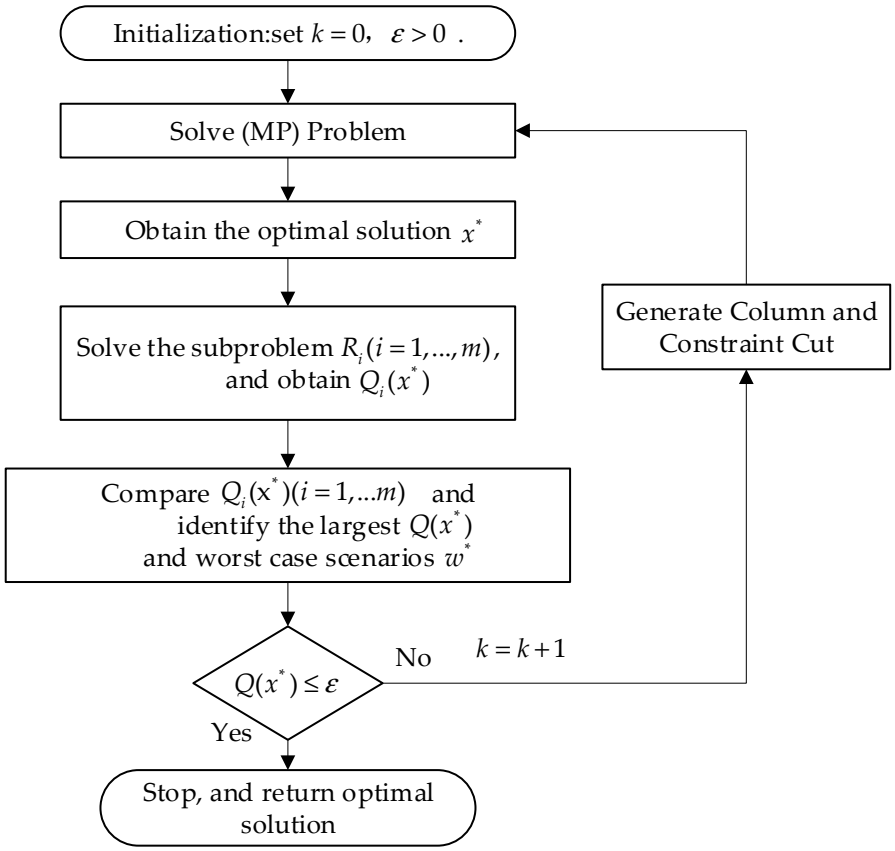


Figure 2. Flowchart of the proposed solution algorithm.

A. Master Problem

The master problem constitutes a relaxed version of problem (11)–(20) at each iteration k. A set of the column and constraint is identified by the subproblem in the previous iteration. For a given iteration k, the master problem is formulated as follows:

$$\begin{aligned} &\min c^T x, \\ &s.t.\ Ax \le b - Cw^0, \\ &By^{(k)} \ge Ax + Cw^{(k)} - b \end{aligned} \tag{25}$$

where $y^{(k)}$ is associated with the worst case uncertainty realization and $w^{(k)}$ is identified by the subproblem at iteration k. w^0 represents the nominal wind power scenarios.

B. Subproblem

For presentation convenience, the subproblem is defined as follows:

$$\begin{aligned} R(x^*) = &\max_{w\in W_1^D\cup W_2^D\cup W_m^D} \min_{y,s^+,s^-} 1^T s \\ &s.t.\ Hx^* + Dy \ge 0 \quad : \lambda \\ &By + s \ge Ax^* + Cw - b \quad : \mu \end{aligned} \tag{26}$$

In this equation, s represents the nonnegative slack variables. If $R = 0$, no wind power spillage and load shedding are required to balance the system. If $R > 0$, there is at least one realization within the uncertainty set for which recourse actions are not feasible. Compared to the traditional uncertainty set, which is a single uncertainty set, the proposed uncertainty set is a union of several uncertainty sets. Therefore, for each uncertainty set, we solve the corresponding subproblem R_i and choose the largest subproblem solution $Q_i(x^*)$ as the final subproblem solution.

By dualizing the inner minimization problem, the max-min R_i problem can be equivalent to bilinear programming as follows:

$$\begin{array}{c} R_i(x^*) = \max\limits_{w \in W_i, u} u^T(Ax^* - b) + u^T C w - \lambda^T H x^* \\ s.t.\ \lambda^T D + u^T B = 0^T, 0 \le u^T \le 1^T, \lambda^T \ge 0^T \end{array} \tag{27}$$

Note that $u^T C w$ is a bilinear function. Ref. [28] and Ref. [29] used the extreme point (EP) approach to linearize the bilinear function. However, the EP approach is intractable when the number of EPs is large. Researchers also employ the outer approximation (OA) [6] method to solve the max-min problem with a good computation efficiency. Although this method has a good computation efficiency, this approach only provides a local optimal solution. In this paper, we used the binary expansion (BE) approach [30] to linearize the bilinear terms. As $u_i \in [0, 1]$, u_i can be expressed as

$$u_i \approx \sum_{n=0}^{N} 2^{-n} v_{i,n},\ \forall i \tag{28}$$

where $v_{i,n}$ is a binary variable associated with the n-th exponential term in the BE expression, and N is the integer parameter used to define the maximum exponential order.

We introduce the auxiliary parameter $\underline{z} \in R^{N_s}$, $\overline{z} \in R^{N_s}$ and let $\underline{z}_i = \min\limits_{w \in W_i} C_i w$, $\overline{z}_i = \max\limits_{w \in W_i} C_i w - \underline{z}_i$. $z = Cw - \underline{z}$. Then, the bilinear term $u^T C w$ can be expressed as

$$\begin{array}{c} u^T C w = u^T(z + \underline{z}) = \sum\limits_{i}^{N_s} u_i z_i + u^T \underline{z} \\ \approx \sum\limits_{i}^{N_s} \sum\limits_{n=0}^{N} 2^{-n} v_{i,n} z_i + u^T \underline{z} \end{array} \tag{29}$$

Defining the new variable $\rho_{i,n} = v_{i,n} z_i$ leads to

$$\rho_{i,n} = \begin{cases} z_i,\ if\ v_{i,n} = 1, \\ 0,\ if\ v_{i,n} = 0. \end{cases} \tag{30}$$

As $0 \le z \le \overline{z}$, the above constraint can be equivalent to

$$0 \le \rho_{i,n} \le \overline{z}_i v_{i,n},\ z_i - \overline{z}_i + \overline{z}_i v_{i,n} \le \rho_{i,n} \le z_i \tag{31}$$

Finally, the robust optimization problem $R_i(x^*)$ is approximated as

$$R_i(x^*) = \max_{w \in W_i, u, \lambda} \sum_{i}^{N_s} \sum_{n=0}^{N} 2^{-n} \rho_{i,n} + u^T(Ax^* - b + \underline{z}) - \lambda^T H x^* \tag{32}$$

$$s.t.\ \lambda^T D + u^T B = 0^T \tag{33}$$

$$\sum_{n=0}^{N} 2^{-n} v_{i,n} \le u_i \le \sum_{n=0}^{N} 2^{-n} v_{i,n} + 2^{-N},\ \forall i \tag{34}$$

$$z = Cw - \underline{z} \tag{35}$$

$$0^T \le u^T \le 1^T,\ u^T \ge 0^T,\ \lambda^T \ge 0^T \tag{36}$$

$$z_i - \overline{z}_i + \overline{z}_i v_{i,n} \le \rho_{i,n} \le z_i,\ 0 \le \rho_{i,n} \le \overline{z}_i v_{i,n}\ \forall i, n \tag{37}$$

$$v_{i,n} \in \{0, 1\},\ \forall i, n \tag{38}$$

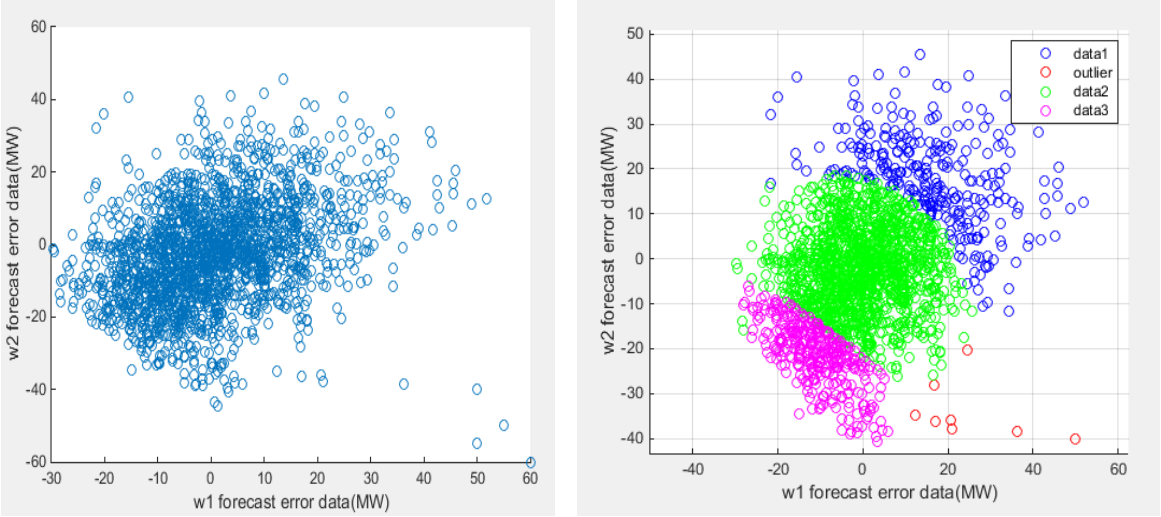
The large size of C leads to a heavy computation burden in the subproblem $R_i(x^*)$, as the size of C is determined by the power balance constraints and power flow constraints. To reduce the size of C, we identified the inactive power flow constraints and reduced the problem size of the subproblem [30].

5. Computational Experiments

In this section, we first used real historical wind power forecast error data to evaluate the performance of the proposed data driven polyhedral uncertainty set. Then, we tested the proposed data driven robust optimization method on the IEEE-5 bus and IEEE-118 bus systems. All optimization problems were solved with Gurobi and implemented on a computer with an Intel (R) Core i7 @2.7GHz and 8GB RAM. The optimality tolerance for Gurobi was set to 0 and the relative optimality gap tolerance for the algorithm was 10^{-4}. The simulation results of the proposed method were compared with those of other robust and distributionally robust optimization methods.

5.1. Performance Evaluation of Proposed Data Driven Polyhedral Uncertainty Set

In this section, we tested the performance of the proposed uncertainty set using real wind power data extracted from the Eastern Wind Dataset provided by the National Renewable Energy Laboratory (NREL) [31]. As the distribution of the wind power forecast error is conditional to the level of forecast values, we proved that the closer the historical forecast value is to the future forecast, the more likely the historical forecast error distribution is to follow the future forecast error distribution. Therefore, we used the data selection strategy [32] to select 2000 wind power forecast error data for two wind farms. This strategy calculates the Euclidean distances between the historical forecast values and the upcoming forecast values. Then, it sorts the sequence in ascending order and chooses 2000 minimal elements as the selected wind power forecast error data. The scatter plot of selected historical forecast error data is shown in Figure 3a. As shown in Figure 3a, the wind power forecast error data is asymmetrical, multimodal, and correlated. We used the DPGMM and variational inference algorithm to estimate the information of the forecast error distribution. The results are shown in Figure 3b. There were three GMM components and outliers that could potentially enlarge the uncertainty set.



(**a**) (**b**)

Figure 3. (**a**) Scatter plot of data set; (**b**) Gaussian mixture model (GMM) data cluster.

In the following, we compared the performance of the proposed data driven robust polyhedral uncertainty set with other data driven robust uncertainty sets in terms of calibration and volume. Calibration reflects the deviations between the expected coverage rates and observed coverage rates. This index helps the operator know in advance the degree of constraint violation upon obtaining the

solution of the optimization problems. Volume reflects the size of uncertainty sets for the required probability guarantee. Because there is no direct method to calculate the uncertainty set volume, we used the Monte Carlo method to generate N' random samples in the feasible range. The feasible range was bounded by the upper and lower boundaries in each dimension of data sets. Then, we calculated the proportion of points that lay in the data driven robust uncertainty set. The volume of the data driven robust uncertainty set is calculated as

$$V^P = N_{in} V^c / N' \tag{39}$$

where N_{in} is the number of wind power forecast error points enveloped by the robust uncertainty set and V^c is the volume of the bounded hyper-cube.

Table 1 gives the comparison results of the proposed data driven robust polyhedral uncertainty sets and other data driven robust uncertainty sets, including the polyhedral uncertainty sets P_1^D, P_∞^D [20] and box uncertainty sets. As can be seen in Table 1, the calibration of polyhedral uncertainty sets P_1^D and P_∞^D are zero. This means that the expected coverage rate of P_1^D and P_∞^D uncertainty sets are the same with the observed coverage rate. The calibration of the box uncertainty set and data driven polyhedral uncertainty sets W_1^D and W_∞^D nearly approach zero. This means that the expected coverage rate of these uncertainty sets are nearly the same with the observed coverage rate. Among the five uncertainty sets, W_1^D showed the best overall performance in terms of volume. This is because the data driven polyhedral uncertainty set used the DPGMM model to capture the asymmetry, multimode of forecast error data, and compact uncertainty of three polyhedral uncertainty sets. However, the polyhedral uncertainty set P_1^D and P_∞^D only used a single polyhedron. The P_1^D covers too much of an unnecessary region because it assumes all the forecast error data follow a uniform distribution and generates a symmetric uncertainty set not in accordance with the asymmetry feature of wind power forecast error data. The volume of polyhedral uncertainty sets P_1^D and P_∞^D are smaller than the box uncertainty set because the polyhedral uncertainty sets capture the correlation information of the forecast error data.

Table 1. Comparison of calibration and volume for different data driven robust uncertainty sets.

Robust Uncertainty Set	Box Uncertainty Set	Polyhedral Uncertainty Set P_1^D	Polyhedral Uncertainty Set P_∞^D	Data Driven Polyhedral Uncertainty Set W_1^D	Data Driven Polyhedral Uncertainty Set W_∞^D
Expected coverage rate	0.999	0.999	0.999	0.999	0.999
Calibration	−0.0005	0	0	0.0005	0.0005
Volume	753.34	719.41	741.76	630.49	681.25

Figure 4 compares the performance of the proposed data driven polyhedral uncertainty sets and other data driven uncertainty sets in dimension 4. As can be seen in Figure 4a, the observed coverage rates of data driven polyhedral uncertainty sets W_1^D and W_∞^D are larger than the expected coverage rate. This is because the data driven polyhedral uncertainty sets are a union of intersected polyhedral uncertainty sets. Each polyhedral uncertainty set not only covers the data belonging to this component, but also other intersected data sets. The difference between the observed coverage rate and expected coverage rate is that there is an approach to zero at 0.99 in the expected coverage rate. Figure 4a also reports that the observed coverage rate of the box uncertainty set, P_1^D and P_∞^D, are nearly the same as the expected coverage rate. Figure 4b gives the estimated volume of various uncertainty sets under different expected coverage rates. As can be seen in Figure 4b, the expected coverage rate increased from 0.91 to 0.99 and the volume of the proposed data driven uncertainty sets was obviously smaller than that of other data driven uncertainty sets, especially when the expected coverage rate reached 0.99. This proved that the proposed data driven robust uncertainty set was less conservative than other data driven robust uncertainty sets with the same data coverage rate.

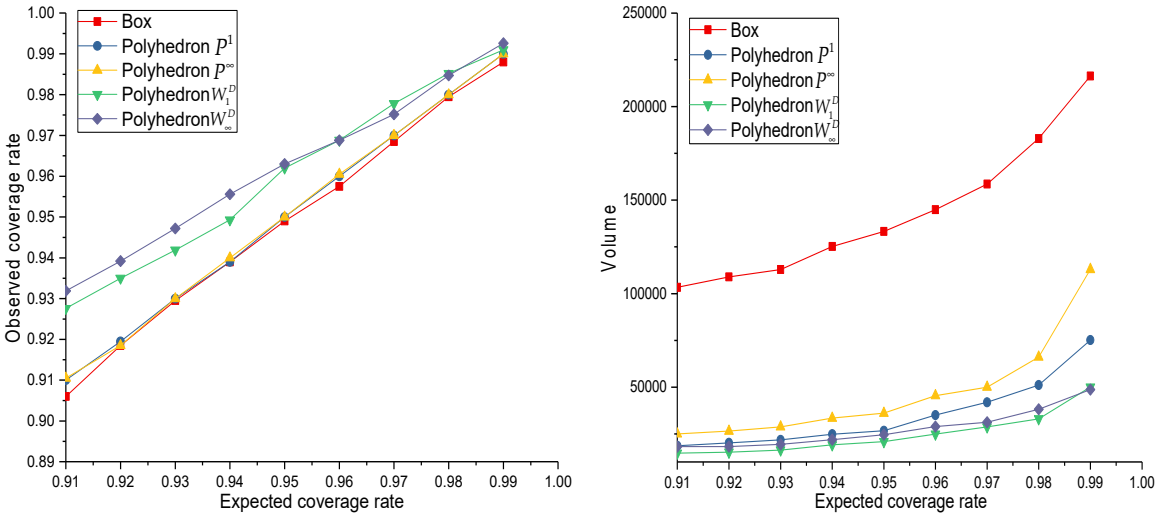


Figure 4. (**a**) The relationship between the expected coverage rate and observed coverage rate, dimension 4; (**b**) the estimated volume of the uncertainty set under a different expected coverage rate, dimension 4.

5.2. Case Study on an PJM-5 Bus System

The PJM 5 bus system diagram is shown in Figure 5. Moreover, there are two wind farms located at bus C and D. The predictive output of wind power is 150 and 100 MW. The detailed unit data are given in Table 2. The transmission line data are given in a previous work [33]. We used the real historical wind power forecast error data mentioned in Section 5.1. A scatter plot of this data is shown in Figure 3a.

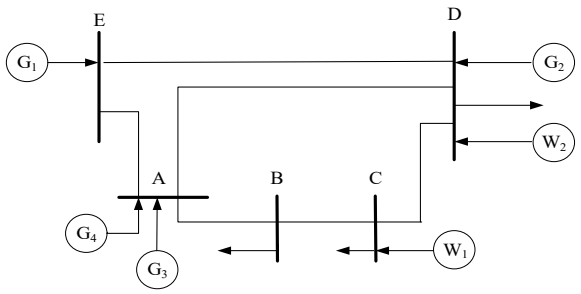


Figure 5. Topology of the five bus system.

Table 2. Parameters of the generator.

Unit No	P_{min}/P_{max} MW	*a* (USD/MW)	b (USD/MW)	c (USD/MW)	Reserve Price (USD/MW)	Ramp MW/h
1	[180 400]	0.01059	8.3391	64.16	30	100
2	[100 300]	0.010875	12.8875	6.78	45	60
3	[150 600]	0.01057	10.76	32.96	54	150
4	[120 500]	0.008401	12.3299	28	40	120

Figure 6a shows that the proposed data driven polyhedral uncertainty set can accurately extract the correlation, asymmetry, and multimode feature of wind power forecast error data. Furthermore, the proposed uncertainty set can identify the outlier data that potentially leads to conservative solutions. The polyhedral uncertainty set P_1^D shown in Figure 6b only uses the mean and covariance information of historical data, and generates a symmetric uncertainty not in accordance with the feature of wind power

forecast error data. The box uncertainty set uses the upper bound and lower bound information of historical data. The gamma uncertainty set only uses the uncertainty deviation information and adjusts the size of the uncertainty set via a scale parameter. Figure 6c,d illustrates the box uncertainty set and gamma uncertainty set that covers too many unnecessary regions, which could lead to overconservative robust energy and reserve dispatch solutions.

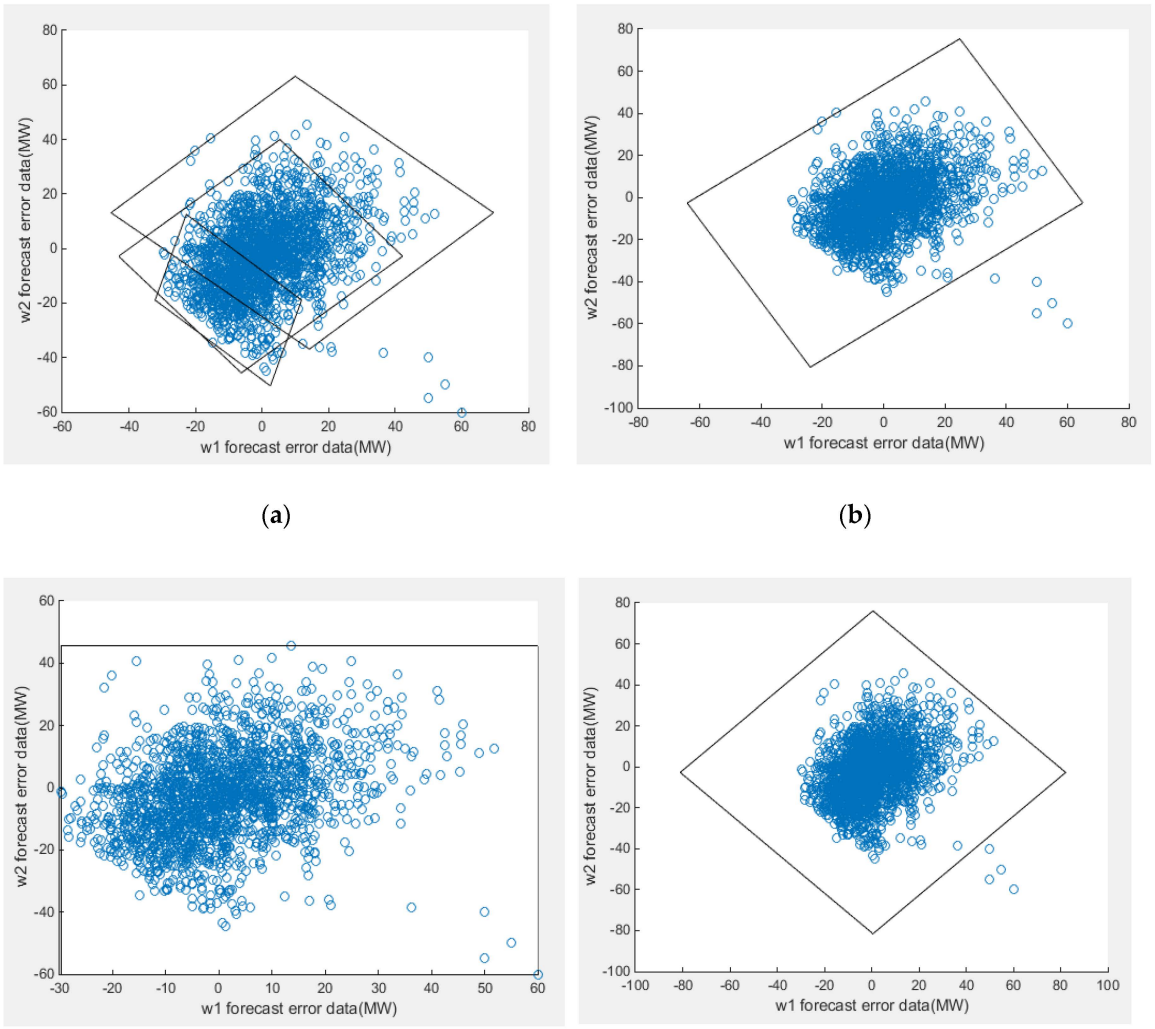


Figure 6. (**a**) Data driven robust uncertainty set W_1^D; (**b**) polyhedral uncertainty set P_1^D; (**c**) box uncertainty set; (**d**) gamma uncertainty set.

We compared the operation cost, up and down reserve capacity, as well as the reserve cost and computational time of the proposed data driven robust optimization methods (DDRO) and other robust optimization methods, such as a robust optimization with the box uncertainty set (BRO), robust optimization with the polyhedral uncertainty set (PRO), and robust optimization with the gamma uncertainty set (GRO). The results are provided in Table 3. The data driven robust optimization (DDRO) methods requires more computational time than other robust optimization models because the DDRO model needs to solve three subproblems in each iteration. The operation cost and reserve cost of DDRO with W_1^D is smaller than DDRO with W_∞^D. This is because the DDRO with W_1^D has a smaller volume than the DDRO with W_∞^D. Compared to other robust optimizations, the DDRO with W_1^D has a much lower reserve cost, which leads to a lower operation cost. More specifically, the reserve cost of DDRO with W_1^D is 46.65% less than BRO, 53.4% less than PRO with P_1^D, and 23.91% less than

GRO. This is because the DDRO with W_1^D captures the asymmetry, multimode and correlation of uncertainty data, and provides a compact uncertainty set that covers fewer regions than other robust optimization methods. The up and down reserve provided by DDRO with W_1^D is different, as it verifies that the proposed uncertainty set captures the asymmetry features of wind power forecast error data. The operation cost and reserve cost of GRO is smaller than BRO and PRO with P_1^D, which indicates that the dispatch strategy obtained by GRO is more economical than the one obtained by BRO and PRO with P_1^D. This is because GRO considers the correlation between two wind farms.

Table 3. Comparison of other robust optimization methods.

Uncertainty Set	Box Uncertainty Set	Polyhedral Uncertainty Set (P_1^D)	Gamma Uncertainty Set	Data Driven Robust Uncertainty Set (W_1^D)	Data Driven Robust Uncertainty Set (W_∞^D)
Data coverage rate	99.9%	99.9%	99.9%	99.9%	99.9%
Operation cost ($)	20,413	20,757	19,570	18,446	19,207
Up reserve (MW)	89.70	104.56	83.74	51.87	65.42
Down reserve (MW)	105.60	100.27	79.50	82.68	93.60
Reserve cost ($)	5915	6193	4996	4036.6	4770.3
Computation time (s)	15	18.36	30.18	40.61	34.50

5.3. Case Study on an IEEE-118 Bus System

In this section, an experiment on a modified IEEE-118 system was carried out to demonstrate the effectiveness and superiority of the proposed data driven robust optimization method. Three wind farms were added at Buses 17, 66, and 94, each with a capacity of 500 MW. The detailed generator line and load data are available online at [34]. The predicted outputs of the three wind generators were 185, 210 and 245 MW.

In this paper, 4000 data based on a multivariate Gaussian mixture model was generated. Figure 7 gives the scatter plot of the forecast error data. Figure 8a,b show the data driven polyhedral uncertainty set W_1^D and W_∞^D with a 99.9% predefined data coverage rate. Further, Figure 8a,b show that the proposed data driven polyhedral uncertainty set based on the DPGMM can accurately extract the number of GMM components in the wind power forecast error data. Figure 8c,d give the plot of the polyhedral uncertainty set P_1^D and the box uncertainty set.

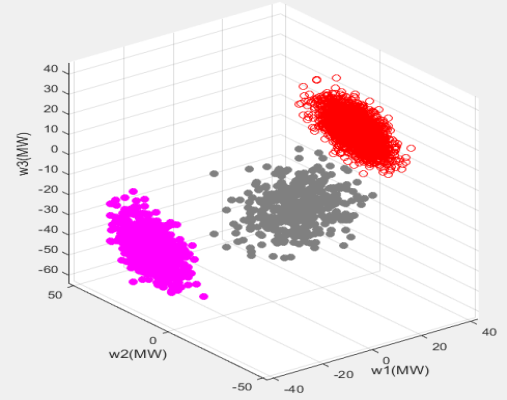


Figure 7. Scatter plot of forecast error data.

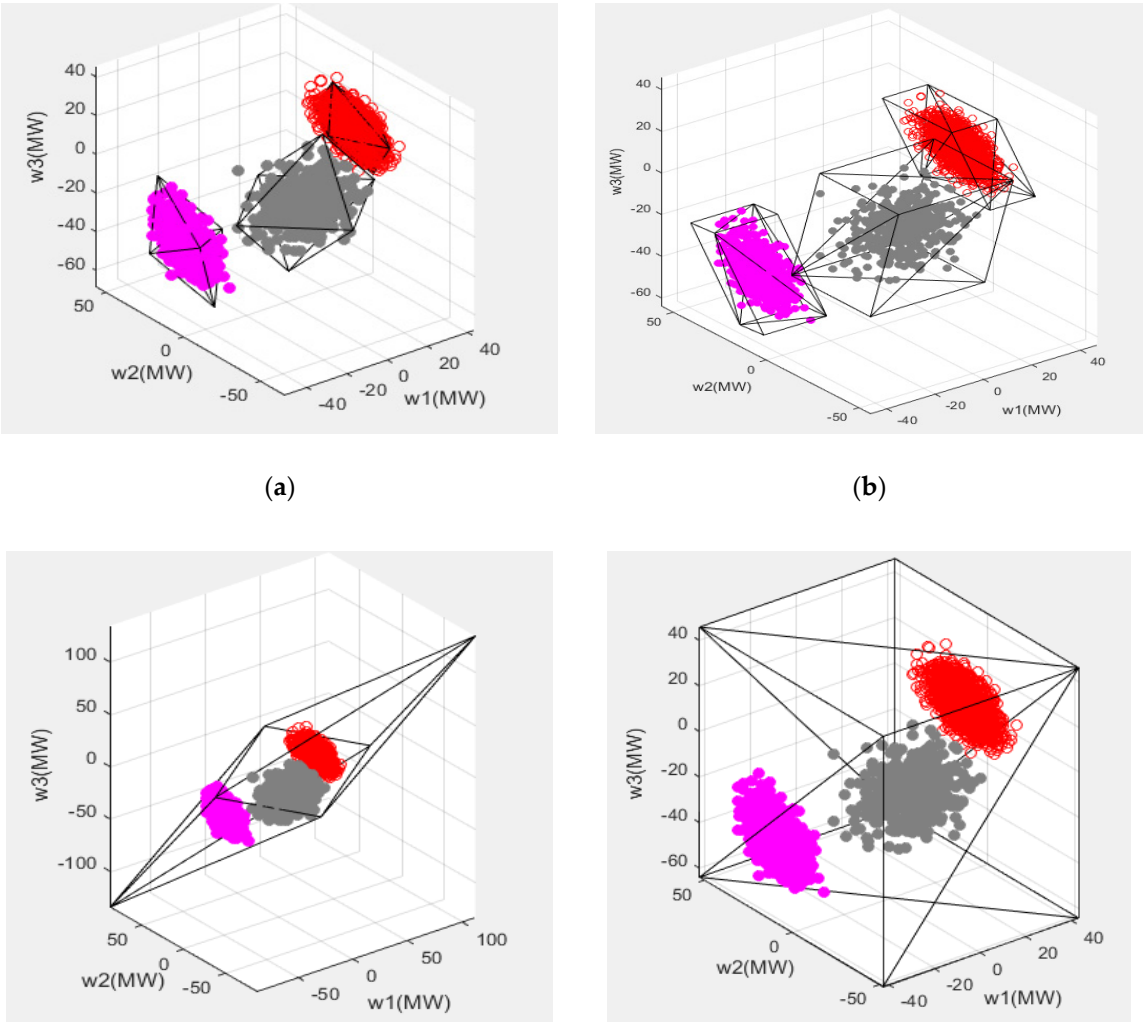


Figure 8. (**a**) Data driven polyhedral uncertainty set W_1^D; (**b**) data driven polyhedral uncertainty set W_∞^D; (**c**) Polyhedral uncertainty set P_1^D; (**d**) Box uncertainty set.

We compared the operation cost, up and down reserve capacity, as well as reserve cost and computation time of the DDRO method with other optimization methods, such as BRO, PRO (with P_1^D), and the moment distributionally robust optimization (MDRO) method. The results are given in Table 4. The reserve cost of DDRO with W_1^D was 86.4% smaller than PRO and 85.51% smaller than BRO. This is because the data driven polyhedral uncertainty set is constructed based on the GMM parameter information extracted via a nonparametric DPGMM. Furthermore, the operation cost of DDRO with P_1^D was the smallest, while the operation cost of PRO was the largest. The difference value between the largest operation cost and other model operation cost was shown in the second line, which demonstrates that the proposed data driven robust optimization methods were obviously less conservative than the other model. The computation time of DDRO was larger than PRO with P_1^D and BRO because the DDRO needed to solve three subproblems in each iteration. Conversely, PRO with P_1^D and BRO only solved one subproblem in each iteration. However, the computation time of DDRO was still acceptable in practical usage and applicable in large scale power system. This is because the W_1^D is a polyhedron and we identified the inactive power flow constraints to reduce the problem size.

Table 4. Comparison of different optimization methods.

	MDRO	**PRO (P_1^D)**	**BRO**	**DDRO (W_1^D)**	**DDRO (W_∞^D)**
Operation cost (USD)	6,181,277	6,181,945	6,181,927	6,180,928	6,181,082
Difference value (USD)	668	0	18	1017	863
Up reserve (MW)	110.0	133.69	158.32	104.91	120.08
Down reserve (MW)	110.07	160.07	141.85	54.93	76.25
Reserve cost (USD)	1526.1	2194.2	2176.2	1177.3	1331.1
Computation time (s)	350	27	53	110.32	120.12

In MDRO, we assumed that the wind power forecast error distribution would follow the probability distribution with the mean and variance information estimated from the historical wind power data and then optimize the expected cost under the worst possible distribution. As can be seen in Table 4, compared with DDRO with P_1^D, both the reserve cost and operation cost of MDRO were larger than DDRO with W_1^D because DDRO with W_1^D fully used the information embedded in the historical wind power forecast error data, while MDRO only used the mean and covariance information derived from the historical data. Furthermore, the computation time of MDRO was larger than DDRO with W_1^D. This is because the MDRO method used delayed constraint generation algorithm and the alternate convex search algorithm to solve this model [35]. In each iteration, this algorithm solved semidefinite programming (solved by SDP3), quadratic programming (solved by KNITRO), and linear programming.

6. Conclusions

In this paper, a data driven robust polyhedral uncertainty set based on the DPGMM was proposed. We used real wind power historical data to test the performance of the proposed uncertainty set. The results indicated that the proposed uncertainty set had a smaller volume than other data driven uncertainty sets with the same predefined data coverage data. This superiority was especially obvious for the high dimensional data because the uncertainty set captured the correlation, multimode, and asymmetry information of the forecast error data. Based on the constructed data driven polyhedral uncertainty set, a data driven robust optimization method for energy and reserve dispatch problems was proposed. Compared with other robust optimization methods, the proposed DDRO method was less conservative with the same predefined coverage rate. Compared with the MDRO, the proposed data driven robust optimization method had a smaller operation cost and reserve cost. This is because the proposed data driven robust optimization method used a compact uncertainty set that is a union of a several basic polyhedral uncertainty sets. The number of uncertainty sets, and mean and covariance information are extracted from the data by using the nonparametric DPGMM algorithm. However, the MDRO only uses the mean and covariance information of data to estabilish the ambiguous probability distributions of wind power. Moreover, the computation time of the DDRO method was acceptable because we used a polyhedral uncertainty set and reduced the problem size by identifying the inactive constraints. This ensures that our model was applicable to real time energy and reserve dispatch.

Author Contributions: L.D. conceived the idea, proposed the optimization model, wrote the original paper, performed simulations and analyzed the data. D.Y. and X.Y. analyed the results and proofread the original paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. U.S. Department of Energy. *20% Wind Energy By 2030: Increasing Wind Energy's Contribution to US Electricity Supply*; Diane Publishing: Collingdale, PA, USA, 2008.
2. IRENA. *REmap 2030: A Renewable Energy Roadmap*; International Renewable Energy Agency: Abu Dhabi, UAE, 2014.
3. International Energy Agency. *China Wind Energy Development Roadmap 2050*; International Energy Agency: Paris, France, 2011.
4. Zheng, Q.P.; Wang, J.; Liu, A.L. Stochastic Optimization for Unit Commitment—A Review. *IEEE Trans. Power Syst.* **2015**, *30*, 1913–1924.
5. Lowery, C.; O'Malley, M.J. Reserves in Stochastic Unit Commitment: An Irish System Case Study. *IEEE Trans. Sustain. Energy* **2015**, *6*, 1029–1038. [CrossRef]
6. Bertsimas, D.; Litvinov, E.; Sun, X.A.; Zhao, J.; Zheng, T. Adaptive Robust Optimization for the Security Constrained Unit Commitment Problem. *IEEE Trans. Power Syst.* **2013**, *28*, 52–63. [CrossRef]
7. Delage, E.; Ye, Y. Distributionally Robust Optimization Under Moment Uncertainty with Application to Data-Driven Problems. *Oper. Res.* **2010**, *58*, 595–612. [CrossRef]
8. Chen, Y.; Wei, W.; Liu, F.; Mei, S. Distributionally robust hydro-thermal-wind economic dispatch. *Appl. Energy* **2016**, *173*, 511–519. [CrossRef]
9. Wei, W.; Liu, F.; Mei, S. Distributionally Robust Co-Optimization of Energy and Reserve Dispatch. *IEEE Trans. Sustain. Energy* **2016**, *7*, 289–300. [CrossRef]
10. Wang, Z.; Bian, Q.; Xin, H.; Gan, D. A Distributionally Robust Co-Ordinated Reserve Scheduling Model Considering CVaR-Based Wind Power Reserve Requirements. *IEEE Trans. Sustain. Energy* **2016**, *7*, 625–636. [CrossRef]
11. Duan, C.; Jiang, L.; Fang, W.; Liu, J. Data-driven Affinely Adjustable Distributionally Robust Unit Commitment. *IEEE Trans. Power Syst.* **2017**, *33*, 1385–1398.
12. Yao, L.; Wang, X.; Duan, C.; Guo, J.; Wu, X.; Zhang, Y. Data-driven distributionally robust reserve and energy scheduling over Wasserstein balls. *IET Gener. Transm. Distrib.* **2018**, *12*, 178–189. [CrossRef]
13. Zhao, C.; Guan, Y. Data-Driven Stochastic Unit Commitment for Integrating Wind Generation. *IEEE Trans. Power Syst.* **2015**, *31*, 2587–2596. [CrossRef]
14. Xiong, P.; Singh, C. A Distributional Interpretation of Uncertainty Sets in Unit Commitment under Uncertain Wind Power. *IEEE Trans. Sustain. Energy* **2018**, *10*, 149–157. [CrossRef]
15. Guan, Y.; Wang, J. Uncertainty Sets for Robust Unit Commitment. *IEEE Trans. Power Syst.* **2014**, *29*, 1439–1440. [CrossRef]
16. Ning, C.; You, F. Data-Driven Adaptive Robust Unit Commitment Under Wind Power Uncertainty: A Bayesian Nonparametric Approach. *IEEE Trans. Power Syst.* **2019**, *34*, 2409–2418. [CrossRef]
17. Golestaneh, F.; Pinson, P.; Azizipanah-Abarghooee, R.; Gooi, H.B. Ellipsoidal Prediction Regions for Multivariate Uncertainty Characterization. *IEEE Trans. Power Syst.* **2018**, *33*, 4519–4530. [CrossRef]
18. Golestaneh, F.; Pinson, P.; Gooi, H.B. Polyhedral Predictive Regions for Power System Applications. *IEEE Trans. Power Syst.* **2018**, *34*, 693–704. [CrossRef]
19. Hu, B.; Wu, L. Robust SCUC With Multi-Band Nodal Load Uncertainty Set. *IEEE Trans. Power Syst.* **2016**, *31*, 2491–2492. [CrossRef]
20. Venzke, A.; Halilbasic, L.; Markovic, U.; Hug, G.; Chatzivasileiadis, S. Convex Relaxations of Chance Constrained AC Optimal Power Flow. *IEEE Trans. Power Syst.* **2018**, *33*, 2829–2841. [CrossRef]
21. Ke, D.; Chung, C.Y.; Sun, Y. A Novel Probabilistic Optimal Power Flow Model with Uncertain Wind Power Generation Described by Customized Gaussian Mixture Model. *IEEE Trans. Sustain. Energy* **2015**, *7*, 200–212. [CrossRef]
22. Wang, Z.; Shen, C.; Liu, F.; Wu, X.; Liu, C.C.; Gao, F. Chance-Constrained Economic Dispatch with Non-Gaussian Correlated Wind Power Uncertainty. *IEEE Trans. Power Syst.* **2017**, *32*, 4880–4893. [CrossRef]
23. Blei, D.M.; Jordan, M.I. Variational inference for Dirichlet process mixtures. *Bayesian Anal.* **2006**, *1*, 121–143. [CrossRef]
24. Teh, Y.W.; I Jordan, M.; Beal, M.J.; Blei, D.M. Hierarchical Dirichlet Processes. *J. Am. Stat. Assoc.* **2006**, *101*, 1566–1581. [CrossRef]
25. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: New York, NY, USA, 2006.

26. Zhu, J.; Ge, Z.; Song, Z. Variational Bayesian Gaussian Mixture Regression for Soft Sensing Key Variables in Non-Gaussian Industrial Processes. *IEEE Trans. Control Syst. Technol.* **2016**, *25*, 1092–1099. [CrossRef]
27. Wu, L. A Tighter Piecewise Linear Approximation of Quadratic Cost Curves for Unit Commitment Problems. *IEEE Trans. Power Syst.* **2011**, *26*, 2581–2583. [CrossRef]
28. Jiang, R.; Wang, J.; Guan, Y. Robust Unit Commitment with Wind Power and Pumped Storage Hydro. *IEEE Trans. Power Syst.* **2012**, *27*, 800–810. [CrossRef]
29. Ye, H.; Li, Z. Robust Security-Constrained Unit Commitment and Dispatch with Recourse Cost Requirement. *IEEE Trans. Power Syst.* **2016**, *31*, 3527–3536. [CrossRef]
30. Ye, H.; Wang, J.; Li, Z. MIP Reformulation for Max-min Problems in Two-stage Robust SCUC. *IEEE Trans. Power Syst.* **2016**, *32*, 1237–1247. [CrossRef]
31. NREL National Renewable Energy Laboratory. Available online: https://www.nrel.gov/wind/data-tools.html (accessed on 24 December 2018).
32. Li, Z.; Qiu, F.; Wang, J. Data-driven real-time power dispatch for maximizing variable renewable generation. *Appl. Energy* **2016**, *170*, 304–313. [CrossRef]
33. Wei, W.; Liu, F.; Mei, S.; Yunhe, H. Robust Energy and Reserve Dispatch Under Variable Renewable Generation. *IEEE Trans. on Smart Grid* **2015**, *6*, 369–380. [CrossRef]
34. 118bus_ro. Available online: http://motor.ece.iit.edu/data/118bus_ro.xls (accessed on 25 March 2020).
35. Dai, L.; You, D.; Yin, X.; Wang, G.; Zou, Q. Distributionally robust dynamic economic dispatch model with conditional value at risk recourse function. *Int. Trans. Electr. Energy Syst.* **2018**, *29*, e2775. [CrossRef]