Plant-wide byproduct gas distribution under uncertainty in iron and steel industry via quantile forecasting and robust optimization

Sheng-Long Jiang[a,*], Meihong Wang[b], I. David L. Bogle[c]

[a] *College of Materials Scinece and Engineering,Chongqing University, Chongqing, 400044, China*
[b] *Department of Chemical and Biological Engineering, The University of Sheffield, Sheffiled, S13JD, U.K.*
[c] *Department of Chemical Engineering, University College London, London, WC1E 7JE, U.K.*

Abstract

In the modern iron and steel industry, the efficient distribution of byproduct gases faces significant challenges due to quantity- and quality-related uncertainties of gases. This study presents an optimal approach to gas distribution that addresses this issue by incorporating the energy flow network and the uncertain surplus gases from the manufacturing system. The uncertain optimization problem is formulated as a two-stage robust optimization (TSRO) model, including "here-and-now" decisions aimed at minimizing the start-stop cost of energy conversion units, as well as "wait-and-see" decisions aimed at minimizing the operating cost of gasholders and the penalties resulting from energy excess or shortage. To facilitate practical implementation, we propose a "first quantify, then optimize" approach: (1) quantifying the uncertainty of surplus gases via a conditional quantile regression (CDQ)-based T-step time series model, and (2) finding the optimal solution through a column-and-constraint generation algorithm. Furthermore, a case study is conducted on an industrial energy system to validate the proposed methodology. Computational results, using evaluation indicators such as MAPE, RMSE, PICP, and PINAW, confirm the effectiveness of the data-driven time series model in accurately quantifying uncertainties in

*Corresponding author: Sheng-Long Jiang: sh.l.jiang.jx@gmail.com +86 13520412520

each period. Sensitivity analysis demonstrates that the proposed TSRO model achieves a favorable balance between robustness and flexibility by selecting the combination of "budget and quantile" and the parameters of storage and conversion units. Comparative results reveal: (1) the optimal objective of the TSRO closely aligns with that of stochastic programming (SP) and is 2.717 times longer than that of deterministic optimization (DO); and (2) the computation time of TSRO is 2.388 times longer than that of DO, yet significantly smaller than that of SP, being only 0.07 times longer. Consequently, TSRO can efficiently find a robust gas distribution solution with the desired level of conservativeness for integrated iron and steel plants.

Keywords:

Byproduct gas, Robust optimization, Uncertainty quantification, Quantile regression, Iron and steel industry

---

1. Introduction

Given that the global population and living standards are improving, steel demand is expected to continuously grow in the coming decades, especially in developing countries such as China, India, and Brazil [1]. However, the iron and steel industry is also an energy-intensive and high-pollution industry, contributing roughly 8% of energy-related consumption and 6% of global CO_2 emissions [2]. This poses a core challenge that the world must face to achieve a low-carbon and sustainable manufacturing future. China is the largest steel producer worldwide, with a crude steel production volume of more than one billion tons in 2021, accounting for approximately 15% of total greenhouse gas emissions [3].

In an integrated iron and steel plant, the processes from iron ore to steel products primarily rely on coal-related resources and generate a significant amount of byproduct gas, which is identified as the most important secondary energy source and primary source of CO_2 emissions in the plant-wide energy system [4]. These byproduct gases come from three different sources, namely

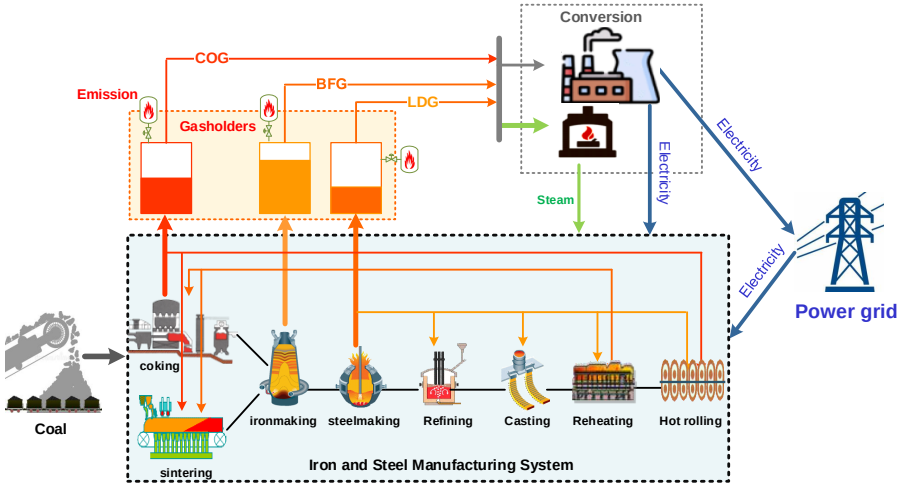


Figure 1: Byproduct gas system in a integrated iron and steel plant.

coke ovens, blast furnaces, and basic oxygen furnaces, and are referred to as coke oven gas (COG), blast furnace gas (BFG), and Linz-Donawitz gas (LDG), respectively. As illustrated in Figure 1, these gases are initially supplied to other production units in the manufacturing system. Then, surplus gases are either stored in dedicated gasholders or converted to other forms of energy, such as steam and electricity, to meet the specific demands of given production units. Generating energy from byproduct gases reduces procurement costs in the market. However, if surplus gases cannot be converted and exceed the capacity of the gasholders, they must be flared, resulting in unnecessary energy loss and environmental pollution. By summarizing these processes, the units are divided into three parts:

1) The manufacturing system, which includes process units, serves as both a supplier and a consumer.
2) The storage system, which includes gasholders, serves as a buffer.
3) The conversion system, which includes boilers, CHP, and CDQ units, serves as an adjusting element.

The primary objective of plant-wide distribution of byproduct gas is to de-

termine the storage and conversion of surplus gases to meet the demands for steam and electricity, while minimizing the environmental pollution from gas emissions. However, the energy system faces uncertainty from the following factors:

1) Quality-related uncertainty, the calories of byproduct gases from the manufacturing system vary because the combustible components (H_2, CH_4, CO) in gases change randomly.
2) Quantity-related uncertainty, the volume of byproduct gases cannot be precisely estimated because the process units in the manufacturing system are not fully controllable.

As these factors are unforeseen, traditional optimal distribution models with perfect assumptions for byproduct gases tend to fail in practice. To overcome these shortcomings, this study presents a robust optimization (RO) method that can provide several flexible strategies to absorb uncertainty and help the byproduct gas system operate with low risks. The contributions of this study can be summarized as follows:

1) Formulating a network flow-based mathematical model to achieve the optimal distribution of byproduct gases in integrated iron and steel plants.
2) Developing a two-stage RO (TSRO) model to make "here-and-now" and "wait-and-see" decisions under uncertainty.
3) Proposing a solution methodology to the TSRO with the "first quantify, then optimize" idea.

The remainder of this study is organized as follows. Section 2 reviews the optimal distribution models in integrated iron and steel plants and the RO technique under uncertainty. Section 3 proposes an optimization-based mathematical programming model for the distribution of byproduct gases and formulates it as a TSRO model under uncertainty. Section 4 reformulates the TSRO as a data-driven two-stage robust optimization (DD-TSRO) model, following the idea of "first quantify, then optimize". Section 5 conducts a case study to verify

the effectiveness of the proposed approach. Section 6 presents the conclusions of the study and provides recommendations for future research.

2. Literature review

The optimal distribution of byproduct gases in integrated iron and steel plants is a complex task that requires simultaneously balancing energy flow and minimizing costs [5]. To achieve this, mathematical modeling approaches, particularly mixed-integer linear programming (MILP), have been widely adopted. The MILP-based distribution model was initially introduced by Akimoto et al. [6], which considered the operational cost of gasholders. Subsequently, MILP-based models have been developed to optimize gas distribution while also considering multiple objectives such as fuel costs [7], varying energy demands [8], switching time of boilers [9], and time-of-use (TOU) electricity pricing [10]. Numerous studies have also adopted the MILP model to describe additional behaviors of the byproduct gas system. For example, Kong et al. [11] focused on the optimal distribution of byproduct gas considering the requirements of the manufacturing system. Zeng et al. [12] assumed that the generation rates of byproduct gases were time-varying. Hu et al. [13] considered the new start-stop behavior of the conversion system. These MILP-based methods have provided valuable insights into energy flow, but they are often criticized for being utopian as they rely on fully known parameters and disregard uncertainty. Consequently, some practitioners have attempted to introduce data-driven techniques to estimate uncertainty related to quantities in the MILP model. For instance, Zhao et al. [14] used a Gaussian kernel-based regression model to predict the uncertain state of outsourcing natural gas and oil, power generation, and gasholder levels. Pena et al. [15] proposed a range of time series models to forecast the generation and consumption of COG, BFG, and LDG, respectively. Subsequently, both of these studies employed a MILP model to minimize operating costs and energy costs using a moving horizon. It is worth noting that all the aforementioned models were solved using off-the-shelf software packages

such as IBM ILOG CPLEX Optimizer and Lingo.

To ensure optimization is achieved more efficiently, several studies have developed heuristic and learning-based methods. Zhao et al. [16] proposed a neural network-based model to predict the real-time flow of COG and its gasholder level. They developed a heuristic algorithm to find the optimal distribution solution. Similarly, Jin et al. [17] applied granular causality techniques and a support vector machine (SVM) model to predict the gasholder level of the LDG system. They developed a particle swarm optimization (PSO) algorithm to find the best solution. However, these methods are only applicable to the distribution problem in a single gas subsystem. To achieve plant-wide distribution, Jin et al. [18] constructed a causal reasoning model to predict the gasholder level and proposed a four-layer causal network involving the operational statuses of boilers, heat quantity, steam demand, and gas mixture proportion to estimate the best solution. Xi et al. [19] focused on the demand for carbon capture, utilization, and storage in iron and steel enterprises. They used a gradient boosting regression tree as the surrogate model of the energy conversion system and the PSO algorithm to determine the optimal distribution decision. To enhance the flexibility of the distribution model, Wang et al. [20] proposed reinforcement learning methods for dynamic distribution, which included a granular prediction model to describe state transitions and a fuzzy rule-based Q function. These studies used data-driven techniques to discover information on quantity-related uncertainty in the byproduct gas system and made more flexible decisions using heuristic and learning-based methods. However, their robustness and applicability are diminished because the uncertainty makes the forecasting results imprecise.

To overcome imprecise predictions in mathematical models, RO is an effective approach for optimization under uncertainty. A comprehensive summary of the development and application of RO can be found in the tutorial by Bertsimas et al. [21], as well as in reviews presented by Gabrel et al. [22] and Rahimian et al. [23]. Rather than precisely predicting uncertain parameters, RO considers a range of possible values within a well-defined "uncertainty set" and seeks robust

and reliable solutions that are optimal under the worst-case scenario within this set. To incorporate risk preferences, various styles of uncertainty sets leveraging historical data have been developed in the RO community [24]. Delage et al. [25] developed a moment-based distributional RO model with mean and variance information. Bertsimas et al. [26] employed hypothesis testing to design uncertainty sets with a probabilistic guarantee and low conservativeness. Shang et al. [27] and Ning et al. [28] derived polyhedron uncertainty sets using machine learning techniques such as SVM, principal component analysis (PCA), and kernel density estimation (KDE). RO models also provide a multi-stage decision process under dynamic environments [29], of which the TSRO is the most representative one. The TSRO model, including here-and-now and wait-and-see decisions, can make decisions as time goes by and has been applied in energy management systems (EMS) for industrial processes [30]. Zhao et al. [31] proposed a TSRO model with KDE-based uncertainty sets to address the operational optimization problem of industrial steam systems under uncertainty. Shen et al. [32] applied a TSRO model with SVM-based uncertainty sets in the energy system of an ethylene plant. In our previous work [33], we developed a TSRO model for the optimal oxygen distribution problem with flexible demands, where the uncertainty set is synthesized via the Gaussian process regression of the iron-making process and the optimal scheduling of the steel-making process.

The existing literature on byproduct gas distribution primarily focuses on addressing quantity-related uncertainties (e.g., gas consumption and gasholder level) and assumes that they are fully known or precisely predictable. However, mathematical models are sensitive to uncertainty, and even small changes in inputs can lead to ineffective gas distribution decisions. This can result in suboptimal adjustments of gasholders and infeasible start-stop operations for conversion units. To fill this research gap, we present a TSRO model for byproduct gas distribution that accounts for uncertain surplus gas from the manufacturing system and develop a solution methodology based on the idea of "first quantify, then optimize".

Nomenclature

Sets,Indices

$\mathcal{K}, k$ Vertex in the energy network

$\mathcal{A}, a$ Arcs in the energy network

$\mathcal{G}, g$ Energy in the energy network

$\mathcal{T}, t$ Time periods of energy distribution

Parameters

$\mathcal{A}^+(k)$, $\mathcal{A}^-(k)$ Input and output arcs of vertex k

a^+, a^- Origin and destination vertex of arc a

$e : a$ Energy of arc a

ω_a Calorie value of the energy flow on arc a

ρ_k conversion efficiency of energy unit on vertex k

$z_{a,t}$ Suplus gas at time pertiod t

$d_{a,t}$ Demand of energy eat time pertiod t

η_k^+ Minimum calorie value of the input energies to conversion unit k

η_k^- Minimum ratio of the ouput energies from conversion unit k to their maximum flows

$\overline{U}_k$ Maximum volume of gasholder k

$\underline{U}_k$ Minimum volume of gasholder k

Δ_k Minimum devivation of gasholder k

$\overline{F}_{k,a}^+$ Maximum flow of input energy a of conversion unit k

$\underline{F}_{k,a}^+$ Minimum flow of input energy a of conversion unit k

$\overline{F}^{-}_{k,a}$ Maximum flow of output energy a of conversion unit k

$\underline{F}^{-}_{k,a}$ Maximum flow of output energy a of conversion unit k

γ_1 Objective coefficient of the start-stop costs

γ_2 Objective coefficient of the gasholders operating costs

$\gamma_{3,k}$ Objective coefficient of surplus and shorages of demand k

Variables

$f_{a,t}$ Continuous variable, which represents the flow of arc a at time period t

$u_{k,t}$ Continuous variable, which represents stroaage level of unit v at time period t

$v_{g,t}$ Continuous variable, which represents shortage or over-stock of energy e at time period t

$O_{k,t}$ Binary variable, which represents the running state (on and off) of conversion unit k at time period t

$S_{k,t}$ Binary variable, which represents the switch state (start up and shut down) of conversion unit k at time period t

3. Optimal distribution under uncertainty

3.1. Network of byproduct gas system

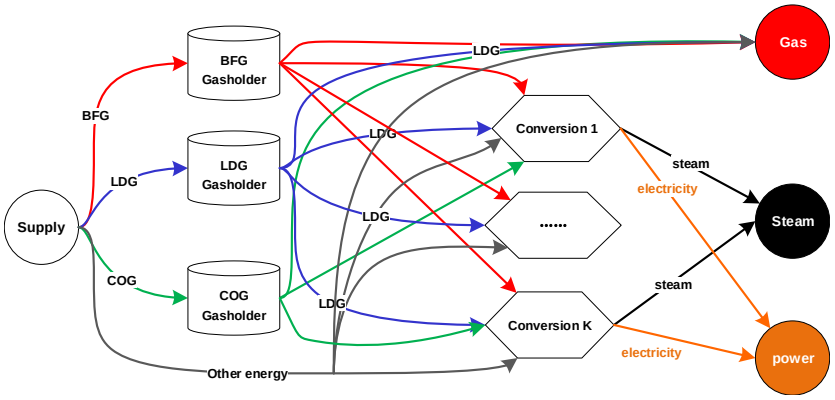


Figure 2: Network flow of byproduct gas system.

This study considers the network of the byproduct gas system $(\mathcal{K}, \mathcal{A})$, where $\mathcal{K}$ denotes the processing units and $\mathcal{A}$ denotes the available flows between these units. Each processing unit is assigned one of the following four roles:

1) Supply ($\mathcal{K}^+$), which supplies surplus gases and other energies () to the system.
2) Storage ($\mathcal{K}^{\mathcal{S}}$), which includes the gasholders of BFG, COG, and LDG.
3) Conversion ($\mathcal{K}^{\mathcal{C}}$), which converts the surplus gases to other energies, i.e., electricity and steam.
4) Demand ($\mathcal{K}^-$), which denotes the energy demands of surplus gases, electricity and steam.

Assuming that the EMS collects all required information, the mathematical model makes optimal decisions for the distribution of byproduct gases over a finite time horizon $\mathcal{T} = \{1, 2, \cdots, t, \cdots, T\}$, where t is the index of a period.

3.2. Deterministic optimization model

3.2.1. Storage units:

In time period t, the surplus gases flow into their dedicated gasholders ($k \in \mathcal{K}^S$), and the varying gasholder level satisfies the mass balance [38]. The

difference between period $t-1$ and t is equal to the surplus volume of byproduct gas minus the total volume of the output flows, which can be described by the following linear equation:

$$u_{k,t} - u_{k,t-1} = \sum_{a \in \mathcal{A}^+(k)} z_{a,t} - \sum_{a \in \mathcal{A}^-(k)} f_{a,t}, \quad k \in \mathcal{K}^S, t \in \mathcal{T} \tag{1}$$

where $u_{k,0}$ is the initial gasholder level.

Since each gasholder must operate within a safe operational region, its level has the following lower and upper bounds:

$$\underline{U}_k \leq u_{k,t}, \leq \overline{U}_k, \quad k \in \mathcal{K}^S, t \in \mathcal{T} \tag{2}$$

Due to operational restraints, the gasholder level change $|u_{k,t} - u_{k,t-1}|$ of gasholder k between periods $t-1$ and t must not exceed Δ_k.

$$u_{k,t} - u_{k,t-1} \leq \Delta_k, u_{k,t-1} - u_{k,t} \leq \Delta_k, \quad k \in \mathcal{K}^S, t \in \mathcal{T} \tag{3}$$

The deviation $v_{k,t}$ of gasholder k from the middle level $u_{k,\text{mid}}$ to the current level $u_{k,t}$ indicates the risk of under-stock or overstock. Thus, the deviation $(v_{k,t})$ can be defined with the following inequalities:

$$u_{k,t} - U_{k,\text{mid}} \leq v_{k,t}, U_{k,\text{mid}} - u_{k,t} \leq v_{k,t}, \quad k \in \mathcal{K}^S, t \in \mathcal{T} \tag{4}$$

3.2.2. Conversion units:

Through a conversion unit (such as a boiler and CHP), input flows can be converted into other output flows, which must satisfy the following input-output balance of energy conversion.

$$\rho_k \sum_{a \in \mathcal{A}^+(k)} f_{a,t} \times \omega_{e:a} = \sum_{a \in \mathcal{A}^-(k)} f_{a,t} \times \omega_{e:a}, \quad k \in \mathcal{K}^C, t \in \mathcal{T} \tag{5}$$

When a conversion unit is turned on $(O_{k,t} = 1)$, its input and output energy flows are restricted by the capacity of conversion unit k.

$$\underline{F}^+_{k,a} O_{k,t} \leq f_{a,t} \leq \overline{F}^+_{k,a} O_{k,t}, \quad k \in \mathcal{K}^C, a \in \mathcal{A}^+(k), t \in \mathcal{T} \tag{6}$$

$$\underline{F}^-_{k,a} O_{k,t} \leq f_{a,t} \leq \overline{F}^-_{k,a} O_{k,t}, \quad k \in \mathcal{K}^C, a \in \mathcal{A}^-(k), t \in \mathcal{T} \tag{7}$$

To ensure that an equipment unit works under normal conditions, the mixed calorific value of the input flows must be greater than or equal to the minimum value of η_k^+.

$$\sum_{a\in\mathcal{A}^+(k)} f_{a,t}\times\omega_{e:a} \geq \eta_k^+ \sum_{a\in\mathcal{A}^+(k)} f_{a,t} \times +bigM(O_{k,t}-1), \quad k\in\mathcal{K}^C, t\in\mathcal{T} \quad (8)$$

Given that the operation of large-capacity equipment with low output is thought to be uneconomic [13], the ratio of output flow to its maximum limits must not be allowed to be lower than the defined threshold value η_k^-. Otherwise, the conversion unit should be closed.

$$\sum_{a\in\mathcal{A}^-(k)} f_{a,t} \geq \eta_k^- \sum_{a\in\mathcal{A}^-(k)} \overline{F}_{k,a}^- + bigM(O_{k,t}-1), \quad k\in\mathcal{K}^C, t\in\mathcal{T} \quad (9)$$

$$\sum_{a\in\mathcal{A}^-(k)} f_{a,t} \leq bigM\times O_{k,t}, \quad k\in\mathcal{K}^C, t\in\mathcal{T} \quad (10)$$

If a conversion unit is being started or stopped ($S_{k,t}$) its on/off status ($O_{k,t}$) is also being changed between periods $t-1$ and t. In a real-world situation, conversion units should operate stably and avoid repeated start-stop changes. Thus, the binary variables $S_{k,t}$ and $O_{k,t}$ are related as follows:

$$O_{k,t}-O_{k,t-1}\leq S_{k,t}, O_{k,t-1}-O_{k,t}\leq S_{k,t} \quad k\in\mathcal{K}^S, t\in\mathcal{T} \quad (11)$$

3.2.3. Demand side:

According to the energy characteristics we divided the energies on the demand side into the following two sets: the emitted gases ($\mathcal{K}_1^-$), including BFG, LDG, and COG; and the generating energy ($\mathcal{K}_2^-$), including electricity and steam. In $\mathcal{K}_1^-$ the gases are not allowed to be emitted into the air; thus, their demands are set to zero, and the excess of supply over demand needs to be minimized. In $\mathcal{K}_2^-$ the shortage of supply under demand will result in a purchase cost. Because the total input energy flow of demand k at period t must meet its demand $d_{k,t}$ we can define the following inequalities:

$$\begin{cases} \sum\limits_{a\in\mathcal{A}^+(k)} f_{a,t} - w_{k,t} \geq d_{k,t}, & k \in \mathcal{K}_1^-, t \in \mathcal{T} \\ \sum\limits_{a\in\mathcal{A}^+(k)} f_{a,t} + w_{k,t} \geq d_{k,t}, & k \in \mathcal{K}_2^-, t \in \mathcal{T}. \end{cases} \tag{12}$$

where $w_{k,t}$ indicates the surplus and shortage of demands in $\mathcal{K}_1^-$ and $\mathcal{K}_2^-$.

3.2.4. Objectives:

Given the assumptions and constraints above, the optimal objective (F) of the studied byproduct gas system within the time horizon ($\mathcal{T}$) can be represented as follows:

$$\min F = \gamma_1 \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}^C} S_{k,t} + \gamma_2 \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}^C} v_{k,t} + \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}^-} \gamma_{3,k} w_{k,t} \tag{13}$$

where the first term (F_1) represents the start-stop cost of conversion units, the second term (F_2) denotes the operating cost of the gasholder deviation from the middle position, and the last term (F_3) indicates the demand penalty caused by the excess or shortage of energies. It should be noted that in (13), γ_1 and γ_2 are unit-unrelated constant coefficients, and $\gamma_{3,k}$ is a unit-related coefficient. Furthermore, γ_2 is represented by a set of piece-wise functions [9].

3.3. TSRO model

It should be noted that, in the gasholder level balance equation (1), the volume of surplus gases ($z_{a,t}$) is equal to the generation of manufacturing units minus their consumption. Given that manufacturing units are not fully controlled, the gas volumes are random and equation (1) needs to hold under the random. In this study, it was assumed that $z_{a,t}$ is an uncertain variable that varies between $[z^\circ_{a,t} - \hat{z}^-_{a,t}, z^\circ_{a,t} + \hat{z}^+_{a,t}]$, where $z^\circ_{a,t}$ is the forecasting point (i.e. nominal value) and $\hat{z}^-_{a,t}, \hat{z}^+_{a,t}$ respectively denote its maximum negative and positive forecasting deviations. Thus, the random surpluses over time can be represented as the following box-uncertainty set.

$$\mathcal{Z}_{\text{box}} := \left\{ \boldsymbol{z} : z_{a,t} = z^\circ_{a,t} + \xi^+_{a,t}\hat{z}^+_{a,t} - \xi^-_{a,t}\hat{z}^-_{a,t}, 0 \leq \xi^-_{a,t}, \xi^+_{a,t} \leq 1, \forall a, t \right\} \tag{14}$$

Since the box-uncertainty set always suffers from over-conservativeness, Bertsimas and Thiele [34] introduced an integer parameter $\Gamma_t (0 \leq \Gamma_t \leq T)$ (budget of uncertainty) to restrict the maximum cumulative deviation with the summation of absolute value constraints.

$$\begin{aligned} \mathcal{Z}_{\text{bud}} := \big\{ \boldsymbol{z} : z_{a,t} = z^{\circ}_{a,t} + \xi^{+}_{a,t} \hat{z}^{+}_{a,t} - \xi^{-}_{a,t} \hat{z}^{-}_{a,t}, \\ 0 \leq \xi^{-}_{a,t}, \xi^{+}_{a,t} \leq 1, \sum_{t=1}^{T} \left(\xi^{+}_{a,t} + \xi^{-}_{a,t} \right) \leq \Gamma_a, \forall a, t \Bigg\} \end{aligned} \tag{15}$$

With the defined uncertainty sets above, a TSRO model can be formulated to make the offline decisions of which conversion units are started or stopped in the first-stage minimization, while the online operating variables can be determined in the second-stage minimization after the potential variations have been realized via the maximization over the uncertainty set $\mathcal{Z}$.

$$\min_{O,S \in \{0,1\}} \left\{ \gamma_1 \sum_{t \in \mathcal{T}} \sum_{k \in \mathcal{K}^C} S_{k,t} + \max_{z \in \mathcal{Z}} \min_{u,v,w,f \in \mathbb{R}} \sum_{t \in \mathcal{T}} \left(\sum_{k \in \mathcal{K}^S} \gamma_2 v_{k,t} + \sum_{k \in \mathcal{K}^-} \gamma_{3,k} \times w_{k,t} \right) \right\} \tag{16}$$

subject to constraints (1)-(12)

4. TSRO formulation and solving

The formulations presented in the previous section provide a comprehensive framework for applying TSRO to byproduct gas distribution.In this section, a reformulation of TSRO into a DD-TSRO model is proposed, following the idea of "first quantify, then optimize". The detailed steps are illustrated in Figure 3.

4.1. Data-driven uncertainty quantification

Since the RO technique is interested in forecasting intervals instead of points, the conditional quantile regression (CQR) proposed by Koenker et al. [35] can be applied to define a time series model. The CQR aims to quantify the uncertain surplus gases at a period τ (z_τ where subscript a is omitted in this section) given data from the past p periods.

$$\tilde{z}_\tau = \psi(z_{\tau-1}, \ldots, z_{\tau-p}; \theta) + \epsilon_\tau \tag{17}$$

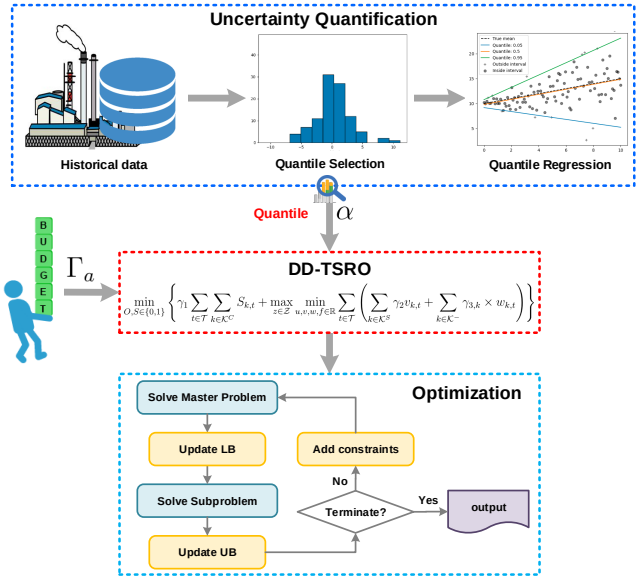


Figure 3: Flowchart of the proposed DD-TSRO framework.

where $\tilde{z}_\tau$ represents the prediction, $\psi(\cdot)$ represents the latent function of unknown form, θ represents its hyper-parameter, and ϵ_τ represents the random noise, but it is not necessarily normally distributed. Instead of forecasting the conditional mean $\mu(\tilde{z}_\tau|z_{\tau-1},\ldots,z_{\tau-p})$, the CQR technique can be applied to estimate a given quantile of $\tilde{z}_\tau$, such as the median. Let L_α be the αth quantile of the cumulative density function of $\tilde{z}_\tau$, i.e. $\Pr(\tilde{z}_\tau \leq L_\alpha) = \alpha$, and the αth conditional quantile function of z_τ can be rewritten as:

$$q_\alpha(\tilde{z}_\tau|z_{\tau-1},\ldots,z_{\tau-p}) := \inf\{z_t \in \mathbb{R} : \Pr(\tilde{z}_\tau|z_{\tau-1},\ldots,z_{\tau-p}) \geq \alpha\} \tag{18}$$

Given the training set $\mathcal{D} = \{(z_{\tau-1}^{(i)},\ldots,z_{\tau-p}^{(i)}; z_\tau^{(i)})\}_{i=1}^N$, classical regression analysis estimates the conditional mean by minimizing the sum of squared residuals on the training set. Analogously, CQR estimates a conditional quantile function q_α of z_τ by minimizing the following loss function:

$$\mathcal{L} = \min_\theta \frac{1}{N}\sum_{i=1}^{N}\left[\alpha/2\max(z_\tau^{(i)} - \tilde{z}_\tau^{(i)}, 0) + (1-\alpha/2)\max(z_\tau^{(i)} - \tilde{z}_\tau^{(i)}, 0)\right] \tag{19}$$

When $\alpha < 0.5$, the loss function places more weight on smaller predictions, and when $\alpha > 0.5$, it places more weight on larger predictions. When $\alpha = 0.5$,

the loss function is reduced to the least absolute deviation regression, which is also known as median regression. Thus, given the training set $\mathcal{D}$ and quantile α, the uncertainty of z_t can now be quantified:

$$\begin{cases} z_\tau^\circ = q_{0.5}(\tilde{z}_\tau | z_{\tau-1}, \ldots, z_{\tau-p}) \\ \hat{z}_\tau^- = z_\tau^\circ - q_\alpha(\tilde{z}_\tau | z_{\tau-1}, \ldots, z_{\tau-p}) \\ \hat{z}_\tau^+ = q_{1-\alpha}(\tilde{z}_\tau | z_{\tau-1}, \ldots, z_{\tau-p}) - z_\tau^\circ \end{cases} \tag{20}$$

In this study, gradient boosting decision tree (GBDT) is applied to find the best latent function $\psi(\cdot)$ to cover the training data. GBDT is a powerful machine learning technique that uses multiple regression trees (e.g. CART). The prediction model can be rewritten as follows:

$$\psi_M(z_{\tau-1}, \ldots, z_{\tau-p}) = \sum_{m=1}^{M} tree(z_{\tau-1}, \ldots, z_{\tau-p}; \Theta_m) \tag{21}$$

where $tree(\cdot;\cdot)$ represents a regression tree, Θ_m denotes parameters of tree m, and M represents the number of regression trees. Then, GBDT uses a gradient boosting and a forward step-wise algorithm to find the optimal hyperparameters.

$$\Theta_m^* = \arg\min_{\Theta_m} \mathcal{L}(\tilde{z}_\tau, \psi_{m-1} + tree(z_{\tau-1}, \ldots, z_{\tau-p}; \Theta_m) \tag{22}$$

The details can be found in the work of [36]. This study also proposes a framework for implementing T-step prediction, as shown in the following:

$$\tilde{z}_{\tau+t} = \psi_M^t(z_{\tau-1}, \ldots, z_{\tau-p}, \alpha), t \in 1, ..., T \tag{23}$$

4.2. Reformulation for TSRO

For formulation simplicity, the TSRO model is presented in the following matrix form:

$$\min \quad \mathbf{c}^\top \mathbf{x} + \max_{\mathbf{z}\in\mathcal{Z}} \min_{y\in\Omega(\mathbf{x},\mathbf{z})} \mathbf{d}^\top \mathbf{y} \tag{24a}$$

$$s.t. \quad \mathbf{A}\mathbf{x} \leq \mathbf{b}, \mathbf{x} \in \{0,1\} \tag{24b}$$

$$\Omega(\mathbf{x},\mathbf{z}) = \begin{cases} \mathbf{G}\mathbf{y} \geq \mathbf{h} \\ \mathbf{Q}\mathbf{y} \geq \mathbf{r} - \mathbf{P}\mathbf{x} \\ \mathbf{W}\mathbf{y} = \mathbf{z} + \mathbf{s} \end{cases} \tag{24c}$$

In formulation (24), vector $\mathbf{x}$ represents the first-stage decisions (O, S), and vector $\mathbf{y}$ represents the second-stage decisions (u, v, w, f). The objective function (24a) is divided into two parts: one depends on the binary variables $\mathbf{x}$, and the other depends on the continuous variables $\mathbf{y}$. Equation (24b) includes all constraints involving only binary variables (6),(7), (10) and (11).In the domain defined by equation (24c), the first term collects constraints (2)-(5) only involving continuous variables, the second term accounts for (8) and (9) that involving mixed variables, and the last term represents the constraints (1) that involves the uncertain gas surplus.

With the simplified formulation, the dual of the second-stage optimization problem is first found to be LP and holds a strong duality. Therefore, its dual problem can be rewritten in the following form:

$$\begin{aligned} \max_{\mathbf{z},\boldsymbol{\lambda},\boldsymbol{\sigma},\boldsymbol{\phi}} \quad & \boldsymbol{\lambda}^\top \mathbf{h} + \boldsymbol{\sigma}^\top (\mathbf{r} - \mathbf{P}\mathbf{x}) + \boldsymbol{\phi}^\top (\mathbf{z} + \mathbf{s}) \\ s.t. \quad & \boldsymbol{\lambda}^\top \mathbf{G} + \boldsymbol{\sigma}^\top \mathbf{Q} + \boldsymbol{\phi}^\top \mathbf{W} = \mathbf{d}^\top \\ & \boldsymbol{\lambda} \geq 0, \boldsymbol{\sigma} \geq 0, \mathbf{z} \in \mathcal{Z} \end{aligned} \tag{25}$$

where $\boldsymbol{\lambda}, \boldsymbol{\sigma}, \boldsymbol{\phi}$ are the Lagrangian multipliers of the formula (24c) and $\boldsymbol{\phi}$ is unbounded. Note that $\boldsymbol{\phi}^\top \mathbf{z}$ are bilinear in the objective function, and therefore needs to be linearized. First, due to the variable z being independent of other variables in the equation (25), the optimal solution must be one of the extreme

points of $\mathcal{Z}$. Next, the uncertain set is inserted into the term ($\boldsymbol{\phi}^\top \mathbf{z}$) to obtain the following equation:

$$\boldsymbol{\phi}^\top \mathbf{z} = \sum_{t=1}^{T} \left(\phi_t z_t^\circ + \phi_t \xi_t^+ \hat{z}_t^+ - \phi_t \xi_t^- \hat{z}_t^- \right)$$

Then, (25) is transformed to the following MILP model using the Big-M method [37], which can be easily solved by commercial solvers.

$$\begin{aligned}
\max_{\boldsymbol{\pi},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\phi}} \quad & \boldsymbol{\lambda}^\top \mathbf{h} + \boldsymbol{\sigma}^\top (\mathbf{r} - \mathbf{P}\mathbf{x}) + \boldsymbol{\phi}^\top \mathbf{s} + \sum_{t=1}^{T} \left(\phi_t z_t^\circ + \pi_t^+ \hat{z}_t^+ + \pi_t^- \hat{z}_t^- \right) \\
s.t. \quad & \boldsymbol{\lambda}^\top \mathbf{G} + \boldsymbol{\sigma}^\top \mathbf{Q} + \boldsymbol{\phi}^\top \mathbf{W} = \mathbf{d}^\top \\
& \pi_t^+ \leq bigM \xi_t^+, \pi_t^+ \leq \phi_t + bigM \left(1 - \xi_t^+\right), \forall t \\
& \pi_t^- \leq bigM \xi_t^-, \pi_t^- \leq bigM \left(1 - \xi_t^-\right) - \phi_t, \forall t \\
& \sum_{t=1}^{T} \left(\xi_t^+ + \xi_t^- \right) \leq \Gamma, \xi_t^+, \xi_t^- \in \{0,1\} \\
& \boldsymbol{\lambda} \geq 0, \boldsymbol{\sigma} \geq 0
\end{aligned} \tag{26}$$

where π_t^+ and π_t^- are auxiliary variables. If ξ_t^+ and ξ_t^- are equal to 1, π_t^+ and π_t^- will be limited to ϕ_t and $-\phi_t$. If ξ_t^+ and ξ_t^- are equal to 0, π_t^+ and π_t^- will be limited 0.

4.3. Column-and-constraint generation algorithm

$$\max_{\mathbf{x},\beta,\mathbf{y}_i | i<I} \quad \mathbf{c}^\top \mathbf{x} + \beta \tag{27a}$$

$$s.t. \quad \mathbf{A}\mathbf{x} \leq \mathbf{b}, \mathbf{x} \in \{0,1\} \tag{27b}$$

$$\beta \geq \mathbf{d}^\top \mathbf{y}_i, i = 1, \ldots, I \tag{27c}$$

$$\mathbf{G}\mathbf{y}_i \geq \mathbf{h}, i = 1, \ldots, I \tag{27d}$$

$$\mathbf{Q}\mathbf{y}_i \geq \mathbf{r} - \mathbf{P}\mathbf{x}, i = 1, \ldots, I \tag{27e}$$

$$\mathbf{W}\mathbf{y}_i = \mathbf{z}_i + \mathbf{s}, i = 1, \ldots, I, \tag{27f}$$

It is difficult to find the optimal solution of (24) directly within a short time using state-of-the-art optimization solvers (e.g. Cplex, Gurobi, and SCIP).

Since the uncertainty set $\mathcal{Z}$ is a polyhedron, the worst-case scenarios of surplus gas are only located at an extreme point of $\mathcal{Z}$; therefore, the number of possible worst-case scenarios is finite [38]. Based on this characteristic, this study applied the column-and-constraint generation (C&CG) algorithm proposed by Zeng and Zhao [38] to seek the optimal solution for DD-TSRO. Initially, the TSRO defined in (24) is relaxed and reformulated as the master problem (27). Then, the possible worst-case scenarios are identified by solving the subproblem (26) and added to (27) as a cutting plane. Finally, the optimal solution can be found by iteratively improving the gap between the lower and upper bounds. The solving procedure is stated in Algorithm 1.

Algorithm 1 Column-and-Constraint Generation (C&CG)

1: Initialization: LB $= -\infty$, UB $= \infty$, $i = 1$
2: Construct a feasible decision at the first stage $\mathbf{x}_1$ (via the deterministic model (1)-(13))
3: **repeat**
4: Solve subproblem (26) with fixed $\mathbf{x}_i$, obtain the optimal solution ξ_i' and the optimal objective value β'. Then, a worst-case scenario $\mathbf{z}_i$ is constructed via the budget-based set (15).
5: Update UB $= \min\{\text{UB}, \mathbf{c}^\top \mathbf{x}_\text{i} + \beta'\}$
6: Solve the master problem defined in (27) by adding the new scenario $\mathbf{z}_i$ and new variables $\mathbf{y}_i$. Let $(\mathbf{x}^*, \beta^*)$ be the optimal solution.
7: Update LB $= \mathbf{c}^\top \mathbf{x}^* + \beta^*$.
8: $i \leftarrow i + 1, \mathbf{x}_i \leftarrow \mathbf{x}^*$.
9: **until** UB $-$ LB $\leq \delta$ or $i > I$ (δ is a user-defined tolerance)

Output: The optimal solution $\mathbf{x}^*$, LB and UB.

5. Case study

To verify the proposed approach, this section presents a case study on the energy system of an integrated iron and steel plant in China. The integrated

iron and steel plant consists of three gasholders, two 35t boilers, two 130t boilers, two sets of CHP, and two sets of CDQ. The operational parameters and the network configuration of the industrial energy system are presented in the attached supplementary materials (ref. Tables S1-S6).

For uncertainty quantification, the open-source scikit-learn package, version 1.2.1 [39], was used. The DD-TSRO model was implemented using Pyomo [40], a Python-coded, open-source modeling language, and solved using "Gurobi 9.0" (`https://www.gurobi.com/products/gurobi-optimizer/`), an academic version with default settings. All computational studies were conducted on a PC with an Intel Core i7 processor (3.60 GHz), 16 GB RAM, and a Windows 10 operating system.

5.1. Uncertainty quantification

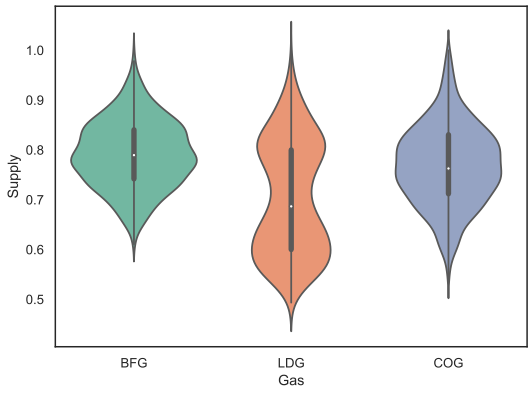


Figure 4: The violin plot of gas surpluses on dataset.

First, we extracted 1000 historical data samples for each gas from the EMS of the iron and steel plant and shared them on the following webpage: `https://github.com/janason/Energy/tree/main/gas`. The violin plot shown in Figure 4 illustrates the visual distribution of normalized samples and suggests that the volume of surplus LDG follows a bimodal distribution. To quantify the uncertain surplus gases, we applied the GBDT technique using scikit-learn's *GradientBoostingRegressor* class. Specifically, we separated all the samples into

training (900 samples) and testing (100 samples) sets and used them to test the CDQ-based T-step time series model.

Next, we set $T = 8$ and $p = 20$ in the proposed time series model. When forecasting the nominals, we set $\alpha = 0.5$, and when forecasting the deviations, we enumerated the quantile α from the set $\{0.01, 0.05, 0.1\}$. Thus, we obtained $8 \times 3 = 24$ groups of prediction results for each type of gas. Figure 5 shows the predictions for $t = 1$, where the "plus" points represent medians, and the gray regions represent intervals of uncertain gases quantified by each quantile. The graphical representation shows that the width of the gray region increases as the quantile decreases, implying that the uncertainty quantification of gases in each period is determined by selecting the appropriate quantile.

Then, we observed the effects of quantile selection by calculating the following four evaluation indicators: (1) mean absolute percentage error (MAPE), (2) root mean square error (RMSE), (3) prediction interval coverage probability (PICP), and (4) prediction interval normalized average width (PINAW):

$$\text{MAPE} = \frac{1}{N}\sum_{\tau=1}^{N}\left|\frac{z_\tau^\circ - \tilde{z}_\tau}{\tilde{z}_\tau}\right| \times 100\% \tag{28}$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{\tau=1}^{N}\left(z_\tau^\circ - \tilde{z}_\tau\right)^2} \tag{29}$$

$$\text{PICP} = \frac{1}{N}\sum_{\tau=1}^{N} Cov_\tau^{(\alpha)} \tag{30}$$

where $Cov_\tau^{(\alpha)} = \begin{cases} 1, & \text{if} \quad z_\tau \in (z_\tau^\circ - \hat{z}_\tau^-, z_\tau^\circ + \hat{z}_\tau^+), \\ 0, & \text{otherwise} \end{cases}$

$$\text{PINAW} = \frac{1}{N(z_{\max} - z_{\min})}\sum_{\tau=1}^{N}(\hat{z}_\tau^- + \hat{z}_\tau^+) \tag{31}$$

where $z_{\min}$ and $z_{\max}$ is the minimum and maximum values of the target value in the testing set, respectively.

Finally, we presented the calculated indicators in Table 1. The results indicate the following: (1) When $\alpha = 0.50$, the forecasted nominals for each gas

are not completely precise, as indicated by their values of MAPE and RMSE being greater than zero. Additionally, they exhibit similar effects, as their values of MAPE and RMSE are close. (2) The values of PICP and PINAW for smaller quantiles are lower, emphasizing the significance of quantile selection for quantifying the uncertainty of the proposed DD-TSRO model.

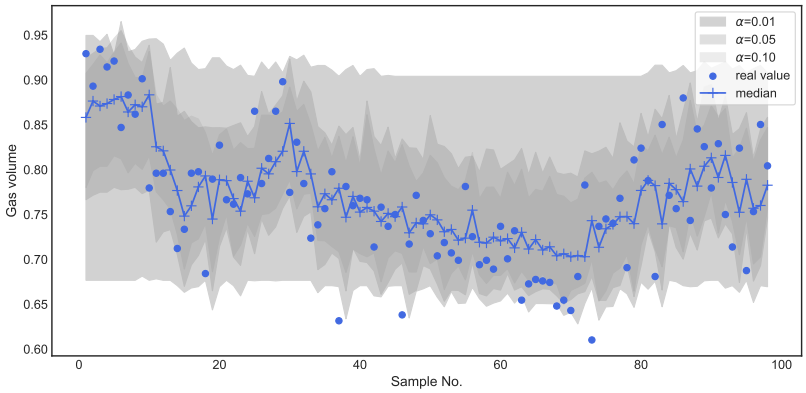
By utilizing the validation set (refer to the attached data file on the webpage), we obtained forecasted nominals $z^{\circ}_{a,t}$ and deviations $z^{-}_{a,t}$, $z^{+}_{a,t}$ for eight periods (see Table S7 in the supplemental material), as illustrated in Figure 6. These values were then used to parameterize the budget-based uncertainty set $\mathcal{Z}$ in the model.

Table 1: Metrics of prediction results

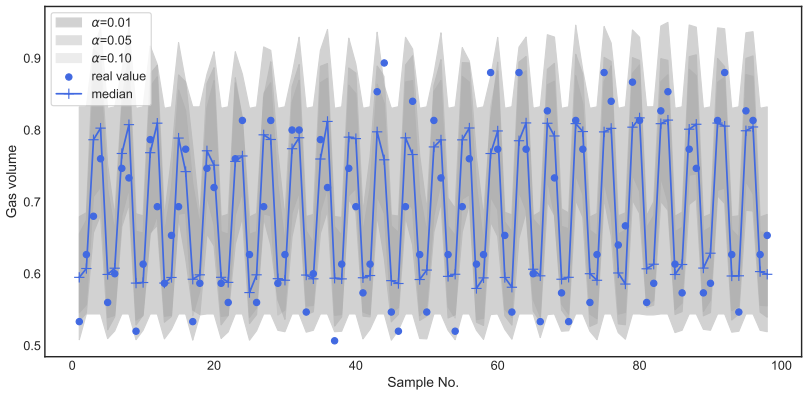
metrics	α	BFG	LDG	COG
MAPE	0.50	6.418	6.488	6.804
RMSE	0.50	0.059	0.056	0.057
	0.01	0.941	0.969	0.981
PICP	0.05	0.825	0.897	0.862
	0.10	0.680	0.784	0.791
	0.01	0.850	0.852	1.011
PINAW	0.05	0.525	0.481	0.550
	0.10	0.392	0.367	0.436

5.2. Sensitivity analysis on Γ *and* α

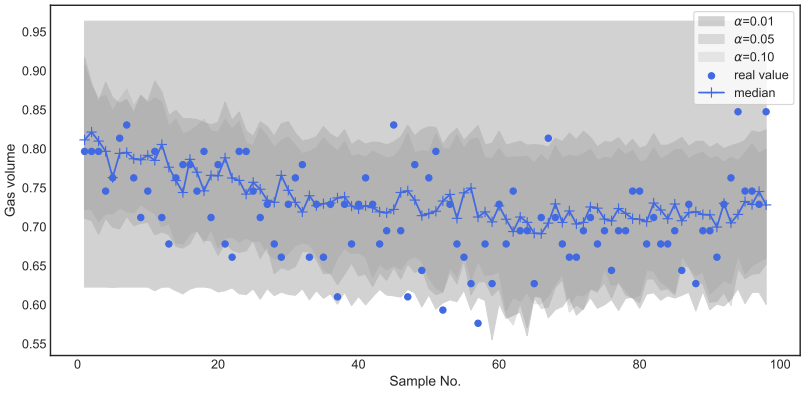
As discussed in Section 3, the quantile α reflects the estimated level of uncertainty of surplus gases in each period, and the budget Γ_a controls how much the uncertain parameters can deviate from their nominal values. In this subsection, a set of scenarios was generated to observe how the optimal objective is affected by budgets and quantiles. Then, the appropriate combination of (α, Γ_a) was determined to avoid over-conservativeness, with the quantile α falling into $\{0.01, 0.05, 0.10\}$, and the budget Γ_a of each gas ranging from 0 to 8. When $\Gamma = 0$, TSRO is defined as a deterministic optimization, representing the most



(a) BFG

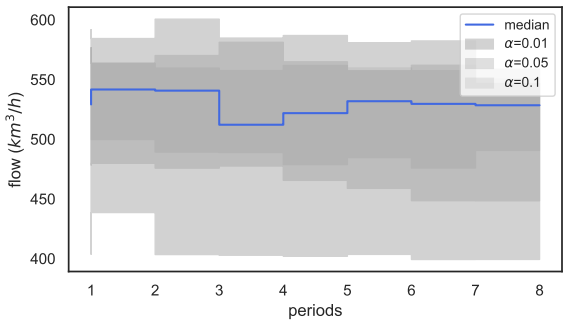


(b) LDG

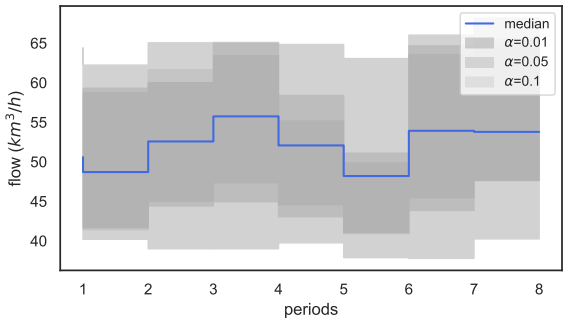


(c) COG

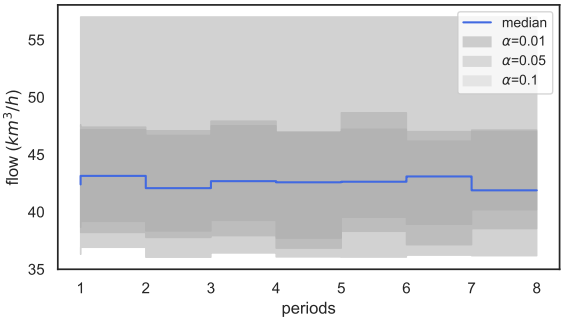
Figure 5: Prediction results for gases when $\tau = 1$



(a) BFG



(b) LDG



(c) COG

Figure 6: Prediction intervals of byproduct gases.

optimistic scenario without any uncertainty. When $\Gamma = 8$, TSRO is defined as an uncertain optimization with a box-based set, representing the most pessimistic scenario.

Figure 7 shows the curves of optimal objectives with different budgets (Γ_a) and quantiles (α). With an increase in the budget of uncertainty, the optimal objectives under all quantiles worsened, as higher budgets represent more uncertain parameters deviating from their nominal values. Specifically, the optimal objectives are at their worst and deteriorate (increase) at the most rapid rate when $\alpha = 0.01$ because it exhibits the most conservative estimation of $z_{a,t}$. To achieve a trade-off between the probability of violation and the effect on the optimal objective, $\alpha = 0.05$ and $\Gamma_a = 4$ were chosen for the subsequent experiments.

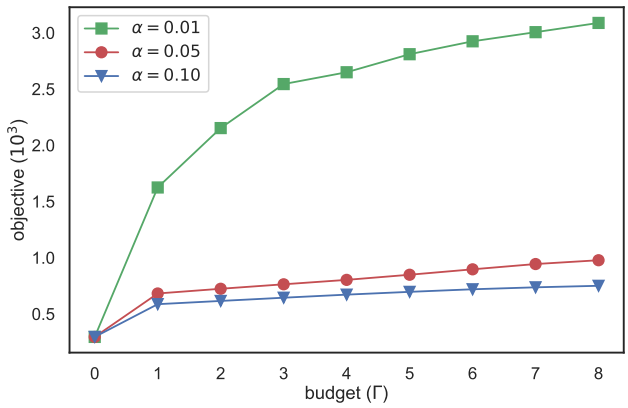


Figure 7: Optimal objective under different budgets and quantiles.

5.3. Sensitivity analysis on Δ_k and η_k^-

In the studied TSRO model, the maximum deviation of each gasholder (Δ_k) and the minimal output ratio of each conversion unit (η_k^-) reflect the flexibility of the byproduct gas system as they can absorb some uncertainties. In the following experiments, we varied their values respectively by $[\Delta_k] = \iota \times \Delta_k$ and $[\eta_k^-] = \iota \times \eta_k^-$, where ι represented the scaling ratio ranging from 0.5 to 2.0 with

a step size of 0.1.

Figure 8 displays the varied objectives output by the TSRO model with different values of $[\Delta_k]$ (left y-axis) and $[\eta_k^-]$ (right y-axis). These curves confirm the following observations:

(1) Increasing $[\Delta_k]$ or decreasing $[\eta_k^-]$ enhances the flexibility of the energy system, as indicated by the decreasing optimal objectives.

(2) The energy system is more sensitive to changes in $[\eta_k^-]$ compared to $[\Delta_k]$. For instance, when $\iota = 1.5$, the objective of $[\eta_k^-]$ is ten times higher than that of $[\Delta_k]$. Moreover, for $\iota \geq 1.6$, the objective of $[\eta_k^-]$ sharply increases.

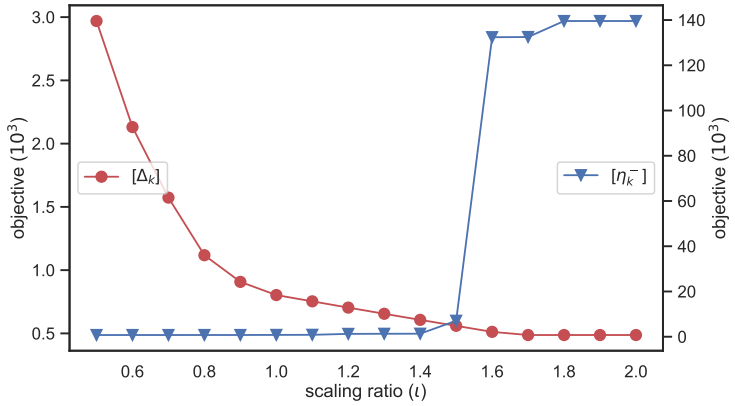


Figure 8: Optimal objective under varied Δ_k and η_k^-

5.4. Comparison with deterministic optimization

To further explore the behaviors of the TSRO model ($\alpha = 0.05, \Gamma_a = 4$), we compared its optimal objectives with that of the deterministic optimization (DO) model ($\Gamma_a = 0$).

Table 2 lists the objectives of the TSRO and RO models, while Figures 9-11 illustrate their solutions for decision variables $O_{k,t}$, $u_{k,t}$, and $v_{e,t}$, respectively. The x-axis represent conversion units, gasholders, and generating energies, the y-axis denote periods, and the z-axis mark their values in each period. These comparative results suggest the following:

(1) The TSRO model has a lower start-stop cost (F_1 in Table 2) since TSRO only shuts down one unit, while RO shuts down two units (as shown in Figure 9), providing adjusting capacity to absorb uncertainties.

(2) The TSRO model incurs significantly higher operating costs for gasholders (F_2 in Table 2) and higher demand penalties (F_3 in Table 2). Figure 10 shows that the gasholder level of TSRO declines more significantly, while Figure 11 shows that more energies have excess and shortage. This is because the TSRO model computes the optimal objectives under the worst-case scenario.

(3) Computation time in Table 2 shows that the TSRO model takes 2.388 times longer than the DO model to solve due to the iterative nature of the C&CG algorithm.

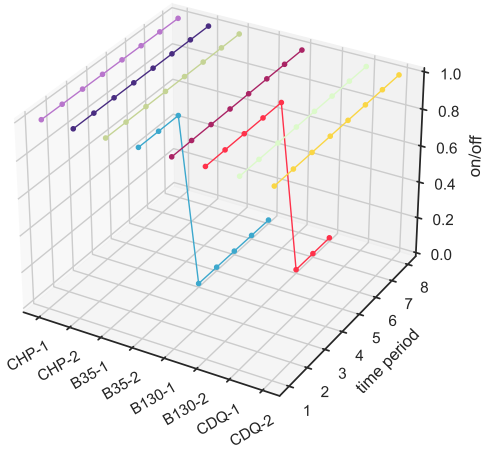
From the above observations, it can be concluded that the TSRO model provides a robust optimal solution considering the worst-case scenario, despite the acceptable cost of increased computation time. This is evidenced by its objective (F) being 2.717 times that of the DO model, and its computation time being within 5 seconds.

Table 2: Comparison results between DO and TSRO model

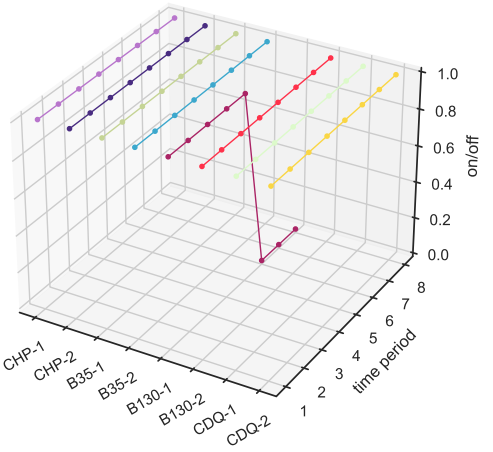
items	DO	TSRO
F_1	20	10
F_2	221.934	371.669
F_3	53.667	421.371
F	295.601	803.040
Computation Time (s)	1.727	4.124

5.5. Comparison with stochastic programming

In this subsection, we compared our TSRO approach with the two-stage stochastic programming (TSSP) approach in solving the test instance. It is

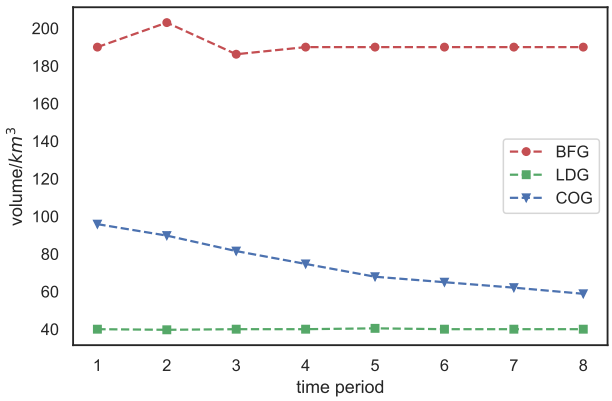


(a) DO

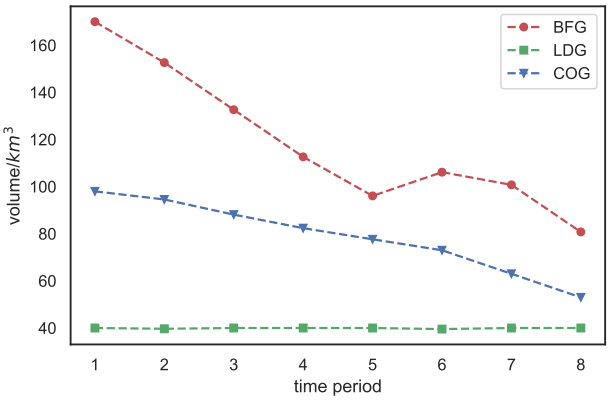


(b) TSRO

Figure 9: On-off status of conversion units.

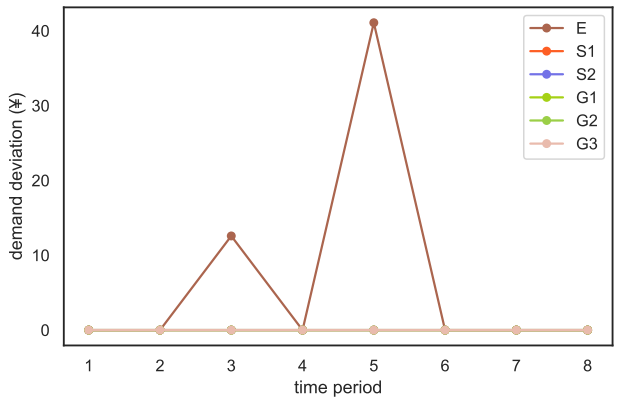


(a) DO

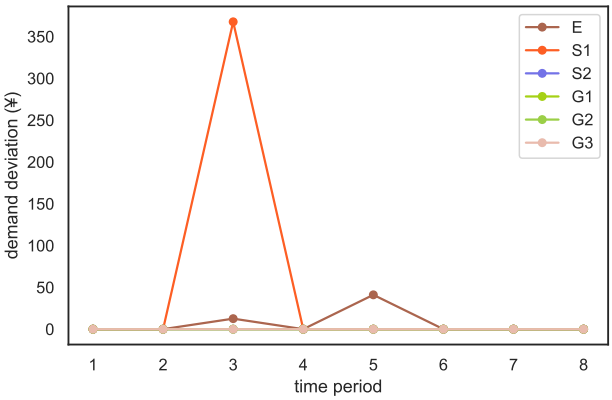


(b) TSRO

Figure 10: Level change of gasholders.



(a) DO



(b) TSRO

Figure 11: Demand penalties caused by energy excess or shortage .

worth noting that the TSSP approach has been widely used in handling uncertainty in previous works on energy management (see [41] and its cited references).

In our TSRO model, we set $\alpha = 0.05$ and $\Gamma_a = 4$. For the TSSP model, we generated 400 scenarios from a truncated normal distribution within the prediction interval. These scenarios were based on the mean z_τ° and standard deviation $(z_\tau^+ + z_\tau^-)/4$, representing the uncertainty of surplus gases. We then used a validation dataset consisting of 800 scenarios, which were randomly drawn from the same distributions as those used in the TSSP model, to compute the objectives of both models.

The simulated results are summarized in Table 3. It is worth noting that the TSRO model incurs a lower start-stop cost (F_1 in Table 3) because it adopts a conservative approach to handle uncertainties, avoiding frequent switching of conversion units. Both models yield similar results for F_2 and F_3 under average and worst-case scenarios. Moreover, the average and worst-case objectives (F) of the TSSP and TSRO models are similar, which indicates that the TSRO model can achieve the desired conservativeness. Notably, the TSRO model runs significantly faster, taking only 0.07 times longer than the stochastic approach.

Table 3: Comparison results between TSSP and TSRO

items	TSSP	TSRO
F_1	20	10
Worst-case F_2	305.696	311.417
Average F_2	229.637	240.199
Worst-case F_3	388.814	388.814
Average F_3	86.337	86.337
Worst-case F	642.277	645.053
Average F	335.974	336.536
Computation Time(s)	65.431	4.592

6. Conclusions

This study presents an optimal gas distribution approach to address the challenge of uncertain surplus gas from the manufacturing system. The proposed approach utilizes a TSRO model that includes "here-and-now" decisions to minimize the start-stop cost of conversion units and "wait-and-see" decisions to minimize the operating costs of gasholders and demand penalties. To facilitate practical implementation, this study also introduces a "first quantify, then optimize" method: using a data-driven time series model to quantify the uncertainty of surplus gas and then applying a C&CG algorithm to find the optimal solution. A comprehensive case study conducted on an industrial energy system of an iron and steel plant provides significant insights, yielding the following key findings:

(1) Computational results of quantified indicators, such as MAPE, RMSE, PICP, and PINAW, validate the effectiveness of the CDQ-based T-step time series model in accurately quantifying uncertainties for each period by selecting the appropriate quantile.

(2) The sensitivity analysis reveals that the proposed TSRO model ensures robustness and flexibility by selecting suitable "budget and quantile" combinations and setting the parameters of storage and conversion units, such as the maximum deviation of gasholders and minimum output ratio. The optimization of these parameters can enhance the performance of the energy system in dealing with uncertainties under various conditions.

(3) Comparative analysis highlights that the TSRO model can obtain a robust solution with low conservativeness while maintaining an acceptable computation speed. Compared to the DO model, the TSRO model achieves a 2.717 times higher objective and takes 2.388 times longer to solve. However, when compared to the stochastic approach, the TSRO model achieves comparable optimal objectives with a computing time that is only 0.07 times as long.

Despite the robustness and low conservativeness offered by the proposed approach, this study still has some limitations. The study primarily focuses on the uncertainty of the surplus volume of byproduct gases and does not account for quality-related uncertainty, particularly the calorific value of each gas. Future research directions should involve developing a novel uncertainty set that incorporates both uncertain surplus and calorific value. Additionally, as the uncertainty information evolves over time, it is necessary to extend the two-stage decision process to a multi-stage decision process.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (No. 61873042).

References

[1] C. Gahm, F. Denz, M. Dirr, A. Tuma, Energy-efficient scheduling in manufacturing companies: A review and research framework, European Journal of Operational Research 248 (3) (2016) 744–757.

[2] Z. Fan, S. J. Friedmann, Low-carbon production of iron and steel: Technology options, economic assessment, and policy, Joule 5 (4) (2021) 829–862.

[3] L. Ren, S. Zhou, T. Peng, X. Ou, A review of co2 emissions reduction technologies and low-carbon development in the iron and steel industry focusing on china, Renewable and Sustainable Energy Reviews 143 (2021) 110846.

[4] H. Zhang, W. Sun, W. Li, G. Ma, A carbon flow tracing and carbon accounting method for exploring co2 emissions of the iron and steel industry: An integrated material–energy–carbon hub, Applied Energy 309 (2022) 118485.

[5] W. Sun, Q. Wang, Y. Zhou, J. Wu, Material and energy flows of the iron and steel industry: Status quo, challenges and perspectives, Applied Energy 268 (2020) 114946.

[6] K. Akimoto, N. Sannomiya, Y. Nishikawa, T. Tsuda, An optimal gas supply for a power plant using a mixed integer programming model, Automatica 27 (3) (1991) 513–518.

[7] J. Kim, H.-S. Yi, C. Han, A novel MILP model for plantwide multiperiod optimization of byproduct gas supply system in the iron-and steel-making process, Chemical Engineering Research and Design 81 (8) (2003) 1015–1025.

[8] J. Kim, H. Yi, C. Han, C. Park, Y. Kim, Plant-wide multiperiod optimal energy resource distribution and byproduct gas holder level control in the iron and steel making process under varying energy demands, in: Computer Aided Chemical Engineering, Vol. 15, Elsevier, 2003, pp. 882–887.

[9] X. Zhao, H. Bai, X. Lu, Q. Shi, J. Han, A MILP model concerning the optimisation of penalty factors for the short-term distribution of byproduct gases produced in the iron and steel making process, Applied energy 148 (2015) 142–158.

[10] X. Zhao, H. Bai, Q. Shi, X. Lu, Z. Zhang, Optimal scheduling of a byproduct gas system in a steel plant considering time-of-use electricity pricing, Applied energy 195 (2017) 100–113.

[11] H. Kong, E. Qi, H. Li, G. Li, X. Zhang, An MILP model for optimization of byproduct gases in the integrated iron and steel plant, Applied Energy 87 (7) (2010) 2156–2163.

[12] Y. Zeng, X. Xiao, J. Li, L. Sun, C. A. Floudas, H. Li, A novel multi-period mixed-integer linear optimization model for optimal distribution of byproduct gases, steam and power in an iron and steel plant, Energy 143 (2018) 881–899.

[13] Z. Hu, D. He, Operation scheduling optimization of gas–steam–power conversion systems in iron and steel enterprises, Applied Thermal Engineering 206 (2022) 118121.

[14] J. Zhao, C. Sheng, W. Wang, W. Pedrycz, Q. Liu, Data-based predictive optimization for byproduct gas system in steel industry, IEEE Transactions on Automation Science and Engineering 14 (4) (2016) 1761–1770.

[15] J. G. C. Pena, V. B. de Oliveira Junior, J. L. F. Salles, Optimal scheduling of a by-product gas supply system in the iron-and steel-making process under uncertainties, Computers & Chemical Engineering 125 (2019) 351–364.

[16] J. Zhao, Q. Liu, W. Wang, W. Pedrycz, L. Cong, Hybrid neural prediction and optimized adjustment for coke oven gas system in steel industry, IEEE transactions on neural networks and learning systems 23 (3) (2012) 439–450.

[17] F. Jin, L. Wang, J. Zhao, W. Wang, Q. Liu, Granular-causality-based byproduct energy scheduling for energy-intensive enterprise, IEEE Transactions on Automation Science and Engineering 17 (4) (2020) 1662–1673.

[18] F. Jin, J. Zhao, Z. Han, W. Wang, A joint scheduling method for multiple byproduct gases in steel industry, Control Engineering Practice 80 (2018) 174–184.

[19] H. Xi, X. Wu, X. Chen, P. Sha, Artificial intelligent based energy scheduling of steel mill gas utilization system towards carbon neutrality, Applied Energy 295 (2021) 117069.

[20] T. Wang, J. Zhao, Q. Xu, W. Pedrycz, W. Wang, A dynamic scheduling framework for byproduct gas system combining expert knowledge and production plan, IEEE Transactions on Automation Science and Engineering 20 (1) (2022) 541–552.

[21] D. Bertsimas, D. B. Brown, C. Caramanis, Theory and applications of robust optimization, SIAM review 53 (3) (2011) 464–501.

[22] V. Gabrel, C. Murat, A. Thiele, Recent advances in robust optimization: An overview, European journal of operational research 235 (3) (2014) 471–483.

[23] H. Rahimian, S. Mehrotra, Distributionally robust optimization: A review, arXiv preprint arXiv:1908.05659 (2019).

[24] C. Ning, F. You, Optimization under uncertainty in the era of big data and deep learning: When machine learning meets mathematical programming, Computers & Chemical Engineering 125 (2019) 434–448.

[25] E. Delage, Y. Ye, Distributionally robust optimization under moment uncertainty with application to data-driven problems, Operations research 58 (3) (2010) 595–612.

[26] D. Bertsimas, V. Gupta, N. Kallus, Data-driven robust optimization, Mathematical Programming 167 (2018) 235–292.

[27] C. Shang, X. Huang, F. You, Data-driven robust optimization based on kernel learning, Computers & Chemical Engineering 106 (2017) 464–479.

[28] C. Ning, F. You, Data-driven decision making under uncertainty integrating robust optimization with principal component analysis and kernel smoothing methods, Computers & Chemical Engineering 112 (2018) 190–210.

[29] H. Bakker, F. Dunke, S. Nickel, A structuring review on multi-stage optimization under uncertainty: Aligning concepts from theory and practice, Omega 96 (2020) 102080.

[30] H. Qiu, W. Gu, P. Liu, Q. Sun, Z. Wu, X. Lu, Application of two-stage robust optimization theory in power system scheduling under uncertainties: A review and perspective, Energy (2022) 123942.

[31] L. Zhao, C. Ning, F. You, A data-driven robust optimization approach to operational optimization of industrial steam systems under uncertainty, in: Computer aided chemical engineering, Vol. 46, Elsevier, 2019, pp. 1399–1404.

[32] F. Shen, L. Zhao, W. Du, W. Zhong, F. Qian, Large-scale industrial energy systems optimization under uncertainty: A data-driven robust optimization approach, Applied Energy 259 (2020) 114199.

[33] S.-L. Jiang, G. Peng, I. D. L. Bogle, Z. Zheng, Two-stage robust optimization approach for flexible oxygen distribution under uncertainty in integrated iron and steel plants, Applied Energy 306 (2022) 118022.

[34] D. Bertsimas, A. Thiele, A robust optimization approach to inventory theory, Operations research 54 (1) (2006) 150–168.

[35] R. Koenker, G. Bassett Jr, Regression quantiles, Econometrica: journal of the Econometric Society (1978) 33–50.

[36] J. H. Friedman, Greedy function approximation: a gradient boosting machine, Annals of statistics (2001) 1189–1232.

[37] Y. Guo, C. Zhao, Islanding-aware robust energy management for microgrids, IEEE Transactions on Smart Grid 9 (2) (2016) 1301–1309.

[38] B. Zeng, L. Zhao, Solving two-stage robust optimization problems using a column-and-constraint generation method, Operations Research Letters 41 (5) (2013) 457–461.

[39] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Scikit-learn: Machine learning in python, the Journal of machine Learning research 12 (2011) 2825–2830.

[40] W. E. Hart, J.-P. Watson, D. L. Woodruff, Pyomo: modeling and solving mathematical programs in python, Mathematical Programming Computation 3 (2011) 219–260.

[41] G. Mavromatidis, K. Orehounig, J. Carmeliet, Design of distributed energy systems under uncertainty: A two-stage stochastic programming approach, Applied energy 222 (2018) 932–950.