Nonparametric probabilistic forecasting based stochastic optimal scheduling of integrated electricity and gas systems

Jinyu Chen and Dawei Chen*

State Grid Fujian Electric Power Research Institute, Fuzhou, China

The volatility and sporadic availability of renewable energy create significant challenges to the optimal scheduling of integrated electricity and gas systems (IEGS). This paper develops a nonparametric probabilistic forecasting based stochastic scheduling approach of IEGS. The quantile at a series of quantile levels can be generated by direct quantile regression method. Given the set of predicted quantiles, a set of representative scenarios for wind power uncertainty can be obtained by using Monte Carlo simulation method and scenario reduction approach. Based on the implicit finite difference scheme, the original partial differential equations of the gas network are discretized to establish an algebraic model, which provide possibility for efficient solution. Then, the nonconvexity caused by the momentum equation is eliminated by the second-order core relaxation. Finally, the stochastic optimal scheduling model is reformulated as a second-order core programming problem. Numerical simulations are performed to showcase the superiority of the established stochastic optimal scheduling model.

KEYWORDS
probabilistic forecasting, stochastic, optimal scheduling, integrated electricity and gas systems, second-order core programming, direct quantile regression

1 Introduction

With the rapid development of natural gas power generation and power-to-gas technology, the power system has established a strong coupling relationship with the gas system (Chen et al., 2021a). Coordinated optimization of integrated electricity and gas systems (IEGS) is of great significance for improving energy efficiency. Besides, natural gas-fired units (NGUs) have the advantage of convenient regulation, which provides a new way to deal with the volatility and intermittency of renewable energy (Shao et al., 2017).

Recently, both academia and industry have been studying the coordinated optimization of IEGS. A dynamic optimal power and gas flow model is developed (Fang et al., 2018) to account for the cushion effects of gas dynamics. By considering the reserve scheduling and renewable uncertainties, a scheduling model of IEGS is formulated (Liu et al., 2019), which focus on enhancing the economic and security of IEGS. A unit commitment model of hybrid power and gas system is proposed (Chen et al., 2019), of which convex envelopes are applied to relax the nonlinear momentum equation. A non-isothermal optimal power and gas flow model is presented (Chen et al., 2021b) to reveal the effects of gas thermodynamics on the power system operation. A model for optimal power and gas flow that takes into account security constraints is developed in (Correa-Posada and Sánchez-Martín, 2014), where linear
sensitivity factors are used to conduct a quick calculation of N-1 contingency of gas pipeline. A model considering operation constraints for unit commitment is put forward in (Liu et al., 2009), of which the benders decomposition method is employed to separately solve the optimization problems of hybrid power and gas systems.

Clean energy, particularly wind power, has made rapid progress with the transformation of the global energy landscape. However, the random changes in wind speed result in strong fluctuations in wind power generation, seriously threatening the stable operation of IEGS. Reference (Alabdulwahab et al., 2017) presents a stochastic dispatch method for hybrid power and gas system that considers the uncertainty of renewable energy and component contingency. A robust dispatch model of IEGS is developed in (Yang et al., 2018), of which a standard gas network method is proposed to consider gas dynamics. A stochastic optimal scheduling approach for IEGS is established in (Zhang and Shahidehpour, 2016) considering hourly electricity demand response. The studies mentioned above all make the assumption that the forecast error of wind power generation conforms to a specific distribution.

However, the stochastic dispatch model of IEGS mainly adopted the point prediction results or the wind power historical data to obtain wind power probability distribution, making it difficult to accurately quantify the time-varying non-stationary wind power prediction uncertainty. Nonparametric probabilistic forecasting can accurately quantify the wind power uncertainty and does not depend on a assumptions regarding the distribution of errors in wind power forecasts, which outperforms parametric probabilistic forecasting with respect to reliability and accuracy (Wan et al., 2014; Wan et al., 2020).

To this end, this paper establishes a stochastic optimal scheduling approach for IEGS, which integrates the advantages of nonparametric probabilistic forecasting to address the randomness of wind power. The method of direct quantile regression (DQR) is applied to generate the quantile at a series of quantile levels. A combination of Monte Carlo simulation and scenario reduction approach is employed to provide a set of representative scenarios by using the predictive quantile. Every representative scenario represents a conglomerate of numerous analogous scenarios. Then, the original partial differential equations of the natural network model are discretized by the implicit difference scheme to establish an algebraic model of the natural gas network. The proposed implicit finite difference scheme has second-order accuracy in both space and time, which provides an accurate approximation of the partial differential equations. The second-order core relaxation eliminates the nonconvexity of the momentum equation, allowing the stochastic optimal scheduling problem to be formulated as a second-order core programming (SOCP) model. This model can be efficiently solved by well-established optimization software. Case studies based on a test system validate the superiority of the developed stochastic scheduling method.

The remainder is organized as follows. The model of IEGS is proposed in Section 2. In Section 3, the formulation of the stochastic scheduling model is presented. The simulation results are shown in Section 4. The paper concludes in Section 5.

## 2 Model of IEGS

### 2.1 Natural gas network model

The flow of natural gas in pipelines is determined by both the momentum equation and the continuity equation. The momentum equation is a description of Newton’s second law (Antenucci and Sansavini, 2018), given as

\[
\frac{\partial \rho v}{\partial x} + \frac{\partial \rho v}{\partial t} + \frac{\partial v^2}{\partial x} + \frac{\rho v}{2d} + \rho g \sin \theta = 0
\]

where \(v\), \(\rho\), and \(p\) represent the velocity, density, and pressure of the gas, respectively; \(\epsilon\) and \(d\) represent the friction factor and the diameter of the pipeline, respectively.

The continuity equation indicates that the natural gas travels continuously along the pipeline (Clegg and Mancarella, 2015), shown as

\[
\frac{\partial p}{\partial t} + \frac{\partial f}{\partial x} = 0
\]

where \(f\) indicates the mass flow rate of pipeline, \(S\) represents the pipeline area.

The state equation expresses the connection between density and pressure, given as

\[
p = \rho c^2
\]

where \(c\) is the speed of sound, which is determined by

\[
c^2 = ZRT
\]

where \(Z\) is the compressibility factor; \(R\) refers to the gas constant; \(T\) indicates the gas temperature.

Based on the steady-state condition, the state variables do not change with time. Thus, the partial differential terms with respect to time equal to 0 and the momentum Eq. 1 as well as the continuity Eq. 2 become (Correa-Posada and Sanchez-Martin, 2015)

\[
\frac{\partial (\rho v)}{\partial t} + \frac{\partial \rho v}{\partial x} + \frac{\partial v^2}{\partial x} + \frac{\rho v}{2d} = 0
\]

\[
\frac{\partial p}{\partial t} + \frac{\partial f}{\partial x} = 0
\]

However, solving Eqs 5, 6 efficiently is challenging due to the form of partial differential. Therefore, the implicit finite difference scheme is introduced to transform the partial differential equations to algebraic equations. The difference scheme is proposed in (Kiuchi, 1994), shown as

\[
\frac{\partial G}{\partial x} = \frac{G^{t+1}_i - G^t_i}{\Delta x} + O(\Delta x^2)
\]

where \(G\) represents the state variables (pressure \(p\) and mass flow rate \(f\)); \(\Delta t\) and \(\Delta x\) denote the temporal and spatial resolution, respectively.

The individual term \(G\) is approximated by

\[
G = \frac{G^{t+1}_i + G^t_i}{2} + O(\Delta x^2)
\]

By substituting Eqs 7, 8 into Eqs 5, 6, the algebraic equations of the natural gas network are formulated as
The flow rate of gas within the pipeline should meet the specified capacity limits:

\[ f_i = f_j = f_p \]  \hspace{1cm} (9)

\[ f_p^2 - \frac{2S_id}{dTRZ} \cdot (p_i^2 - p_f^2) = 0 \]  \hspace{1cm} (10)

where \( f_p^2 \) determine the range of mass flow rate for pipeline \( p \).

The nodal pressure is constrained by

\[ p_i^{\text{min}} \leq p_i \leq p_i^{\text{max}} \]  \hspace{1cm} (12)

where \( p_i^{\text{min}} \) and \( p_i^{\text{max}} \) determine the range of pressure of node \( i \).

The output of gas supplier is limited by

\[ f_{i,t}^{\text{min}} \leq f_{i,t} \leq f_{i,t}^{\text{max}} \]  \hspace{1cm} (13)

where \( f_{i,t} \) is the output of gas supplier \( i \); \( f_{i,t}^{\text{min}} \) and \( f_{i,t}^{\text{max}} \) determine the range of the output for gas supplier \( i \).

Moreover, the gas demand and the gas consumption of NGUs are satisfied by the gas supplier, given as (Zhang et al., 2018)

\[ \sum_i f_{i,j} = \sum_i \left( \sum_j f_{i,j}^{\text{min}} + \sum_j f_{i,j}^{\text{max}} + \sum_p f_{i,p}^{\text{com}} \right) \]  \hspace{1cm} (14)

where \( f_{i,j}^{\text{com}} \) indicates the gas demand of node \( j \); \( f_{i,j}^{\text{min}} \) indicates the gas consumption for natural gas-fired unit \( k \).

The operational pressure of NGUs has restrictions, shown as

\[ P_{i,t}^{G,\text{min}} \leq P_{i,t} \leq P_{i,t}^{G,\text{max}} \]  \hspace{1cm} (15)

where \( P_{i,t}^{G} \) is the pressure of NGU \( i \) at time \( t \); \( P_{i,t}^{G,\text{min}} \) and \( P_{i,t}^{G,\text{max}} \) determine the range of pressure for NGU \( i \).

The proposed natural gas network model contains nonconvex quadratic constraint (10), which poses challenges to solving the model. Thus, second-order cone relaxation is employed to convexify the quadratic constraint (10), given as

\[ f_p^2 - \frac{2S_id}{dTRZ} \cdot (p_i^2 - p_f^2) \leq 0 \]  \hspace{1cm} (16)

Eq. 20 can be transformed into standard SOC form:

\[ \begin{bmatrix} f_p \\ \frac{2S_id}{dTRZ} \cdot p_i \end{bmatrix} \leq 2S_id \hspace{1cm} (17) \]

where \( HP_{p}^{\text{com}} \) is the horsepower of compressor cross pipeline \( p \) at time \( t \); \( \chi_{p}^{\text{com}} \) and \( \nu_{p} \) are the horsepower constants of compressor \( p \); \( GHV \) indicates the gross heating value of natural gas; \( \gamma_p \) refers to the energy conversion coefficient of compressor \( p \); \( k_p^{\text{min}} \) and \( k_p^{\text{max}} \) determine the range of compression ratio of compressor \( p \), respectively; \( f_p^{\text{com, min}} \) and \( f_p^{\text{com, max}} \) determine the range of mass flow rate of compressor \( p \), respectively.

### 2.2 Power system model

The DC power flow model is adopted in this paper to consider the scheduling of active power, expressed as

\[ P_{i,k} = B_k \left( \phi_{i,k} - \phi_{j,k} \right) \]  \hspace{1cm} (22)

where \( P_{i,k} \) indicates the active power flow of the transmission line \( k \) at time \( t \); \( B_k \) indicates the susceptance of transmission line \( k \); \( \phi_{i,k} \) and \( \phi_{j,k} \) denote the voltage angle of bus \( i \) and \( j \), respectively.

In the power system, the power requirements are satisfied by coal-fired generators (CGs) and NGUs, given as

\[ \sum_i \left( \sum_j f_{i,j}^{G} + \sum_j f_{i,j}^{p} \right) = \sum_i p_{i,q}^{C} \]  \hspace{1cm} (23)

where \( p_{i,q}^{C} \) indicates the active power output of CG \( i \); \( p_{j,q}^{C} \) indicates the active power output of NGU \( j \); \( p_{i,q}^{D} \) is the power demand of bus \( q \) at time \( t \).

The DC power system should satisfy the capacity limit, shown as

\[ P_{i,t}^{G,\text{min}} \leq P_{i,t} \leq P_{i,t}^{G,\text{max}} \]  \hspace{1cm} (24)

where \( P_{i,t}^{G,\text{min}} \) and \( P_{i,t}^{G,\text{max}} \) determine the range of active power output of CG \( i \).

The power output of NGU \( i \) should also meet the capacity limit, given as

\[ P_{i,t}^{G,\text{min}} \leq P_{i,t} \leq P_{i,t}^{G,\text{max}} \]  \hspace{1cm} (25)

where \( P_{i,t}^{G,\text{min}} \) and \( P_{i,t}^{G,\text{max}} \) determine the range of active power output of NGU \( i \).

The transmission line imposes constraints on the flow of active power.

\[ P_{k}^{\text{min}} \leq P_{k} \leq P_{k}^{\text{max}} \]  \hspace{1cm} (26)

where \( P_{k}^{\text{min}} \) and \( P_{k}^{\text{max}} \) determine the range of active power flow through transmission line \( k \).

Voltage phase angle is limited by

\[ \phi_{i,t}^{\text{min}} \leq \phi_{i,t} \leq \phi_{i,t}^{\text{max}} \]  \hspace{1cm} (27)

where \( \phi_{i,t}^{\text{min}} \) and \( \phi_{i,t}^{\text{max}} \) determine the range of voltage phase angle of bus \( i \).

### 3 Nonparametric probabilistic forecasting based stochastic scheduling of IEGS

#### 3.1 Objective function

The stochastic optimization algorithm takes into account the probability distribution of uncertain parameters and typically
utilizes the probability density function to depict their variations (Qadran et al., 2014). Afterwards, various scenarios can be generated at random through techniques such as Monte Carlo simulation.

The primary goal of the stochastic scheduling model is to minimize the anticipated operational cost of IEGS, expressed as

$$\min \sum \phi_x \times \left( \sum_{i=1}^{n} a_i x_{i,t}^c + \sum_{j=1}^{m} b_{j,t} x_{j,t}^e \right)$$  \hspace{1cm} (28)

where $\phi_x$ is the occurrence probability of wind power scenario $x$, and $a_i$ and $b_{j,t}$ indicate the cost factors of CG $i$ and gas supplier $j$, respectively.

### 3.2 Nonparametric probabilistic forecasting method

The volatility and sporadic availability of renewable energy create significant challenges to the optimal scheduling of IEGS. The utilization of probabilistic forecasting for wind power can offer critical insights for scheduling of power system in the presence of a substantial amount of wind power.

Typically, probabilistic forecasting methods encompass both parametric and nonparametric approaches. Parametric probabilistic forecasting relies on assumptions regarding the distribution of errors in wind power forecasts. Nonparametric probabilistic forecasting, however, can accurately measure the uncertainty of wind power and is not reliant on a specific probability distribution of forecasting errors, resulting in better reliability and accuracy.

In this paper, the DQR method is utilized to quantify the uncertainty of wind power. The quantile $q^\omega$ can be defined by

$$\Pr(x_t \leq q^\omega) = \omega \hspace{1cm} (29)$$

$$q^\omega = Y_1^{-1}(\omega) \hspace{1cm} (30)$$

where $\Pr ()$ denotes the probability operator; $Y_1$ denotes the cumulative distribution function of wind power; $x_t$ indicates the random variable of wind power; $\omega$ indicates quantile level of the quantile.

Based on the obtained quantile $q^\omega$, a set of predicted quantiles for wind power can be obtained through the nonparametric probability prediction method of DQR, expressed as

$$Y_{\text{q,kg}} = \left\{ q_{\text{q,kg}}^\omega | 0 \leq \omega_1 < \omega_2 < \cdots < \omega_n \leq 1 \right\} \hspace{1cm} (31)$$

where $q_{\text{q,kg}}^\omega$ is the approximation of real quantile $q_{\text{q,kg}}^\omega$; $Y_{\text{q,kg}}$ indicates the predicted quantile with proportions within the range of $0-1$, $y_1$ indicates the probability density function; $Y_1$ indicates the cumulative distribution function.

The DQR method based on the extreme learning machine can transform the probabilistic forecasting into a linear programming model that can be efficiently solved. By introducing the DQR method, the predictive quantile series with proportions can be easily obtained. Given a set of predicted quantiles, the scenarios of wind power output can be obtained by Monte Carlo Simulation. Nevertheless, the straightforward implementation of a significant number of uncertainty scenarios would significantly prolong the computation time. Therefore, the scenario reduction method (Jiang et al., 2020) is applied to reduce the number of scenarios. As an effective tool for the scenario reduction, SCENRED provided by the General Algebraic Modelling System (Zhang et al., 2016) is applied in this paper.

Moreover, for deterministic method that do not consider the probabilistic information involved in wind power uncertainty, spinning reserve should be deployed to address the wind power uncertainty, given as

$$SR_t = \sum_k \left( P_{k,t}^{\text{max}} - P_{k,t}^{\text{e}} \right) \geq SR_{\text{min}}$$  \hspace{1cm} (32)

where $SR_t$ denotes the reserve that can be scheduled; $SR_{\text{min}}$ indicates the minimum reserve of the IEGS.

### 3.3 Model summary

Figure 1 displays the flowchart of the stochastic scheduling model for IEGS. First, a set of predicted quantiles can be obtained by the DQR method. Afterwards, numerous wind power scenarios are created using the Monte Carlo simulation. The scenario reduction method clusters the generated scenarios into several representative scenarios. The nonconvex constraints of gas network are relaxed based on the SOC relaxation. In conclusion, the stochastic scheduling model of IEGS can be expressed as a SOCP model that can be efficiently solved.

### 4 Case study

#### 4.1 System configuration

The topology of the test system is given in Figures 2, 3. The testing system is made up of a gas system with 6 nodes and a power...
A system with 6 buses. The power demands of the test system are satisfied by 1 wind farm, 1 NGU and 1 CG. The wind farm is situated on bus 6, while the NGU is positioned on bus 1 in the power system and node 5 in the natural gas system. The CG is installed at bus 2. All simulations are performed on CPLEX-MATLAB solver on a PC with Inter Core i7 3.60GHz and 32GB RAM. To demonstrate the effectiveness of the proposed stochastic scheduling model for IEGS, two cases are considered.

Case 1: Optimal scheduling of IEGS without stochastic condition.

Case 2: Nonparametric probabilistic forecasting based stochastic optimal scheduling of IEGS.

Case 1 is a deterministic scenario in which spinning reserve is necessary to mitigate the uncertainty of the wind power. In Case 2, the consideration of wind power uncertainty includes the incorporation of probabilistic information.

4.2 Simulation results

Case 1: The method of single-point wind forecasting can be used to obtain wind power generation. The scheduling results for Case 1 are presented in Figure 4. It can be observed that the power demands are satisfied by the CG, NGU, and wind power generation. The CG and NGU adjust their output to ensure
that the wind power can be fully consumed. The natural gas system and the power system are coupled by the NGUs. Hence, the power generation of NGUs has an impact on the equilibrium of the natural gas system. Figure 5 displays the rate of mass flow within pipeline for Case 1. It can be seen that the incoming gas equals to the outgoing gas at each node.

Case 2: The security and economy of IEGS face significant challenges due to the volatility and sporadic availability of renewable energy. The probabilistic information involved in wind power generation is considered in Case 2 to address the uncertainty of renewable energy. Based on the set of predicted quantiles, 3,000 wind power generation scenarios are obtained by Monte Carlo simulation. To reduce computational burden, 3,000 scenarios are reduced to a total of 7 representatives by the scenario reduction method. Seven scenarios are capable of approximating original model.

The obtained 7 representative scenarios of wind power generation are given in Figure 6. The power output of wind farms differs in every scenario. Table 1 lists the probability and total costs of IEGS in different scenario. It can be seen that the total costs of IEGS in scenarios are different from each other. This is due to the fact that every representative scenario is comprised of a combination of comparable scenarios. By utilizing the nonparametric probability forecasting method, the secure operation of power system with high penetration of wind power can be can guaranteed.

The comparison of results for different cases is depicted in Table 2. Due to the different output of wind farm, it can be seen that the operation costs in scenario S2, S3, and S5 are higher than the expected total operation cost. Moreover, it can be easily observed that the expected total operation cost for 7 representative scenarios is $739,172, which is almost 1% lower that of Case 1, validating the significance of stochastic method in cost reduction.

**TABLE 1 Comparison of results for different scenarios in Case 3.**

<table>
<thead>
<tr>
<th>Scenario</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S6</th>
<th>S7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>15.6%</td>
<td>14.95%</td>
<td>12.4%</td>
<td>14.05%</td>
<td>14.0%</td>
<td>13.55%</td>
<td>15.45%</td>
</tr>
<tr>
<td>Obj ($10^5$)</td>
<td>7.3846</td>
<td>7.4299</td>
<td>7.4673</td>
<td>7.3516</td>
<td>7.4049</td>
<td>7.3602</td>
<td>7.3534</td>
</tr>
</tbody>
</table>

**TABLE 2 Comparison of results for different cases.**

<table>
<thead>
<tr>
<th>Case</th>
<th>Case 1</th>
<th>Case 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obj ($10^5$)</td>
<td>7.4861</td>
<td>7.3917</td>
</tr>
</tbody>
</table>

The comparison of results for different cases is depicted in Table 2. Due to the different output of wind farm, it can be seen that the operation costs in scenario S2, S3, and S5 are higher than the expected total operation cost. Moreover, it can be easily observed that the expected total operation cost for 7 representative scenarios is $739,172, which is almost 1% lower that of Case 1, validating the significance of stochastic method in cost reduction.
probability information of wind power uncertainty, the operation strategy of IEGS has better economy and security.

4.3 Larger test system

A larger system composed of a gas system with 40 nodes and a power system with 118 buses is given in Figures 7, 8. Detailed topology data can be found in (Chen et al., 2021b). The Monte Carlo simulation method generates 5,000 stochastic scenarios. The scenario reduction method reduces all stochastic scenarios to 7 representative scenarios.

The probably and total costs of the 7 scenarios are presented in Table 3. The total operation cost in various scenarios is provided in Table 4. The expected total operation cost in Case 2 is $4,935,092, which is 0.94% lower than that of Case 1. By considering the stochastic condition, the IEGS can operate more economically.

5 Conclusion

To address the volatility and sporadic availability of wind power, this paper presents a nonparametric probabilistic forecasting based stochastic optimal scheduling approach for IEGS. The DQR method is used to generate predictive quantile series for wind power. Given a set of predicted quantiles, a combination of Monte Carlo simulation and scenario reduction approach is employed to provide a set of representative scenarios. The original partial differential equations of the natural network model are discretized by the implicit finite difference scheme to establish an algebraic model of the gas network. The nonconvexity of the momentum equation is eliminated by the second-order core relaxationF, and the proposed stochastic optimal scheduling problem is cast into a SOCP model. The simulation results validate that the proposed stochastic scheduling model outperforms the deterministic model, achieving a nearly 1% reduction in total operation cost. The proposed stochastic scheduling model offers even greater benefits in large systems, opening up a new
path towards secure and cost-effective operation of a significant percentage of renewable energy power systems.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Author contributions

JC. Methodology, Writing—original draft. DC. Writing—review and editing.

Funding

The authors declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Science and Technology Project of State Grid Fujian Electric Power Company Limited (52130423000).

Conflict of interest

JC and DC were employed by State Grid Fujian Electric Power Research Institute.

The authors declare that this study received funding from State Grid Fujian Electric Power Company Limited. The funder had the following involvement in the study: data collection and analysis.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References


