# Data-Driven Multistage Scheduling Optimization for Refinery Production under Uncertainty: Systematic Framework, Modeling Approach, and Application Analysis

Biao Han*

**Abstract**

The widespread existence of various uncertainties makes the inherently complex refinery production scheduling problem even more challenging. To address this issue, this paper proposes a viable systematic data-driven multistage scheduling optimization framework and develops a corresponding structured modeling methodology. Under this paradigm, unit-level advanced control and plant-level intelligent scheduling are coordinated to jointly deal with endogenous and exogenous uncertainties while ensuring the implementability of the solution. Historical production data are leveraged to learn the operation mode-specific process models and to train the overall stochastic dynamic scheduling model. Operation mode indicative matrices with associated "shift registers" are introduced to describe different operational status of the production units and also the transition processes induced by operational switching. The model is established based on the policy graph representation and its tractability can be achieved by resorting to the stochastic dual dynamic programming algorithm, although it is mathematically a tricky mixed integer model. The scheduling policy is allowed to be trained according to different risk preferences of the decision-makers and can be applied to online interactive sequential scheduling over finite or infinite horizons. Model refining and policy retraining mechanisms based on human feedback equip the scheduling policy with good generalization capability and evolutionary adaptability. An industrial-scale application case study is conducted based on Julia language, and the in-sample and out-of-sample Monte Carlo simu-

*Corresponding author.
*Email address:* hanb16@tsinghua.org.cn (Biao Han)

lations demonstrate the feasibility and effectiveness of the proposal in this paper.

*Keywords:*
refinery scheduling, uncertainty, data-driven, stochastic programming, reinforcement learning

---

## 1. Introduction

Petroleum is a vital resource on the earth. It not only profoundly influences the development of global politics and economy, but is also closely related to the daily lives of ordinary people. Petroleum refining industry plays a key role in converting crude oil into required fuels and other necessities (Larraz, 2021). The scheduling optimization level of refining production directly affects the resource consumption and pollution emission level of the refinery (Joly et al., 2002; Joly, 2012; Khor and Varvarezos, 2017). Even a small improvement in production scheduling can bring significant benefits to the refinery (Yuan et al., 2017). Therefore, refinery production scheduling optimization has received increasing attention from both academia and industry (Joly et al., 2018).

However, the production scheduling optimization of a refinery is far from a trivial issue. Since the entire refining process is very complex, involving a variety of production units and processing procedures, how to establish a proper model is first of all a challenging problem. Scholars have tried to establish process models of different production units by means of mechanism analysis and process simulation (Zhang et al., 2001; Menezes et al., 2013; Alhajri et al., 2008; Gao et al., 2015). Based on various characterization methods of process units, many endeavors on refinery production scheduling optimization modeling have been made (Pinto et al., 2000; Göthe-Lundgren et al., 2002; Luo and Rong, 2007; Shah and Ierapetritou, 2011). As a matter of fact, an optimized schedule makes sense only if it can be well implemented in practice, while few literatures discuss the realization of production unit operations. In addition, due to the large inertia characteristic of the continuous production process industry, the transition process caused by unit operation switching can hardly be ignored. Following this spirit, various modeling methods have been developed successively and have achieved expected results (Shi et al., 2014; Zhang et al., 2015; Chen and Jiang, 2024). The refinery production scheduling model is usually formulated as a mixed

integer programming in mathematics, and their efficient solution is another challenging problem. Especially when it is hoped that the scheduling scheme can be effective over a longer scheduling horizon, the existing models need to consider a larger number of time intervals, which will lead to a disastrous increase in the scale of the programming problem, making it difficult for classical algorithms and off-the-shelf solvers to give a solution of ideal quality within the allowed time. Therefore, techniques such as decomposition algorithms (Jia and Ierapetritou, 2004; Shah et al., 2009; Shi et al., 2015; Li et al., 2025) and heuristic algorithms (Chen et al., 2021, 2024) have emerged to improve the solution efficiency. Nevertheless, with the increasingly urgent requirements for long-term and fast-response scheduling, these stopgap measures are ultimately difficult to escape the "curse of dimensionality", and are therefore not suitable for effective application in actual production scheduling, especially for online scheduling scenarios. Besides, more importantly, none of the above-mentioned literatures considers the uncertainties in the refining production process at a systematic level, which further hinders their industrial application.

In fact, uncertainties are widespread amoung the volatile and dynamic refinery production environment, such as unstable material prices, fluctuating process parameters, unpredictable customer demands, etc (Leiras et al., 2011; Al-Shammari and Ba-Shammakh, 2011; Yang and Barton, 2016). It has been recognized that even a small perturbation in the nominal value of the parameter may cause the optimization result based on the deterministic model to become suboptimal or completely infeasible. There are two main routes commonly used for handling uncertainties in mathematical optimization, namely stochastic programming (Birge and Louveaux, 2011) and robust optimization (Ben-Tal et al., 2009). Because of a popular critique of stochastic programming that the probability distribution of the uncertainty must be known *a priori*, the robust optimization route seems to have gained more preference in refining production optimization in recent years (Wang et al., 2021; Zhang et al., 2022; Liu et al., 2022). However, robust optimization is not a "free lunch". It only offers the worst-case performance guarantee due to the extreme risk aversion, which will lead to overly conservative and impractical scheduling results, bringing unnecessary extra costs to the refining production (Bertsimas and Sim, 2004). Nowadays, the big data era is witnessing the explosion of data availability in many fields particularly the petroleum refining industry, where there are plenty of data yielded from the previous operations and laboratory evaluations, but they are far

from being fully mined and leveraged (Yang and Barton, 2016; Joly et al., 2018). These data can not only be used for process modeling, but also for learning the distribution of uncertainty, which paves the way for the use of stochastic programming. One benefit of stochastic programming is that it can well incorporate different risk measures, such as expectation, conditional value-at-risk (CVaR) (Rockafellar and Uryasev, 2002), etc., so as to conveniently control the conservatism of scheduling results. Objectively speaking, scholars have recognized that a more flexible modeling scheme that aligns naturally with the refining production process under uncertainty should be multistage sequential decision-making (Li and Grossmann, 2021; Zhang et al., 2022; He et al., 2023). It supports efficient adaptive recourse decisions as uncertainties are revealed gradually over a long time horizon, based on a pretrained policy, which enhances decision transparency and interoperability for decision-makers (Han et al., 2021; Han, 2024; Ou and Bi, 2024; Al-Shafei et al., 2025). And stochastic programming can be well extended to this dynamic decision-making mode (Dowson, 2020; Zhang et al., 2023; Guevara et al., 2024). Despite these advantages, the gap between mathematical theory and engineering practice has limited the broad application of multistage stochastic programming in refinery production scheduling problems (Li and Grossmann, 2021).

In order to bridge this gap, this paper focuses on the refinery production practice, harnesses the process data resource, and inaugurates a novel viable multistage scheduling solution for oil refining production under uncertainty. The contribution of this paper mainly lies in the following aspects.

- A data-driven evolvable systematic framework integrating control and scheduling is proposed to address the endogenous and exogenous uncertainties in refinery production scheduling optimization. The framework can be effectively applied to dynamic refinery production scheduling over finite or infinite horizons, and the risk preference of decision-makers can be adjusted as needed.

- Under the above framework, a multistage stochastic dynamic scheduling modeling methodology for refining production process is developed based on the policy graph representation. The operation mode indicative matrices and their corresponding "shift registers" are introduced to describe the operation status changes of the processing units, and the dynamic characteristics caused by operational switching are also taken into account.

- An industrial-scale application case study of refinery production scheduling is presented along with rich graphical results. The application case is fully implemented in Julia, a new promising open-source high-performance dynamic high-level programming language, which is rarely used in the existing predominant refinery scheduling literatures, thus pioneering its popularization amoung the practitioners in related fields.

Note that the underlying logic of our proposed framework has a lot in common with the more prevailing reinforcement learning (RL) (Bertsekas, 2019; Sutton and Barto, 2018). When training the scheduling policy, we are essentially performing Q-learning where the value function is approximated using mathematical programming duality (Dowson, 2020). And the evolution of the framework is driven by the reinforcement learning from human feedback (RLHF) mechanism (Kaufmann et al., 2024; Lambert, 2025). Therefore, this paper organically applies the idea of RL to systematically cope with the dynamic scheduling decision-making for the entire large-scale, long-horizon and strongly-coupled refinery production process with various uncertainty sources, which, to the best of our knowledge, remains largely unexplored in published scholarly works to date.

The rest of this paper is organized as follows. Section 2 presents a brief overview on the refinery prodution process. In Section 3, based on the discussion on the uncertain nature of the refining production process, the systematic control and scheduling framework of this paper is proposed. As the backbone of this framework, we provide the multistage stochastic dynamic scheduling modeling methodology in detail in Section 4. Section 5 studies the application of the proposed framework on a practical industrial-scale refinery production scheduling case. We finally conclude the paper in Section 6.

## 2. Brief Overview on Refinery Production Process

A typical oil refining production process is shown in Figure 1, which is a simplified flowchart of the production process of a real fuel oil refinery in North China. From the production procedure point of view, the whole oil refining process can be divided into three main parts: feeding of crude oils, production and processing of component oils, and blending and delivery of product oils.

*Feeding of Crude Oils.* Feeding of crude oils is the first procedure of the whole refinery production process. Before this, there are often crude oil manage-

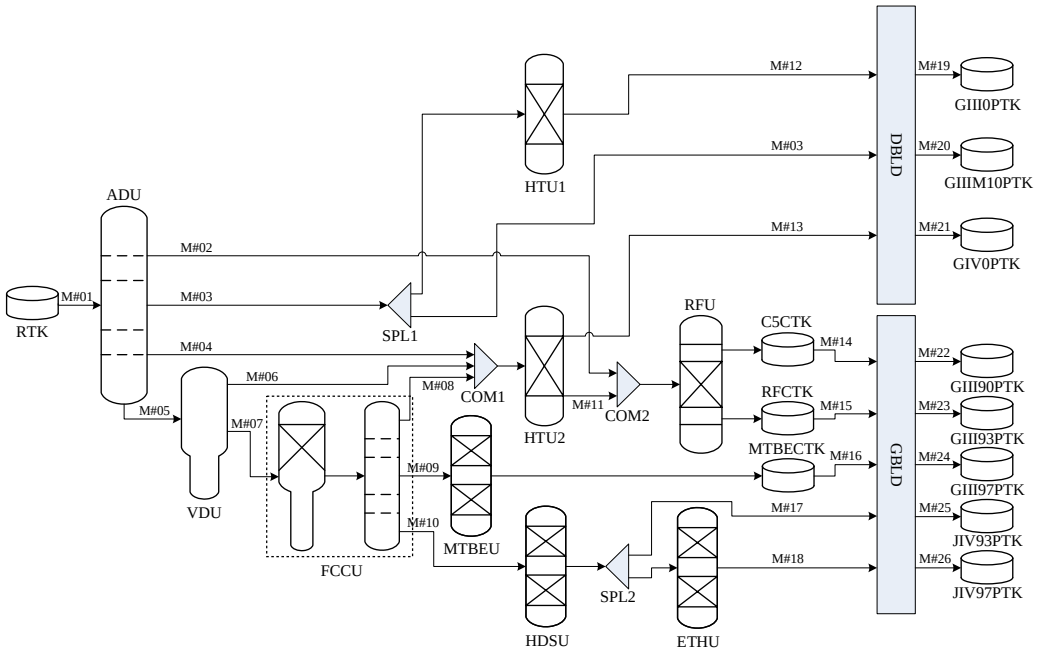


Figure 1: Simplified Flowchart of A Typical Refinery Production Process[1]

ment and scheduling procedures such as crude oil procurement planning, oil tank inventory management of crude oil entering the terminal and the plant area, and management of crude oil blending and supply tanks. For the refinery production process studied in this paper, we assume that the previous crude oil management and scheduling ensure the stability of crude oil properties and sufficient supply. Therefore, the crude oil feeding procedure can be simplified to transporting the crude oil from the raw material storage tank with uniform property and sufficient reserve to the primary processing unit.

*Production and Processing of Component Oils.* Different processing units process the raw oil step by step to produce several key component oils, which is the second and core procedure of the refining production process. Typical processing units (PRUs) mainly include: atmospheric distillation units (ADUs), vacuum distillation units (VDUs), fluid catalytic cracking units (FCCUs), hydrodesulfurization units (HDSUs), etherification units (ETHUs),

[1]The specific meanings of the labels and indices in the figure are listed in Appendix A.

hydrotreating units (HTUs), methyl tert-butyl ether production units (MTBEUs), catalytic reforming units (RFUs), etc. Each unit processes its input material under certain operating conditions, produces output materials with properties that meet certain standards at corresponding yields, and transports them to subsequent units through pipelines (PPLs) for further processing, while incurring certain operating cost. When the operating condition of a unit changes, its product yields, property index values, and operating cost will also change accordingly.

*Blending and Delivery of Product Oils.* After crude oil is processed by the above-mentioned units, several key component oils for product oil blending can be obtained. These component oils are directly transported to blending units (BLDs) through pipelines for product oil blending, or they will be transported to component oil tanks (CTKs) for buffering if they have stable properties, so that they can be used at any time during the blending process. Blending units are generally divided into gasoline blending units (GBLDs) and diesel blending units (DBLDs). The key component oils are blended by them according to relevant quality standards stipulated by the state to form different types of gasoline product oils and diesel product oils, which are then stored in their respective product oil tanks (PTKs) and delivered and sold according to market demand.

From a systematic point of view, all the objects involved in the whole oil refining production process mentioned above can be generally abstracted into 8 categories according to their functional characteristics, namely, raw material tanks (RTKs), processing units (PRUs), component oil tanks (CTKs), blenders (BLDs), product oil tanks (PTKs), pipelines (PPLs), and dummy units combiners (COMs) and splitters (SPLs) (Han et al., 2022). Among them, RTK is a type of object located at the most upstream position of the production process. It is the source of all materials and can be regarded as an object with only a single output but no input. At the most downstream end, the market is the sink of all materials that has multiple inputs but no output. The other units between the two are objects with both input and output materials, among which, PRU can be regarded as an object with single input and multiple outputs, BLD can be regarded as an object with multiple inputs and multiple outputs, and CTK, PTK and PPL can all be regarded as objects with single input and single output (the output end of PTK flows directly to the market and is omitted in Figure 1). The dummy units COM and SPL are responsible for material collection and separation,

and they can be regarded as multiple-input single-output and single-input multiple-output objects respectively.

In summary, the whole oil refining production process involves a wide variety of production units, and the same type of production unit often has multiple instances. All production units are interconnected through their multiple inputs and outputs, thus forming a intricate production and processing network.

## 3. Uncertainty Sources and Systematic Coping Framework

### *3.1. Uncertain Nature of Refining Production Process*

In the refinery production process shown in Figure 1, the entire production process has been abstracted and simplified into a network consisting of several production units with specific functions. In fact, each abstract production unit may actually contain more than one key production device entity. Taking FCCU as an example, it generally includes a reaction-regeneration system and a main distillation tower, and here we just consider them as a whole. In addition, each key device usually also has some additional devices as auxiliary, which do not need to be considered at the scheduling level. In other words, each production and processing unit itself is an object with complex dynamic characteristics. During the production process, the operator intervenes in the production unit through the process control system based on experience, so that the process unit can switch between different working points as needed, which naturally leads to diversity and uncertainty in the operation status of the process unit. On the other hand, unlike general motion control systems, the refinery production process unit, as a typical process control system, often has a relatively large time lag and a relatively slow response speed. Therefore, the frequent and intensive switching of the unit operating conditions will not only cause fluctuations that are difficult to determine in the unit itself, but also cause fluctuations in its associated downstream units and even the entire production process. Compared with some classic discrete manufacturing processes, the oil refining process does not have a limited number of processing methods and task states that are naturally separated. Due to the continuity nature of the oil refining process variables, its processing methods and operating states may present infinite possibilities. All of the above make it difficult to clearly describe and accurately estimate the operation modes, transition times, output yields and

various property index values as well as the operation costs of the oil refining production processing units.

In addition to the processing unit, other production units are also affected by uncertainties. First, the price of crude oil supplied by the upstream raw material tank and the prices of product oils sold to the market may change in the future. may change in the future. This market price fluctuation directly affects the cost and revenue of the refinery. Secondly, the storage cost of component oils and product oil storage tanks will also show certain uncertainty as the storage conditions change, such as temperature changes, oil volatilization, etc., which will affect the storage cost. Compared with other manufacturing processes, the refining production process has a more significant feature. Its final products are very diverse and directly connected to the market. The demand is often very intensive and compact. Changes in demand have a profound and unpredictable impact on production and processing. It is a crucial high-dimensional uncertainty factor that significantly affects the overall economic benefits of the entire refinery. This market price fluctuation directly affects the cost and revenue of the refinery. Secondly, the storage costs of component oil and product oil storage tanks will also become uncertain as the storage conditions change, such as temperature changes, oil volatilization, etc. The final products of the oil refining production process are very diverse and directly connected to the market. The demand is often very intensive and compact. The change in demand has a far-reaching and difficult to predict impact on production and processing. It is a crucial high-dimensional uncertainty factor that significantly affects the overall economic benefits of the entire refinery.

It can be seen that complex uncertainties exist widely in all parts of the refining production process, which brings great difficulties to the modeling and implementation of the control and scheduling of the entire refining production process, and even forms an irreconcilable deadlock that hinders each other — the uncertainties at the control level will cause the mismatch of the scheduling model and lead to the inaccurate implementation of the scheduling optimization solution, while the uncertainties at the scheduling level will in turn aggravate the chaos of the control system (Gao et al., 2016). Therefore, production scheduling has never been an isolated and simple mathematical programming problem, but a systematic and practical comprehensive problem. Only by fully considering and effectively dealing with various uncertainties and taking into account all levels of the production process with systematic thinking can we obtain a scheduling solution with practical and

implementable significance (Andrés-Martínez and Ricardez-Sandoval, 2022; Gao et al., 2016; Shobrys and White, 2002).

### *3.2. Systematic Control and Scheduling Framework*

From the perspective of scheduling optimization, the uncertainties in the refining production process discussed above can be summarized into two categories: endogenous uncertainties and exogenous uncertainties (Zhang and Feng, 2020; Feng et al., 2021; Motamed Nasab and Li, 2021; Lappas and Gounaris, 2018). The endogenous uncertainties include the duration of the transition process caused by the operator switching the operation mode of the processing unit, the yield of each output of the processing unit under a specific operation state, the various property index values of each output, and the operation cost of the unit, while the exogenous uncertainty include the market price of the crude and product oils, the storage costs of component and product oil tanks, and the market demand for each product oil.

This paper proposes a systematic solution framework that is to deal with the endogenous uncertainties with unit-level advanced control and to deal with the exogenous uncertainties with plant-level scheduling optimization. Specifically, the implementation of unit-level advanced control enables the unit to operate under a limited number of stable optimized operation modes (Lyu et al., 2010; Gao et al., 2014, 2015). These potential operation modes can be discovered through data analytics on the large amount of operation data accumulated by the process control system (PCS) using some data mining methods (Gao, 2014; Gao et al., 2016). These operation modes often have clear practical meanings. For example, Gao (2014); Shi et al. (2014) pointed out that ADU has two stable operation modes termed as gasoline mode and diesel mode, which in practice correspond to the operation states of maximizing the gasoline component and diesel component in the output materials respectively. Under a certain stable operation mode, the endogenous uncertain parameters such as the output yields, output property index values and operation cost of the unit will show some specific patterns. Based on the unit-level advanced control, these finite number of discrete realizable stable operation modes constitute the basic unit operation strategies for the plant-level scheduling optimization.

At the plant level, the historical operational data such as oil prices, inventory costs, and demand accumulated by the manufacturing execution system (MES) provide support for the characterization of exogenous uncertainties. Then we are able to establish an integrated model for plant-level scheduling

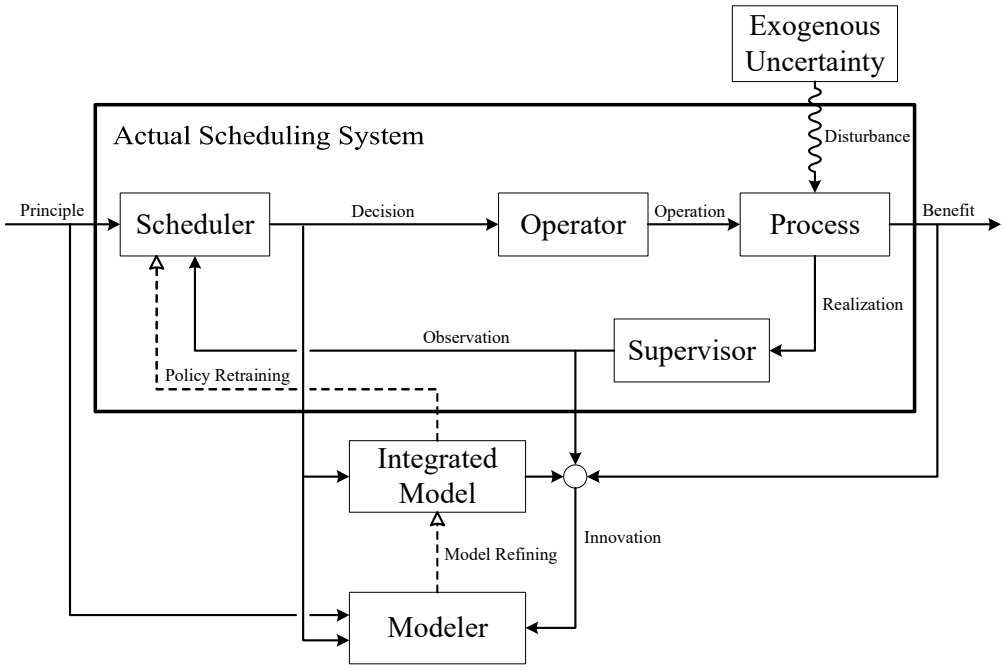


Figure 2: Systematic Control and Scheduling Framework

through multistage stochastic programming method, and to train an optimal strategy as a decision guide for the scheduler (which is regarded as the central decision-making body of the scheduling system) based on these exogenous uncertain historical data. By making full use of the idea of cybernetics (Maggio et al., 2014), the entire scheduling system forms a closed loop with an online update mechanism as shown in Figure 2. Under this framework, higher-level managers such as those responsible for enterprise resource planning (ERP), issue fundamental scheduling policy (principle) to the scheduling system, such as their value orientation and risk preference. The scheduler makes scheduling decision based on the supervisor's observation, that is, the production state observed from the production process (such as the operation mode of the processing units, the inventory level of the tanks, etc.) and the specific realization of exogenous uncertainty (such as the current demand orders received, etc.). The operator then comprehends the decision information from the scheduler and converts it into corresponding actual operations, which are applied to the production process units equipped with

advanced control systems. The production process thus enters a new production state and generates corresponding economic benefits that meet the scheduling principle (cost saving or profit expansion). The latest state of the production process and the latest exogenous uncertainty effect are immediately fed back to the scheduler by the supervisor again, thereby generating new targeted scheduling decision, thus forming a closed loop of the scheduling system.

**Remark 1.** In the framework shown in Figure 2, incomplete or inaccurate observation and operation of the supervisor and the operator are also possible (Dowson et al., 2020), although in this article we only restrict our attention to the case of complete and accurate observation and operation.

It is worth mentioning that according to the framework presented by Figure 2, the scheduling decision made by the scheduler in the actual scheduling system will be simultaneously passed to the integrated model for simulation and deduction. The output of the model is compared and evaluated with the corresponding observations in the actual production process to form innovation and pass it to the modeler, which will timely refine and update the integrated model based on the latest production operation data to avoid model mismatch. At the same time, policy training is also timely re-performed to update the policy for the scheduler to better adapt to the real scheduling environment. In a sense, this mechanism is essentially a kind of reinforcement learning from human feedback (RLHF) (Kaufmann et al., 2024; Lambert, 2025).

In summary, the unit-level advanced control provides a limited number of achievable operation modes and a large amount of operation data for plant-level scheduling optimization, which is conducive to making the scheduling model more accurate, reliable and realistic. At the same time, the decision-making framework based on multistage stochastic dynamic scheduling provides a more flexible and reasonable operation guidance for the processing units, which is more conducive to the implementation of advanced control at the unit level, so that the effect of control optimization becomes more prominent. With the engergization of online timely model and policy update mechanism, control and scheduling work together to deal with endogenous and exogenous uncertainties in the refining production process, which will form a benign interactive situation that effectively balances exploration and expoitation, and promotes the flat development of enterprise architecture. It

is obvious but worth mentioning that the framework given by Figure 2 can be used not only to organize traditional manned factories, but also to develop intelligent unmanned factories, where the main modules in the schema are composed of computer software.

As the basis of the integrated model, the above-mentioned systematic control and scheduling framework includes two core contents, namely unit-level advanced control integrated process model representation and plant-level multistage stochastic dynamic scheduling schema. The following two sections will introduce them respectively.

*3.2.1. Unit-Level Advanced Control Integrated Process Model Representation*

Under the aforementioned framework, we structure the advanced control integrated process model of each processing unit into a square table or matrix, where the entries in different positions represent the models of the unit under different operation states, thus forming the model library. For example, the yield models of the output material $o$ of the processing unit $u$ are organized as the following $\mathcal{M}^u \times \mathcal{M}^u$ square matrix:

$$\mathbf{\Psi}_o^u = \left[\psi_{o,m,m'}^u\right]_{\mathcal{M}^u \times \mathcal{M}^u}, \quad \forall o \in \mathcal{O}^u, u \in \mathcal{PRU}, \tag{1}$$

where $\mathcal{M}^u$ denotes the index set of basic steady operation modes of the unit $u$ and $m, m' \in \mathcal{M}^u$, and the entry $\psi_{o,m,m'}^u$ of $\mathbf{\Psi}_o^u$ denotes the yield of the unit when it operates at the operation mode $m$-$m'$. Specifically, the main diagonal entry $\psi_{o,m,m}^u$ of $\mathbf{\Psi}_o^u$ represents the yield of the unit when it operates at the stable operation mode $m$, while the off-diagonal entry $\psi_{o,m,m'}^u, m \neq m'$ represents the yield at the transitional operating state when switching from the stable mode $m$ to another stable mode $m'$. The yield $\psi_{o,m,m'}^u$ can be modeled as a nominal constant scalar, or it is also allowed to be modeled as an uncertain parameter or a function on some revealed fragments of the historical states and the underlying independent uncertainties and even the current decision that is to be made, where of course the corresponding computational tractability also needs to be considered additionally. Usually in practice, especially when the unit-level advanced control system is considered to be operating relatively ideally, we can succinctly construct a single-value model for the yield $\psi_{o,m,m'}^u$ based on statistical analysis of the operation data of the unit under different operation modes, and resort to the aforementioned systematic control and scheduling framework with a timely model update mechanism to ensure the implementability of the scheduling solution (Shi et al., 2014, 2015; Zhang et al., 2015; Gao, 2014; Han, 2022).

Similar to the above yield model matrix, the models of the prodution processing unit also include matrices on the property indices of the output materials

$$\mathbf{P}^u_{r,o} = \left[\rho^u_{r,o,m,m'}\right]_{\mathcal{M}^u \times \mathcal{M}^u}, \quad \forall r \in \mathcal{R}^o, o \in \mathcal{O}^u, u \in \mathcal{PRU} \tag{2}$$

and an operation cost matrix

$$\mathbf{\Phi}^u = \left[\phi^u_{m,m'}\right]_{\mathcal{M}^u \times \mathcal{M}^u}, \quad \forall u \in \mathcal{PRU}, \tag{3}$$

where $\rho^u_{r,o,m,m'}$ represents the model of the property index $r$ of the output material $o$ when the unit $u$ operates at the operation mode $m$-$m'$, and $\phi^u_{m,m'}$ represents the unit time operation cost model when the $u$ operates at the operation mode $m$-$m'$.

In addition, very similar but slightly different, the processing unit also has the following time constant model matrix:

$$\mathbf{T}^u = \left[\tau^u_{m,m'}\right]_{\mathcal{M}^u \times \mathcal{M}^u}, \quad \forall u \in \mathcal{PRU}, \tag{4}$$

where $\tau^u_{m,m'} \in \mathbb{Z}_+$. For $\mathbf{T}^u$, its main diagonal entry $\tau^u_{m,m}$ represents the minimum duration of the stable operation mode $m$ of the unit $u$ stipulated by the refinery for production safety and stability considerations, while the off-diagonal entry $\tau^u_{m,m}, m \neq m'$ represents the duration of the transition process of the unit $u$ from the stable mode $m$ to another stable mode $m'$. They are all equivalently converted into discrete-time representations (i.e., the numbers of "stages" in the following text). Generally, in actual production, the duration of the transition process caused by operation switching of an upstream unit is longer than that of a downstream unit, and usually the stable operation mode of the former is also required to be maintained for a longer time than the latter.

*3.2.2. Plant-Level Multistage Stochastic Dynamic Scheduling Schema*

Based on the unit-level advanced control process model representation, we propose to model the plant-level refining production scheduling problem into multistage stochastic programming (MSP). Since MSP shares a lot of similarities with fields like Markov decision processes (Puterman, 1994), approximate dynamic programming (Powell, 2011), stochastic optimal control (Bertsekas, 2012a,b) and reinforcement learning (Bertsekas, 2019; Sutton and Barto, 2018), to avoid heterogeneous languages, here we mainly follow the terminology and notion clearly defined by Dowson (2020).

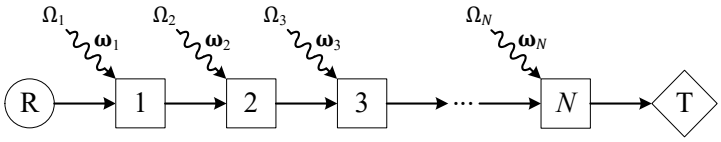


Figure 3: Linear Policy Graph

The discrete time representation commonly used in refinery production scheduling modeling provides a natural convenience for the implementation of the multistage decision-making scheme, so that we can naturally use the linear policy graph shown in Figure 3 to characterize the entire scheduling time horizon. In the linear policy graph, the scheduling horizon of the refinery production is evenly divided into $N$ scheduling decision stages that are depicted in the figure as square Hazard-Decision nodes marked with ordinal numbers, which means the scheduling decision is made after the stagewise-independent exogenous uncertainty $\boldsymbol{\omega}_n \in \Omega_n$ is revealed in each stage. This type of decision is also known as a Wait-and-See decision. The nodes in Figure 3 are connected by directed arcs, representing the transition direction of the refinery production state (such as the operation modes of the units, the storage levels of the tanks, etc.) and the order of scheduling decisions. Every arc has an origin node and a destination node that are parent and child to each other, and the outgoing state variable $\mathbf{s}_n^{\text{out}}$ of the parent node is set to be the incoming state variable $\mathbf{s}_{n+1}^{\text{in}}$ of its child node. The parent node of node 1 is the root node, represented by a circle marked with "R" in the figure, and the child node of node $N$ is the terminal node, represented by a diamond shape marked with "T" in the figure. They stand for the start and end of the scheduling horizon respectively and store the production states of the refinery at the initial and final moments.

In Figure 3, inside each Hazard-Decision stage node, there is a dynamic model of the refinery-wide production scheduling with unit-level advanced control integrated, which is generally consist of a stage decision rule, a stage transition function and a stage objective function. The stage decision rule $\boldsymbol{\pi}_n(\mathbf{s}_n^{\text{in}}, \boldsymbol{\omega}_n)$ is a funtion that maps the observation of the incoming production state $\mathbf{s}_n^{\text{in}}$ and the exogenous uncertainty $\boldsymbol{\omega}_n$ to a scheduling decision $\mathbf{u}_n$ within some admissible set $U_n(\mathbf{s}_n^{\text{in}}, \boldsymbol{\omega}_n)$. The stage transistion funtion $T_n(\mathbf{s}_n^{\text{in}}, \mathbf{u}_n, \boldsymbol{\omega}_n)$ maps the incoming production state, the scheduling decision, and the exogenous uncertainty to the outgoing production state $\mathbf{s}_n^{\text{out}}$. The stage objective funtion $C_n(\mathbf{s}_n^{\text{in}}, \mathbf{u}_n, \boldsymbol{\omega}_n)$ is a benefit evaluation of the schedul-

ing decision and the production dynamics within that stage, for example, it can be the stagewise production cost or sales profit of the refinery.

**Remark 2.** Here we only discuss the Hazard-Decision or Wait-and-See type of decision because a Decision-Hazard or Here-and-Now node is possible to be transformed into a deterministic node followed by a Hazard-Decision node with an expanded state-space (Dowson, 2020; Street et al., 2020).

According to Bellman's principle of optimality (Bellman, 1954), our overall goal of plant-level multistage stochastic dynamic scheduling is to train an optimal policy $\boldsymbol{\pi}$ (which is a set of all stage decision rules, i.e., $\{\boldsymbol{\pi}_n\}_{n=1}^N$) based on the historical data collection of the stagewise-independent exogenous uncertainty sample spaces $\{\Omega_n\}_{n=1}^N$ over the scheduling horizon, so as to minimize or maximize some risk measure $\mathbb{F}[\cdot]$ of the so-called cost-to-go or value-to-go function $V_n^{\boldsymbol{\pi}}(\mathbf{s}_n^{\text{in}}, \boldsymbol{\omega}_n)$ from the root node (which is also the input principle in the proposed systematic framework, see Figure 2), i.e., (We only take the minimizing sense as an example in the following formulations.)

$$\min_{\boldsymbol{\pi}} \quad \mathop{\mathbb{F}_{\text{R}}}_{\boldsymbol{\omega}_1 \in \Omega_1} \left[ V_1^{\boldsymbol{\pi}} \left( \mathbf{s}_{\text{R}}, \boldsymbol{\omega}_1 \right) \right], \tag{5}$$

where

$$\begin{cases} V_n^{\boldsymbol{\pi}}(\mathbf{s}_n^{\text{in}}, \boldsymbol{\omega}_n) = C_n(\mathbf{s}_n^{\text{in}}, \mathbf{u}_n, \boldsymbol{\omega}_n) + \mathop{\mathbb{F}_n}_{\boldsymbol{\omega}_{n+1} \in \Omega_{n+1}} \left[ V_{n+1}^{\boldsymbol{\pi}}(\mathbf{s}_n^{\text{out}}, \ \boldsymbol{\omega}_{n+1}) \right], \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall n \in [N-1], \\ V_N^{\boldsymbol{\pi}}(\mathbf{s}_N^{\text{in}}, \boldsymbol{\omega}_N) = C_N(\mathbf{s}_N^{\text{in}}, \mathbf{u}_N, \boldsymbol{\omega}_N) \end{cases} \tag{6}$$

$$\tag{7}$$

and $\mathbf{u}_n = \boldsymbol{\pi}_n(\mathbf{s}_n^{\text{in}}, \boldsymbol{\omega}_n) \in U_n(\mathbf{s}_n^{\text{in}}, \boldsymbol{\omega}_n)$, $\mathbf{s}_n^{\text{out}} = T_n(\mathbf{s}_n^{\text{in}}, \mathbf{u}_n, \boldsymbol{\omega}_n) \triangleq \mathbf{s}_{n+1}^{\text{in}}$ with the initial state $\mathbf{s}_{\text{R}} = \mathbf{s}_1^{\text{in}}$. We point out that, the operator $\mathbb{F}[\cdot]$ in (5) and (6) can be taken as any coherent risk measure (Artzner et al., 1999), or more generally, any convex risk measure (Dowson et al., 2022; Föllmer and Schied, 2002; Rockafellar, 2007; Liu, 2010), dependinng on the input scheduling principle from the higher level of the refinery (see Figure 2).

With the above schema in mind, to build the multistage stochastic dynamic scheduling model for the refinery, we only need to clearly define the following scheduling subproblem from the perspective of each stage node $n \in [N]$:

$$\text{SP}_n(\mathbf{s}_n^{\text{in}}, \boldsymbol{\omega}_n): \begin{cases} \min\limits_{\mathbf{u}_n, \mathbf{s}_n^{\text{out}}} & C_n(\mathbf{s}_n^{\text{in}}, \mathbf{u}_n, \boldsymbol{\omega}_n) \\ \text{s.t.} & \mathbf{s}_n^{\text{out}} = T_n(\mathbf{s}_n^{\text{in}}, \mathbf{u}_n, \boldsymbol{\omega}_n) \\ & \mathbf{u}_n \in U_n(\mathbf{s}_n^{\text{in}}, \boldsymbol{\omega}_n). \end{cases} \tag{8}$$

All of the components in this subproblem will be concretized with the help of the modeling methodology discussed in Section 4.

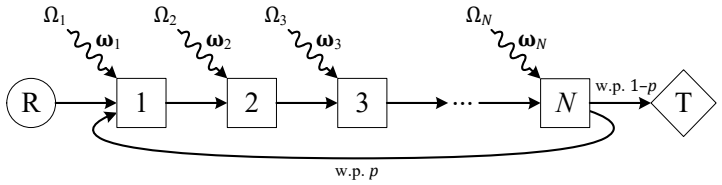


Figure 4: Cyclic Policy Graph

**Remark 3.** In actual refining production, we usually hope that the policy obtained after one training can serve for as long as possible, or even be effective in an infinite scheduling horizon, unless the policy really needs to be retrained based on the latest refined integrated model to avoid model mismatch, or encounter regular or irregular routine shutdown and maintenance of the refinery. We can introduce the cyclic policy graph shown in Figure 4 to model this application scenario (Dowson, 2020). It means that when a periodic scheduling horizon is completed, the refinery will continue the production operation of the next cycle with probability $p$, or perform model refining or shutdown and maintenance with probability $1-p$ (Bertsekas, 2012a; Puterman, 1994). In this case, we only need to replace the formula (7) in the aforementioned plant-level multistage stochastic dynamic scheduling schema with the following formula:

$$V_N^{\boldsymbol{\pi}}(\mathbf{s}_N^{\text{in}}, \boldsymbol{\omega}_N) = C_N(\mathbf{s}_N^{\text{in}}, \mathbf{u}_N, \boldsymbol{\omega}_N) + p \cdot \underset{\boldsymbol{\omega}_1 \in \Omega_1}{\mathbb{F}_1} \left[ V_1^{\boldsymbol{\pi}}(\mathbf{s}_N^{\text{out}}, \boldsymbol{\omega}_1) \right]. \tag{9}$$

## 4. Multistage Stochastic Dynamic Scheduling Modeling Methodology[2]

### *4.1. Mathematical Model*

In Section 2, the basic elements of the oil refining production process were abstracted into 8 model categories. In this section we will give their modeling methods respectively. According to the framework proposed in Section 3.2,

[2]Since this section involves many model symbols, readers can query Appendix A to learn their specific meanings.

we need to establish the refinery production process scheduling model from the perspective of the subproblem of multistage stochastic programming, that is, to characterize the model with respect to each node or stage. When modeling, we should take into account constraints including the operation mode switching rules of the processing units, the balance between material input and output, product quality requirements, and the evaluation of costs and benefits. Since our subproblem model is applicable to all stage nodes, to avoid messy notation, we omit the subscript $n$ representing node number in the model.

We first discuss the modeling of PRUs. Since they involve the switching of operation modes, we first introduce the following operation mode indicative matrix with binary elements (Han et al., 2022):

$$\mathbf{X}^u = \left[x^u_{m,m'}\right]_{\mathcal{M}^u \times \mathcal{M}^u} \in \{0,1\}^{|\mathcal{M}^u| \times |\mathcal{M}^u|}, \quad \forall u \in \mathcal{PRU}.$$

If $x^u_{m,m'} = 1$, it means that the processing unit $u$ is working at the operation mode $m$-$m'$ in the current stage (if $m = m'$, it refers to the steady state $m$, otherwise it refers to the corresponding transition state, which corresponds to the previous formula (1)$\sim$(4)). Based on our settings, the unit can only working at a single steady or transitional operation mode in any stage, which means

$$\sum_{m,m' \in \mathcal{M}^u} x^u_{m,m'} = 1, \quad \forall u \in \mathcal{PRU}. \tag{10}$$

Furthermore, in order to characterize the operation mode switching rules of the processing unit during the entire scheduling horizon, we need to access the unit's operation modes in the previous stages in the subproblem of the current stage, so we additionally introduce an "operation mode shift register" $\mathbf{XR}^u = \{\mathbf{XR}^u_t\}_{t=1}^{\overline{\tau}^u}$, where

$$\mathbf{XR}^u_t = \left[xr^u_{t,m,m'}\right]_{\mathcal{M}^u \times \mathcal{M}^u} \in \{0,1\}^{|\mathcal{M}^u| \times |\mathcal{M}^u|}, \quad \forall t \in [\overline{\tau}^u], u \in \mathcal{PRU}.$$

It is used to dynamically store the unit's operation mode states in the previous $t = 1, \cdots, \overline{\tau}^u$ stages. It is essentially a set of auxiliary state variables, which, like other state variables, has two counterparts $\mathbf{XR}^{\text{in},u}$ and $\mathbf{XR}^{\text{out},u}$, representing the states of the register when coming in and going out of the current node of stage. In another word, mathematically, we let its transition function follow the following rules, for all $u \in \mathcal{PRU}$:

$$\begin{cases} \mathbf{XR}^{\text{out},u}_1 = \mathbf{X}^u, \\ \mathbf{XR}^{\text{out},u}_t = \mathbf{XR}^{\text{in},u}_{t-1}, \quad t = 2, \cdots, \overline{\tau}^u. \end{cases} \tag{11}$$

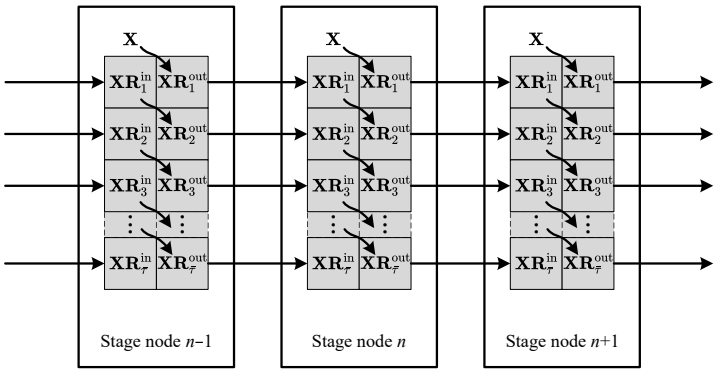


Figure 5: Transfer Flow of Operation Mode Information

Note that, the above rule, together with the implicit mechanism of any state variable passing between stage nodes (i.e., the outgoing state of the previous node is equal to the incoming state of the current node) will enable the register to store the unit operation states of the previous nodes in a "first in, first out" rolling manner with memory $\overline{\tau}^u$. Thus, we can easily access the unit's operation mode indicative matrix of the previous $t$th stage from the current stage node, that is, $\mathbf{XR}_t^{\text{in},u}$. Figure 5 shows the flow of the operation mode information of each processing unit $u$ between different stage nodes. We point out that it will be enough for our modeling if we take the memory

$$\overline{\tau}^u = 1 + \max_{m,m' \in \mathcal{M}^u} \tau_{m,m'}^u, \quad \forall u \in \mathcal{PRU}.$$

Based on this, we are able to characterize the operation mode switching rules for the processing unit. If the unit switches from a steady state to another steady state, it must go through the corresponding transition state in the middle. That is, if the previous mode is a steady state, the operation mode of the current stage can still be the steady state or a transition state:

$$\sum_{m' \in \mathcal{M}^u} x_{m,m'}^u \geq xr_{1,m,m}^{\text{in},u}, \quad \forall m \in \mathcal{M}^u, u \in \mathcal{PRU}; \tag{12}$$

if the previous mode is a transition state, the next mode can still be this transition state or the corresponding steady state to which it is switching:

$$x_{m,m'}^u + x_{m',m'}^u \geq xr_{1,m,m'}^{\text{in},u}, \quad \forall m, m' \in \mathcal{M}^u, m \neq m', u \in \mathcal{PRU}. \tag{13}$$

The duration of any transition state must be equal to the duration of the transition process $\tau^u_{m,m'}$, which is guaranteed by the following three inequalities:

$$\begin{cases} \sum_{t=1}^{\tau^u_{m,m'}} xr^{\text{in},u}_{t,m,m'} \geq \tau^u_{m,m'} \cdot \left( xr^{\text{in},u}_{1,m,m'} + x^u_{m',m'} - 1 \right), & (14) \\ xr^{\text{in},u}_{\tau^u_{m,m'}+1,m,m} \geq xr^{\text{in},u}_{1,m,m'} + x^u_{m',m'} - 1, & (15) \\ \sum_{t=1}^{\tau^u_{m,m'}} xr^{\text{in},u}_{t,m,m'} + x^u_{m,m'} \leq \tau^u_{m,m'}, & (16) \end{cases}$$

$$\forall m, m' \in \mathcal{M}^u, m \neq m', u \in \mathcal{PRU}.$$

The duration of any steady state should be not less than the duration $\tau^u_{m,m}$ required by production safety:

$$\sum_{t=1}^{\tau^u_{m,m}} xr^{\text{in},u}_{t,m,m} \geq \tau^u_{m,m} \cdot \left( xr^{\text{in},u}_{1,m,m} + \sum_{m' \neq m} x^u_{m,m'} - 1 \right), \quad \forall m \in \mathcal{M}^u, u \in \mathcal{PRU}. \tag{17}$$

In addition to the above rules, sometimes according to production practice, the operation mode of some downstream units needs to be consistent with their upstream units, so in this case a small amount of additional state correlation constraints also need to be added to characterize this dependency (Shi et al., 2014; Zhang et al., 2015).

The PRUs must satisfy the material balance constraints, that is, the flow rate $g^u_o$ of a certain output material of a unit should be equal to the load $f^u$ of its feed material multiplied by the corresponding yield:

$$g^u_o = \left( \sum_{m,m' \in \mathcal{M}^u} x^u_{m,m'} \cdot \psi^u_{o,m,m'} \right) \cdot f^u, \quad \forall o \in \mathcal{O}^u, u \in \mathcal{PRU}. \tag{18}$$

For the sake of simplicity, for units with only a single input/output material, we omit the subscripts $i/o$ representing the input/output material index in the input flow $f^u_i$/output flow $g^u_o$, when it does not cause confusion, such as the $f^u$ in the above formula.

For those processing units directly connected to the blenders, it is necessary to consider the key property values of the output materials:

$$q^u_{r,o} = \sum_{m,m' \in \mathcal{M}^u} x^u_{m,m'} \cdot \rho^u_{r,o,m,m'}, \quad \forall r \in \mathcal{R}^o, o \in \mathcal{O}^u, u \in \mathcal{PRU}. \tag{19}$$

In each stage, the processing unit will incur an operation cost proportional to its processing load and processing time:

$$z^u_{\text{opr}} = \left( \sum_{m,m' \in \mathcal{M}^u} x^u_{m,m'} \cdot \phi^u_{m,m'} \right) \cdot f^u \cdot \Delta t, \quad \forall u \in \mathcal{PRU}. \tag{20}$$

For BLDs, the sum of the flow rates of each component oil used to blend various product oils is equal to the feed flow rate of this kind of component oil:

$$\sum_{o \in \mathcal{O}^u} w^u_{i,o} = f^u_i, \quad \forall i \in \mathcal{I}^u, u \in \mathcal{BLD}; \tag{21}$$

and the output material flow rate of each product oil is equal to the sum of the flow rates of various component oils used to blend this product oil:

$$g^u_o = \sum_{i \in \mathcal{I}^u} w^u_{i,o}, \quad \forall o \in \mathcal{O}^u, u \in \mathcal{BLD}. \tag{22}$$

The $w^u_{i,o}$ in the formula represents the flow rate of the feed component oil $i$ used by the BLD in the current stage to blend the product oil $o$, which must also meet the natural upper and lower bound constraints

$$0 \leq w^u_{i,o} \leq f^u_i, \quad \forall i \in \mathcal{I}^u, o \in \mathcal{O}^u, u \in \mathcal{BLD} \tag{23}$$

and the product component content ratio restrictions required by production

$$\underline{\kappa}^u_{i,o} \cdot g^u_o \leq w^u_{i,o} \leq \overline{\kappa}^u_{i,o} \cdot g^u_o, \quad \forall i \in \mathcal{I}^u, o \in \mathcal{O}^u, u \in \mathcal{BLD}. \tag{24}$$

Some key properties of the blended product oils, such as the condensation point factor (CPF) and cetane number index (CNI) of diesels, the research octane number (RON) of gasolines, and the sulfur content (SC) of the both, need to meet the upper and lower limits specified by national or local quality standards. The values of these property indices can be approximated by linear calculation formulas, so the product quality constraints can be expressed as (Shi et al., 2015, 2014; Hou, 1991):

$$\underline{\rho}^u_{r,o} \cdot g^u_o \leq \sum_{i \in \mathcal{I}^u} p^u_{r,i} \cdot w^u_{i,o} \leq \overline{\rho}^u_{r,o} \cdot g^u_o, \quad \forall r \in \mathcal{R}^o, o \in \mathcal{O}^u, u \in \mathcal{BLD}. \tag{25}$$

Please note that the property $p^u_{r,i}$ of the input materials of a blender $u$ is consistent with the property $q^{u^\star}_{r,i}$ (which is determined by (19)) of the corresponding output materials of its direct upstream processing unit that we

denote by $u^\star$. The operation costs of the blending units are negligible in the model.

For storage tanks, RTKs, CTKs and PTKs are treated differently. The RTK is at the upstream of oil refining production and it is the source of raw material. We assume that the supply of raw material is always sufficient throughout the entire scheduling time horizon, and only consider the cost of the raw material consumed by the processing units that are transported outward from the RTK, that is,

$$z^u_{\text{mat}} = \alpha^u \cdot g^u \cdot \Delta t, \quad \forall u \in \mathcal{RTK}. \tag{26}$$

For CTKs and PTKs, we need to consider their tank inventory balances

$$v^{\text{out},u} = v^{\text{in},u} + (f^u - g^u) \cdot \Delta t, \quad \forall u \in \mathcal{CTK} \tag{27}$$

and

$$v^{\text{out},u} = v^{\text{in},u} + f^u \cdot \Delta t - s^u, \quad \forall u \in \mathcal{PTK}, \tag{28}$$

and storage costs

$$z^u_{\text{sto}} = \beta^u \cdot \left(v^{\text{in},u} + v^{\text{out},u}\right) \cdot \Delta t/2, \quad \forall u \in \mathcal{CTK} \cup \mathcal{PTK}. \tag{29}$$

The $v^u$ in formulas (27) and (28) is a continuous-valued state variable, and its two counterparts $v^{\text{in},u}$ and $v^{\text{out},u}$ represent the tank inventory levels at the start and end of the current stage, respectively. The $s^u$ in formula (28) refers to the amount of the corresponding oil product actually delivered or sold to the demander from PTK $u$, which should not exceed the demand $\delta^u$ for this kind of oil product:

$$0 \le s^u \le \delta^u, \quad \forall u \in \mathcal{PTK}. \tag{30}$$

In the scheduling optimization model of this article, we allow that the demand for product oils may not always be met, but the refinery should bear the penalty caused by shortage:

$$z^u_{\text{pen}} = \gamma^u \cdot (\delta^u - s^u), \quad \forall u \in \mathcal{PTK}. \tag{31}$$

This enables the manager or decision-maker to characterize the tolerance for the occurrence of related product shortages in practice by adjusting the shortage penalty coefficient $\gamma^u$ in the model. In addition, sometimes we also consider the income from oil sales:

$$z^u_{\text{inc}} = \mu^u \cdot s^u, \quad \forall u \in \mathcal{PTK}. \tag{32}$$

For the sake of production capacity and production safety, the feed load $f_i^u$ of the processing units and the storage level $v^u$ of the storage tanks discussed above must also be subject to some upper and lower limit constraints:

$$\underline{f}_i^u \leq f_i^u \leq \overline{f}_i^u, \quad \forall i \in \mathcal{I}^u, u \in \mathcal{PRU} \cup \mathcal{BLD} \cup \mathcal{CTK} \cup \mathcal{PTK}, \tag{33}$$

$$\underline{v}^u \leq v^{\text{out},u} \leq \overline{v}^u, \quad \forall u \in \mathcal{CTK} \cup \mathcal{PTK}. \tag{34}$$

The output material flow of a dummy unit COM is equal to the sum of its feed flows:

$$g^u = \sum_{i \in \mathcal{I}^u} f_i^u, \quad \forall u \in \mathcal{COM}. \tag{35}$$

Correspondingly, SPLs follow a similar material balance relationship, that is,

$$\sum_{o \in \mathcal{O}^u} g_o^u = f^u, \quad \forall u \in \mathcal{SPL}. \tag{36}$$

At the same time, the properties of the materials flowing through the SPLs do not change:

$$q_{r,o}^u = p_r^u, \quad \forall r \in \mathcal{R}^o, o \in \mathcal{O}^u, u \in \mathcal{SPL}. \tag{37}$$

Finally, the input and output materials of all the units discussed above are connected through pipelines $l_{o,i}$ according to the topology of the refinery, where $o$ represents the output material index of the upstream unit, i.e., the input end of the pipeline, and $i$ represents the input material index of the corresponding downstream unit, i.e., the output end of the pipeline. The pipeline does not change the flow rate and properties of the material, so there is

$$f_i^u = g_o^{u^-}, \quad \forall i \in \mathcal{I}^u, o \in \mathcal{O}^{u^-}, l_{o,i} \in \mathcal{PPL}, u \in \mathcal{U} \setminus \mathcal{RTK} \tag{38}$$

and

$$p_{r,i}^u = q_{r,o}^{u^-}, \quad \forall\, r \in \mathcal{R}^i, i \in \mathcal{I}^u, o \in \mathcal{O}^{u^-}, l_{o,i} \in \mathcal{PPL}, u \in \mathcal{U} \setminus \mathcal{RTK}, \tag{39}$$

where the $u^-$ in the formulas represents the upstream unit that is directly connected to $u$ by a pipeline.

In any stage, the overall goal of our scheduling optimization is to minimize the overall production cost generated in that stage

$$z_{\text{cost}} = \sum_{u \in \mathcal{RTK}} z_{\text{mat}}^u + \sum_{u \in \mathcal{PRU}} z_{\text{opr}}^u + \sum_{u \in \mathcal{CTK} \cup \mathcal{PTK}} z_{\text{sto}}^u + \sum_{u \in \mathcal{PTK}} z_{\text{pen}}^u \tag{40}$$

or maximize the overall profit

$$z_{\text{profit}} = \sum_{u \in \mathcal{PTK}} z^u_{\text{inc}} - z_{\text{cost}}. \tag{41}$$

In summary, from the perspective of each stage node $n \in [N]$, the subproblem (8) is concretized as: (We only take minimizing the cost as an example.)

$$\text{SP}: \begin{cases} \min & z_{\text{cost}} \\ \text{s.t.} & (10) \sim (39). \end{cases} \tag{42}$$

*4.2. Model Analysis*

Table 1: Model Statistics

| Category | #Variables | #Constraints |
|---|---|---|
| PRU | $\sum_{u \in \mathcal{PRU}} \left( \lvert\mathcal{M}^u\rvert^2 (2\overline{\tau}^u + 1) + \lvert\mathcal{O}^u\rvert + 2 + \sum_{o \in \mathcal{O}^u} \lvert\mathcal{R}^o\rvert \right)$ | $\sum_{u \in \mathcal{PRU}} \left( \lvert\mathcal{M}^u\rvert^2 (\overline{\tau}^u + 4) - 2\lvert\mathcal{M}^u\rvert + 2\lvert\mathcal{I}^u\rvert + \lvert\mathcal{O}^u\rvert + 2 + \sum_{o \in \mathcal{O}^u} \lvert\mathcal{R}^o\rvert \right)$ |
| BLD | $\sum_{u \in \mathcal{BLD}} (\lvert\mathcal{I}^u\rvert + \lvert\mathcal{O}^u\rvert + \lvert\mathcal{I}^u\rvert \lvert\mathcal{O}^u\rvert)$ | $\sum_{u \in \mathcal{BLD}} \left( 3\lvert\mathcal{I}^u\rvert + \lvert\mathcal{O}^u\rvert + \lvert\mathcal{I}^u\rvert \lvert\mathcal{O}^u\rvert + 2\sum_{o \in \mathcal{O}^u} \lvert\mathcal{R}^o\rvert \right)$ |
| RTK | $\lvert\mathcal{RTK}\rvert$ | $\lvert\mathcal{RTK}\rvert$ |
| CTK | $5\lvert\mathcal{CTK}\rvert$ | $4\lvert\mathcal{CTK}\rvert + 2\sum_{u \in \mathcal{CTK}} \lvert\mathcal{I}^u\rvert$ |
| PTK | $7\lvert\mathcal{PTK}\rvert$ | $8\lvert\mathcal{PTK}\rvert + 2\sum_{u \in \mathcal{PTK}} \lvert\mathcal{I}^u\rvert$ |
| COM | $\sum_{u \in \mathcal{COM}} (\lvert\mathcal{I}^u\rvert + 1)$ | $\lvert\mathcal{COM}\rvert$ |
| SPL | $\sum_{u \in \mathcal{SPL}} \left( \lvert\mathcal{I}^u\rvert + 1 + 2\sum_{o \in \mathcal{O}^u} \lvert\mathcal{R}^o\rvert \right)$ | $\lvert\mathcal{SPL}\rvert + \sum_{u \in \mathcal{SPL}, o \in \mathcal{O}^u} \lvert\mathcal{R}^o\rvert$ |
| PPL | 0 | $2\lvert\mathcal{PPL}\rvert$ |

Table 1 lists the numbers of scalar variables and constraints contained in each category of units in the subproblem model (42). We figure out that, among all the variables, the state variables include the operation mode registers $\mathbf{XR}^u$ of the processing units as binary variables and the tank inventory levels $v^u$ of the component oil and product oil tanks as continuous variables, which together have a total of $2\left(\sum_{u \in \mathcal{PRU}} \lvert\mathcal{M}^u\rvert^2 \overline{\tau}^u + \lvert\mathcal{CTK}\rvert + \lvert\mathcal{PTK}\rvert\right)$ scalar variables. All other variables are general control variables, which stand for actions taken (implicitly or explicitly) by the oprators of the refinery within a node of stage that modify the state variables. Please note that in practical

production, operators only need to focus on and accurately execute some of the key or independent control variables, and the other variables will be determined thereby by the objectively existing production process constraint laws. For example, for the refinery shown in Figure 1, such kind of key independent control variables include: the operation mode indicative variables $\mathbf{X}^u$ of all PRUs, the feed loads $f^u$ of ADU, HTU1 and ETHU, the flow allocation variables $w_{i,o}^u$ of the blenders, and the sales quantity $s^u$ of each type of product oil. In this model, the uncertain variables we can consider include but are not limited to crude oil price $\alpha^u$, storage tank inventory cost $\beta^u$, product oil demand $\delta^u$, product oil price $\mu^u$, etc. Their observed values can be known to the refinery at the beginning of each stage before the scheduling decision-making of that stage.

Note that the subproblem model (42) itself is a mixed integer nonlinear programming problem. The binary variables come from the operation mode indicative variables $\mathbf{X}^u$ and the operation mode registers $\mathbf{XR}^u$ of the processing units, totaling $\sum_{u\in\mathcal{PRU}} |\mathcal{M}^u|^2 (2\overline{\tau}^u + 1)$. Its nonlinearity comes from the material balance constraints (18) and operation cost constraints (20) of the processing units, as well as the product quality constraints (25) of the blenders. Specifically, take the former constraint as an example, there exist bilinear terms that include $x_{m,m'}^u \cdot f^u$ in it. It will be a better practice if we linearize the nonlinear constraints before resorting to the solver. Fortunately, the methed in the literature (Shi et al., 2014, 2015; Zhang et al., 2015) to dealing with bilinearity can also be applied to the subproblem model presented in this paper. That is, if we augment the feed load $f^u$ of the processing unit into a matrix $\left[f_{m,m'}^u\right]_{\mathcal{M}^u\times\mathcal{M}^u}$ that corresponds to each of the operation modes of the unit, the nonlinear constraint (18) can then be replaced by the following linear ones:

$$\begin{cases} g_o^u = \sum\limits_{m,m'\in\mathcal{M}^u} f_{m,m'}^u \cdot \psi_{o,m,m'}^u, \quad \forall o \in \mathcal{O}^u, u \in \mathcal{PRU}, \\ \underline{f}^u \cdot x_{m,m'}^u \le f_{m,m'}^u \le \overline{f}^u \cdot x_{m,m'}^u, \quad \forall u \in \mathcal{PRU}. \end{cases}$$

Other nonlinear constraints can be reformulated linearly in the same way and will not be repeatedly described here. In a nutshell, by appropriately adding some variables and constraints, the subproblem model (42) can be converted into a mixed integer linear programming form, which is convenient for solving.

The entire multistage refinery production scheduling model we constructed

above can be solved using the stochastic dual dynamic programming (SDDP) algorithm (Pereira and Pinto, 1991; Zou et al., 2019; Dowson, 2020). Its core idea is to convert the complicated problem into a tractable form by applying Kelley's cutting plane algorithm to approximate the nasty recursive $V_n^{\boldsymbol{\pi}}(\mathbf{s}_n^{\text{in}}, \boldsymbol{\omega}_n)$ functions in the cost/value-to-go term (6). The algorithm generates the cuts through simulations based on the uncertainty data set, which can be viewed as a RL mechanism where it tries to train an optimal action-value function. In the algorithm, it is usually necessary to introduce one extra dummy variable and constraint for each state variable in order to use the so-called *fishing dual* technique to conveniently obtain valid subgradients of $V_n^{\boldsymbol{\pi}}(\mathbf{s}_n^{\text{in}}, \boldsymbol{\omega}_n)$ (Girardeau et al., 2015; Dowson, 2020). This as a result will cause a moderate increase in the actual number of variables and constraints in the model.

## 5. Application Case Study

### *5.1. Case Description*

We apply the systematic framework proposed in this paper to a real-scale refinery production scheduling case. The production process of the refinery under study has been given as an example in Section 2 above, namely Figure 1. It contains one set of ADU, VDU, FCCU, HDSU, ETHU, MTBEU, RFU, and two sets of HTU, producing a total of 5 component oils for blending gasoline products and 3 component oils for blending diesel products. Some of the component oils are stored in separate component oil tanks once produced for blending at any time, while the rest are not equipped with intermediate storage tanks and are transported to the blending units through pipelines for direct blending. The above component oils are blended through GBLD and DBLD respectively, and finally 5 types of gasoline and 3 types of diesel are produced and stored in different product oil storage tanks. As described in Section 2, some key property indicators of these gasoline and diesel products that will be sold to the market must meet the corresponding quality standards.

In this case, the exogenous uncertain parameters we focus on are the demand for all gasoline and diesel products. The refinery divides the daily production cycle into 12 decision stages at 2-hour intervals. The demand for 8 types of refined oil products in each stage every day for the past 3 months (90 days) was recorded as a training data set. In order to better show the distribution of these uncertain demand historical data, we visualize them

according to the daily production cycle through a violin plot, as shown in Figure 6, where the demand for gasoline and diesel is displayed separately. The horizontal axis in the figure represents the different time stages of a day, and the vertical axis represents the demand level. Each violin-shaped area with different colors represents the distribution of historical demand data for a type of diesel or gasoline. The horizontal line segment in the middle of the color area represents the median demand level of the oil product at the same stage in history, the line segments at the upper and lower ends of the color area represent the extreme values of the historical demand data, and the horizontal width at different heights of the color area represents the proportion of samples at the corresponding demand level in the historical data.

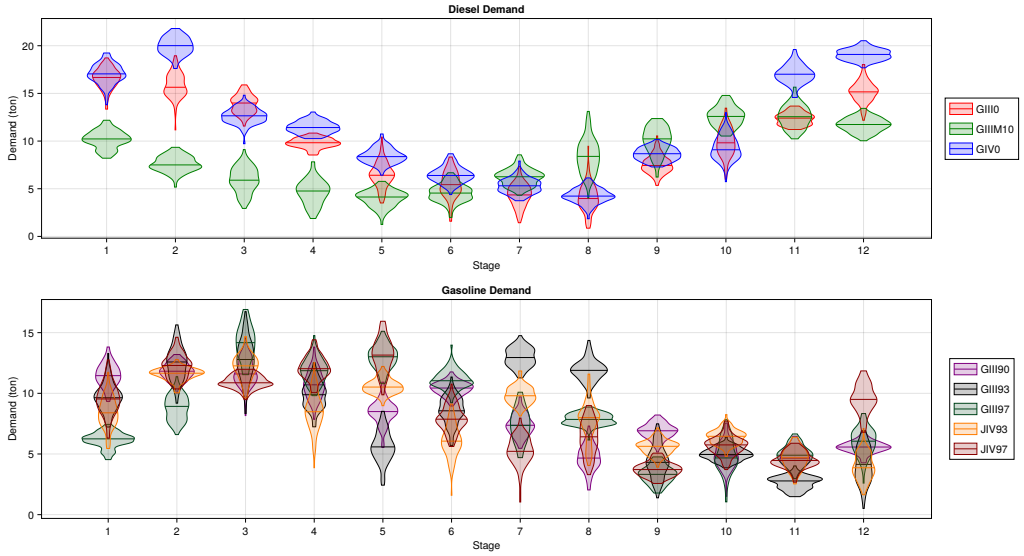


Figure 6: Distribution of Historical Demand Data over the Past 3 Months

How to formulate a risk-neutral scheduling policy to minimize the expected production cost while meeting the uncertain market demand as much as possible constitutes our refinery production scheduling optimization problem.

### *5.2. Modeling and Training*

According to the modeling method proposed above, this paper establishes the production scheduling model of the entire refinery using Julia v1.10.3

and impliments the SDDP slgorithm to solve the corresponding multistage production scheduling optimization problem with the help of the open-source package `SDDP.jl` (Dowson and Kapelevich, 2021) v1.7.0 on a computer with 1.80 GHz Intel Core i7 CPU and 16 GB of RAM. The underlying optimization solver we use is COPT (Ge et al., 2022) v7.1.1, which is invoked through its Julia interface `COPT.jl` v1.1.18.

It is worth mentioning that in order to meet the needs of long-term stable operation of the refinery production scheduling system, we introduced the cyclic policy graph when modeling to characterize the application scenario of infinite scheduling horizon. The discount factor is set to $p = 0.9$, which means that we tolerate the scheduling system to perform model refining or shutdown and maintenance with a probability of 0.1. Therefore, taking a day as a production cycle and every 2 hours as a natural decision stage, there are 12 stage nodes in our cyclic policy graph. The scheduling principle of this problem is to minimize the expectation of the cost-to-go evaluation from each stage node once the uncertain demand of the current stage has been revealed. So the general risk measure $\mathbb{F}[\cdot]$ in the formulation (5) of this problem should be replaced by the expectation operator $\mathbb{E}[\cdot]$. When training the policy, the sample average approximation method is adopted (King and Rockafellar, 1993; Shapiro, 1993; Homem-de Mello and Bayraksan, 2014; Shapiro et al., 2021). The uncertain demand for the eight types of oil products in each stage is treated as an 8-dimensional stage-independent random variable, which follows a uniform distribution on the given 3-month historical demand samples in that stage. Model statistics show that our problem contains a total of 1026 scalar variables and 1461 scalar constraints, of which the number of binary variables is 336.

Although theoretically, SDDP-type algorithms usually converge to an optimal policy almost surely in a finite number of iterations (Girardeau et al., 2015; Guigues, 2016; Philpott and Guan, 2008), this number may be very large, and such a result may not hold for the infinite-horizon setting (Dowson, 2020). As a matter of fact, in practice we only need to seek for a policy that is good enough for use. In our problem, we let the training of the policy terminate once the deterministic bound fails to improve by more than an absolute tolerance $\epsilon = 1$ (accordingly, the relative value of this tolerance with respect to the absolute value of the bound is less than $10^{-6}$) for more than $k = 10$ consective iterations, or the training time exceed a time limit of $t_l = 20$ min. The actual training result shows that the algorithm terminated because of bound stalling and the training process lasted 17.95 min.

*5.3. Simulation Analysis*

Monte Carlo simulation is an effective means to evaluate the quality of the trained scheduling policy before the scheduling system is actually put into operation. Depending on the different ways of generating uncertain demand observations, we performed both in-sample and out-of-sample Monte Carlo simulations on the scheduling policy respectively. The so-called in-sample Monte Carlo simulation means that we simulate the trained policy using the same distributions of the uncertain demand terms that were defined during the training phase. Obviously, this may not be a good approximation of of the underlying probabilistic nature of the true production and sales process, and the generalization ability of the trained policy can not be well tested by this way. Therefore, we also carried out the out-of-sample Monte Carlo simulation. Specifically, we first performed maximum likelihood estimation based on the historical demand data of the 8 types of product oils shown in Figure 6 to obtain the joint Gaussian distribution of the uncertain demand in each stage to model the demand characteristics of the market, where we allowed for the correlation between the demand for different product oils in a same stage. Subsequently, in the Monte Carlo simulation process, samples drawn from the corresponding probability distribution were used as simulated observations of uncertain demand. In both the in-sample and out-of-sample simulation schemes, we replicated the simulation for 100 times under the same initial conditions, and in each replication we let the simulation last 5 days (i.e., 60 stages) without interruption for the convenience of illustration, although the policy we trained could theoretically be applied to longer or even infinite scheduling horizons.

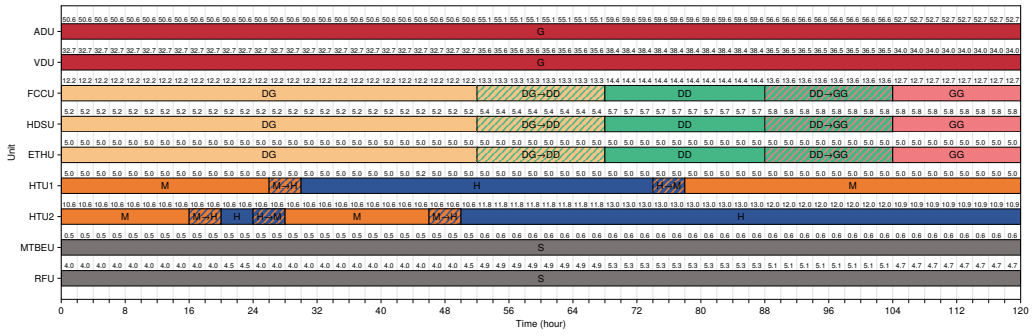


Figure 7: Operation Gantt Chart of Processing Units Obtained by Simulation

Here in this paper, we only focus on the main results of the out-of-sample

Monte Carlo simulation experiment. In these 100 replications, there are 41 different combinations of operation mode switching sequences for all processing units (dubbed operation mode trajectories for short). Even in replications with the same operation mode trajectory, the values of continuous manipulated variables such as feed load of various units are also diverse, which reflects the flexibility of our trained policy, that is, it can fully mobilize production units to respond to different market demand situations with various production strategies. Figure 7 presents the operation Gantt chart of all processing units in one of the replications corresponding to the operation mode trajectory with the highest occurrence frequency (9%). In Figure 7, the horizontal axis represents the entire simulation time horizon, with a minimum interval of 2 hours (i.e., one decision stage), and the vertical axis lists all processing units in the refinery. The single-color bars in the figure represent the stable operation modes of the units, and the two-color zebra stripe bars represent the transition processes of switching between different operation modes, which are marked with the names of the corresponding operation modes or transition processes. The feed load of each processing unit in each stage is marked above the bar, in units of ton/h. As can be seen from the figure, ADU and VDU, which are at the upstream of the entire production process, always maintain a stable operation mode to avoid large production fluctuations and increase production costs. They only control the total output by adjusting the production load to adapt to the changes of market demand. The production strategies of the downstream FCCU, HDSU, ETHU, HTU1 and HTU2 are more flexible, and their operation mode switches are more diverse. Since MTBEU and RFU only have one stable operation mode, they only make some load fine-tuning in the entire production process.

In addition, in the part of blending and delivery of product oils, the blending loads of the blenders and the delivery amounts of the products in each stage are also important manipulated variables of the refining production process. Figure 8 takes DBLD as an example to show the load variation trajectories and their distributions of the 3 diesel component oils used to blend the 3 types of diesel products in these 100 simulation replications, where each of the subplots represents the usage direction of a kind of component oil and is indicated by its title. Figure 9 shows the variation trajectories and the distributions of the delivery amounts of the 8 product oils. Each subplot in these figures displays ribbons of the $0 \sim 100$, $10 \sim 90$ and $25 \sim 75$ percentiles, and depicts the median (i.e., 50th percentile) in the middle with dark solid line. It can be seen from the figures that driven by the uncertain demand

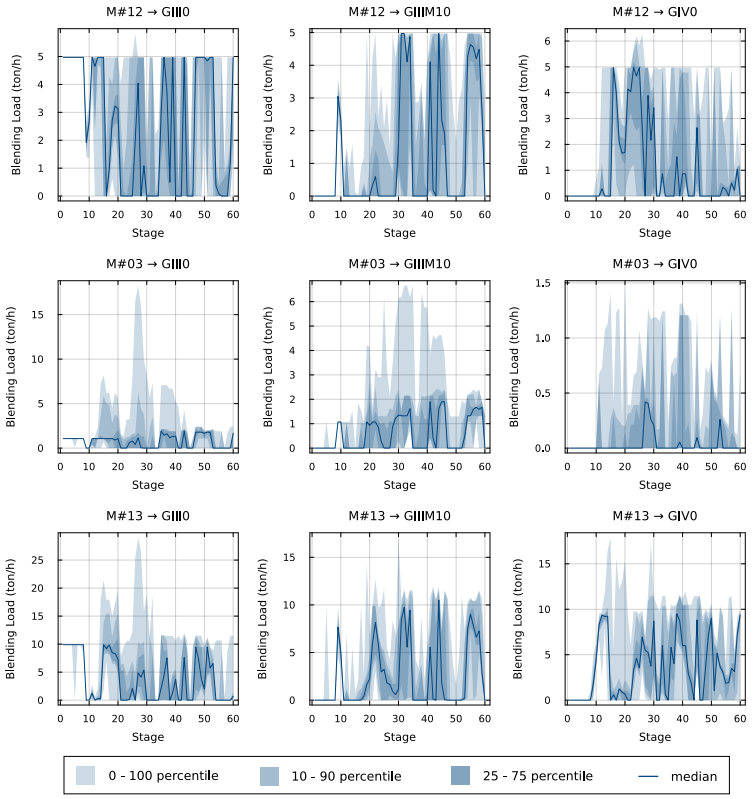


Figure 8: Blending Loads of DBLD Obtained by Monte Carlo Simulation

sequences, the blending units at the downstream of the production process respond agilely to blend various product oils that meet market demand as much as possible for delivery, so that the actual delivery amounts show periodic variation trends that are consistent with the daily cycles in demand. The storage levels of the tanks are typical state variables in the production process and an important focus of production safety. Their variations are shown in Figure 10. As a matter of fact, they are always maintained within allowable reasonable ranges in all simulation replications. To sum up, our trained scheduling policy can ensure the safe, stable, long-term and optimized operation of the refining production process.

We point out that the results of the Monte Carlo simulation can also provide inspiring hints for the development of the refinery at a higher level. For example, as can be seen from Figure 10, the storage level of the product

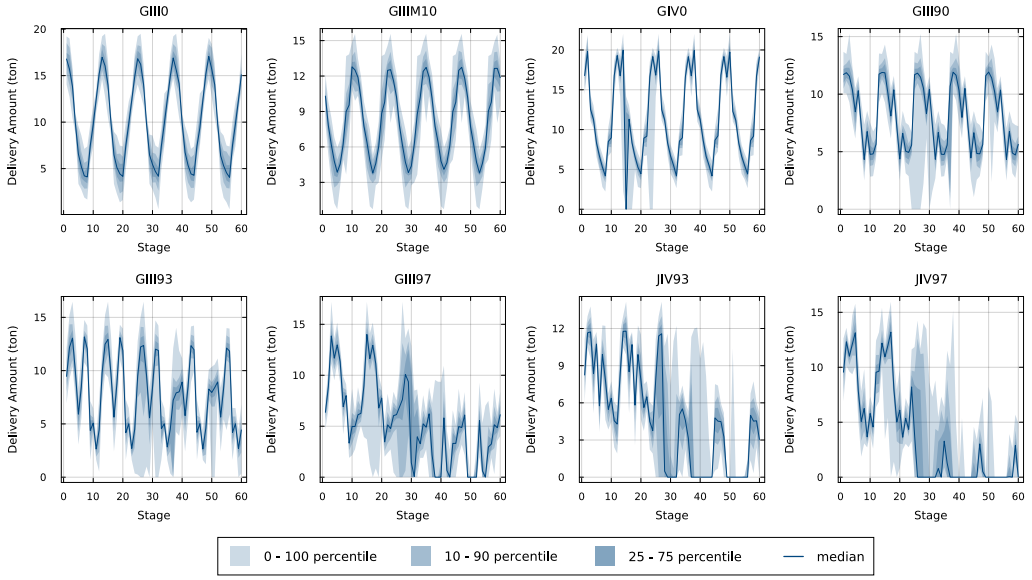


Figure 9: Delivery Amounts of 8 Product Oils Obtained by Monte Carlo Simulation

diesel GIII $0^{\#}$ storage tank may be much higher than that of other tanks, so the excess stock reflects potential production overcapacity; and multiple gasoline products may have a certain degree of production undercapacity, for they consume a large amount of inventory to meet delivery needs, and the reason may be traced back to the scarcity of C5 (M#14) and catalytic reformed gasoline (M#15) component oils. Therefore, for the managers of the refinery, these situations may imply the necessity of adjusting and upgrading the production structure of the refinery. Specifically, the future development direction may be to increase the storage capacity of diesel GIII $0^{\#}$ or open up new and larger sales channels, as well as upgrade RFU-related equipments or purchase a certain amount of gasoline component oils in a timely manner, so as to better adapt to the demand structure.

Due to space limitations, for more results of the out-of-sample simulation and the results of the in-sample simulation, please refer to Appendix B.

## 6. Concluding Remarks

This paper established an innovative data-driven multistage scheduling structured framework for long-term full-process refinery production under

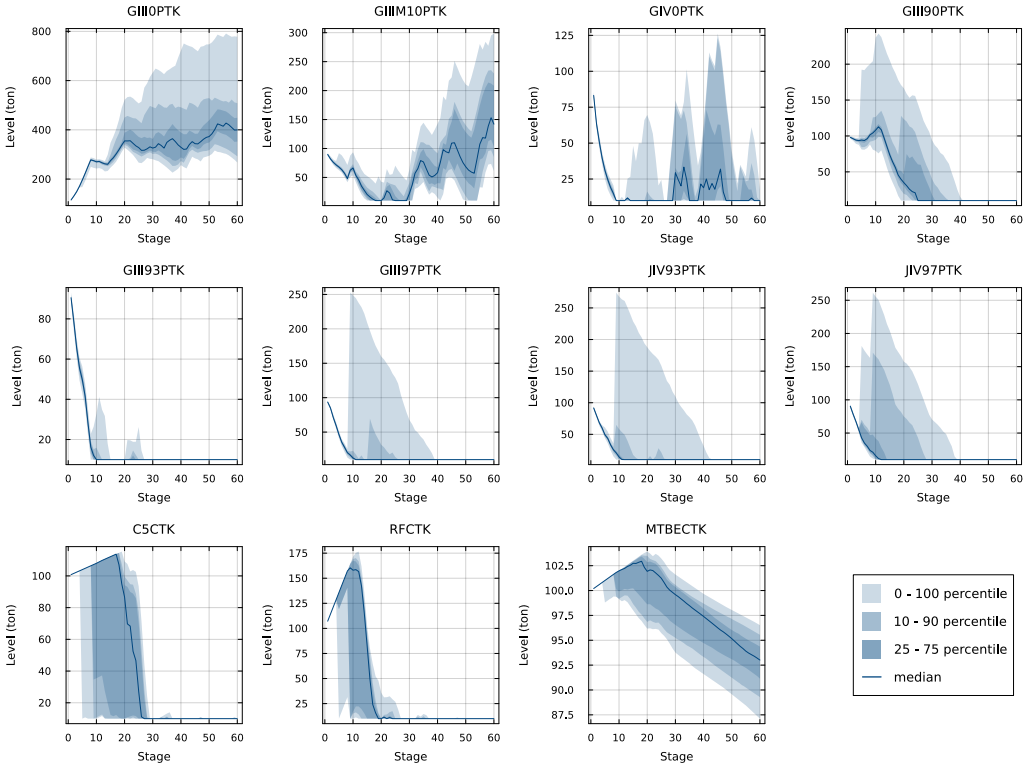


Figure 10: Storage Levels of All Tanks Obtained by Monte Carlo Simulation

uncertainty. This framework organically integrates unit-level advanced control and plant-level intelligent scheduling to systematically deal with endogenous and exogenous uncertainties in the refining production process. Under this framework, this paper further presented a detailed multistage stochastic dynamic scheduling modeling methodology based on the policy graph representation, in which the transition processes induced by operational switching of processing units were incorporated. The tractability of the model can be achieved by resorting to the existing SDDP algorithm. Following this paradigm, the scheduling model is driven by the continuously generated data from production reality, and the risk measure is allowed to be adjusted according to the decision-maker's risk preference, so that the implementability of the scheduling results is effectively guaranteed. The dynamic multistage sequential decision-making mode enhances the interactivity and interpretability of refinery production scheduling and achieves a balance between

exploitation and exploration. The integration of RL and RLHF mechanisms in the proposed framework enables the scheduling policy to have good generalization capability and evolutionary adaptability. Based on Julia language, this paper conducted modeling, training and Monte Carlo simulation analysis on an industrial-scale application case and provided rich results, which demonstrates the feasibility and effectiveness of the proposed framework.

This paper delineates a viable blueprint for propelling refinery production scheduling under uncertainty toward data-driven multistage decision-making, which opens up a wide space for future research. There are some promising directions for further development and improvement. In terms of modeling, the process model of each production unit under each operation mode adopted by the application case in this paper is succinct single-value model, and future work could extend it to decision-dependent functional models or probabilistic models to characterize different application scenarios. In terms of model solving or training, the computational burden brought by more complex models could be tried to mitigate through parallel High-Performance Computing (HPC). In addition, since the use of the SDDP algorithm has some caveats, such as the relatively complete recourse assumption and the stagewise-independent uncertainty assumption, corresponding countermeasures could be explored in the future to further broaden the scope of application of the proposed framework.

**Acknowledgements**

The author would like to thank Dr. Oscar Dowson, the main contributor of `SDDP.jl`, for the many useful discssions with him that inspired the author a lot.

The author also would like to thank Dr. Ying Guan, research fellow at Harvard University, for providing substantial support.

## Appendix A. Notation

**Labels in Figure 1**

| | |
|---|---|
| RTK | Raw petroleum storage tank. |
| ADU | Atmospheric distillation unit. |

| | |
|---|---|
| VDU | Vacuum distillation unit. |
| FCCU | Fluid catalytic cracking unit. |
| HDSU | Hydrodesulfurization unit. |
| ETHU | Etherification unit. |
| HTU1 | Hydrotreating unit #1. |
| HTU2 | Hydrotreating unit #2. |
| MTBEU | Methyl tert-butyl ether production unit. |
| RFU | Catalytic reforming unit. |
| COM1 | Dummy unit combiner #1. |
| COM2 | Dummy unit combiner #2. |
| SPL1 | Dummy unit splitter #1. |
| SPL2 | Dummy unit splitter #2. |
| C5CTK | Refined C5 component oil storage tank. |
| RFCTK | Catalytic reformed gasoline storage tank. |
| MTBECTK | Methyl tert-butyl ether storage tank. |
| GIII0PTK | Product diesel GIII $0^{\#}$ storage tank. |
| GIIIM10PTK | Product diesel GIII $-10^{\#}$ storage tank. |
| GIV0PTK | Product diesel GIV $0^{\#}$ storage tank. |
| GIII90PTK | Product gasoline GIII $90^{\#}$ storage tank. |
| GIII93PTK | Product gasoline GIII $93^{\#}$ storage tank. |
| GIII97PTK | Product gasoline GIII $97^{\#}$ storage tank. |
| JIV93PTK | Product gasoline JIV $93^{\#}$ storage tank. |
| JIV97PTK | Product gasoline JIV $97^{\#}$ storage tank. |

**Indices in Figure 1**

| | |
|---|---|
| M#01 | Raw petroleum. |
| M#02 | Straight-run gasoline. |
| M#03 | Light straight-run diesel. |
| M#04 | Heavy straight-run diesel. |
| M#05 | Atmospheric residue. |
| M#06 | Vacuum distillate. |
| M#07 | Vacuum residue. |
| M#08 | Catalytic cracking diesel. |
| M#09 | Catalytic cracking rich gas. |
| M#10 | Catalytic cracking gasoline. |
| M#11 | Naphtha. |
| M#12 | Refined diesel #1. |
| M#13 | Refined diesel #2. |
| M#14 | Refined C5 component oil. |
| M#15 | Catalytic reformed gasoline. |

| | |
|---|---|
| M#16 | Methyl tert-butyl ether. |
| M#17 | Hydrodesulfurization gasoline. |
| M#18 | Etherified gasoline. |
| M#19 | Product diesel GIII $0^{\#}$. |
| M#20 | Product diesel GIII $-10^{\#}$. |
| M#21 | Product diesel GIV $0^{\#}$. |
| M#22 | Product gasoline GIII $90^{\#}$. |
| M#23 | Product gasoline GIII $93^{\#}$. |
| M#24 | Product gasoline GIII $97^{\#}$. |
| M#25 | Product gasoline JIV $93^{\#}$. |
| M#26 | Product gasoline JIV $97^{\#}$. |

**Sets**

| | |
|---|---|
| $\mathcal{M}^u$ | Index set of the basic steady operation modes of processing unit $u$. |
| $\mathcal{I}^u$ | Input material index set of unit $u$. |
| $\mathcal{O}^u$ | Output material index set of unit $u$. |
| $\mathcal{R}^i, \mathcal{R}^o$ | Key property index set of input/output material $i/o$. |
| $\mathcal{U}$ | Universal set of all units of the refinery. |
| $\mathcal{PRU}$ | Set of all processing units. |
| $\mathcal{BLD}$ | Set of all blenders. |
| $\mathcal{RTK}$ | Set of all raw material tanks. |
| $\mathcal{CTK}$ | Set of all component oil tanks. |
| $\mathcal{PTK}$ | Set of all product oil tanks. |
| $\mathcal{COM}$ | Set of all combiners. |
| $\mathcal{SPL}$ | Set of all spliters. |
| $\mathcal{PPL}$ | Set of all pipelines. |

**Parameters**

| | |
|---|---|
| $\mathbf{T}^u$, $\tau^u_{m,m'}$ | Time constant matrix and its scalar element of processing unit $u$ at operation mode $m$-$m'$. |
| $\mathbf{\Psi}^u_o$, $\psi^u_{o,m,m'}$ | Yield matrix and its scalar element of output material $o$ of processing unit $u$ at operation mode $m$-$m'$. |
| $\mathbf{P}^u_{r,o}$, $\rho^u_{r,o,m,m'}$ | Property matrix and its scalar element of property index $r$, output material $o$ of processing unit $u$ at operation mode $m$-$m'$. |
| $\mathbf{\Phi}^u$, $\phi^u_{m,m'}$ | Operation cost matrix and its scalar element of processing unit $u$ at operation mode $m$-$m'$. |
| $\overline{\tau}^u$ | Memory of the operation mode shifting register of processing unit $u$. |
| $\underline{\kappa}^u_{i,o}$, $\overline{\kappa}^u_{i,o}$ | Required lower and upper bounds of the proportion of component oil $i$ in product oil $o$ that is blended by blender $u$. |

| | |
|---|---|
| $\underline{\rho}^u_{r,o}$, $\overline{\rho}^u_{r,o}$ | Required lower and upper bounds of property index $r$ of product oil $o$ that is blended by blender $u$ . |
| $\alpha^u$ | Unit price of the raw petroleum in raw material tank $u$. |
| $\beta^u$ | Unit storage cost of tank $u$. |
| $\delta^u$ | Demand of the product oil in tank $u$. |
| $\gamma^u$ | Unit penalty of the product oil in tank $u$ if its demand isn't met. |
| $\mu^u$ | Unit selling price of the product oil in tank $u$. |
| $\underline{f}^u_i$, $\overline{f}^u_i$ | Required lower and upper bounds of the flow rate of input material $i$ of unit $u$. |
| $\underline{v}^u$, $\overline{v}^u$ | Required lower and upper bounds of the inventory level of tank $u$. |
| $\Delta t$ | Duration of each stage. |
| $N$ | Total number of stage nodes within each periodic scheduling horizon. |

**Variables**

| | |
|---|---|
| $\mathbf{X}^u$, $x^u_{m,m'}$ | Operation mode indicator matrix and its scalar element that inditates if processing unit $u$ is at operation mode $m$-$m'$. |
| $\mathbf{XR}^u$ | Operation mode shifting register of processing unit $u$. |
| $\mathbf{XR}^u_t$, $xr^u_{t,m,m'}$ | The $t$-th position of the operation mode shifting register and its scalar element. |
| $f^u_i$ | Feed load (i.e., flow rate) of input material $i$ of unit $u$. |
| $g^u_o$ | Flow rate of output material $o$ of unit $u$. |
| $p^u_{r,i}$ | Value of property index $r$, input material $i$ of unit $u$. |
| $q^u_{r,o}$ | Value of property index $r$, output material $o$ of unit $u$. |
| $z^u_{\text{opr}}$ | Operation cost of processing unit $u$. |
| $w^u_{i,o}$ | Blending load (i.e., allocated flow rate) of blender $u$ when its input component oil $i$ is used to blend output product oil $o$. |
| $z^u_{\text{mat}}$ | Raw material cost of raw material tank $u$. |
| $v^{\text{in},u}$, $v^{\text{out},u}$ | Inventory levels of tank $u$ at the begining and the end of current stage. |
| $s^u$ | Delivery amount (i.e., sales amount) of the product oil in tank $u$. |
| $z^u_{\text{sto}}$ | Storage cost of tank $u$. |
| $z^u_{\text{pen}}$ | Penalty cost of the product oil in tank $u$ if its demand isn't met. |
| $z^u_{\text{inc}}$ | Sales income of the product oil in tank $u$. |
| $z_{\text{cost}}$ | Total production cost of the refinery during current stage. |
| $z_{\text{profit}}$ | Total profit of the refinery during current stage. |

## Appendix B. Supplementary Materials for Section 5

The supplementary materials include the given historical demand data, the trained scheduling policy, and the results of both the in-sample and out-of-sample Monte Carlo simulations. There are also spaghetti plots of all the control and state variables, which makes it possible for the reader to highlight individual trajectories in an interactive way by hovering the cursor over them. These materials are available at `https://github.com/hanb16/Han2025Data`.

## References


Al-Shafei, A., Amjady, N., Zareipour, H., Cao, Y., 2025. Power system transition planning: An industry-aligned framework for long-term optimization. arXiv preprint arXiv:2505.01331 doi:10.48550/arXiv.2505.01331.

Al-Shammari, A., Ba-Shammakh, M.S., 2011. Uncertainty analysis for refinery production planning. Industrial & Engineering Chemistry Research 50, 7065–7072. doi:10.1021/ie200313r.

Alhajri, I., Elkamel, A., Albahri, T., Douglas, P., 2008. A nonlinear programming model for refinery planning and optimisation with rigorous process models and product quality specifications. International Journal of Oil, Gas and Coal Technology 1, 283–307. doi:10.1504/IJOGCT.2008.019846.

Andrés-Martínez, O., Ricardez-Sandoval, L.A., 2022. Integration of planning, scheduling, and control: A review and new perspectives. The Canadian Journal of Chemical Engineering 100, 2057–2070. doi:10.1002/cjce.24501.

Artzner, P., Delbaen, F., Eber, J., Heath, D., 1999. Coherent measures of risk. Mathematical Finance 9, 203–228. doi:10.1111/1467-9965.00068.

Bellman, R., 1954. The theory of dynamic programming. Bulletin of the American Mathematical Society 60, 503–515. doi:10.1090/S0002-9904-1954-09848-8.

Ben-Tal, A., Nemirovski, A., El Ghaoui, L., 2009. Robust Optimization. Princeton University Press.

Bertsekas, D., 2012a. Dynamic Programming and Optimal Control: Volume I. volume 4. Athena Scientific.

Bertsekas, D., 2012b. Dynamic Programming and Optimal Control: Volume II: Approximate Dynamic Programming. volume 4. Athena Scientific.

Bertsekas, D., 2019. Reinforcement Learning and Optimal Control. volume 1. Athena Scientific, Belmont, Massachusetts.

Bertsimas, D., Sim, M., 2004. The price of robustness. Operations Research 52, 35–53. doi:10.1287/opre.1030.0065.

Birge, J.R., Louveaux, F., 2011. Introduction to Stochastic Programming. Springer Science & Business Media.

Chen, Y., Ding, J., Chai, T., 2021. A knowledge transfer based scheduling algorithm for large-scale refinery production. IEEE Transactions on Industrial Informatics 18, 869–879. doi:10.1109/tii.2021.3074157.

Chen, Y., Ding, J., Chen, Q., 2024. A reinforcement learning based large-scale refinery production scheduling algorithm. IEEE Transactions on Automation Science and Engineering 21, 6041–6055. doi:10.1109/tase.2023.3321612.

Chen, Y., Jiang, Z., 2024. A refinery production scheduling model with operation mode transition based on variable driven modelling approach. Journal of Chemical Engineering of Japan 57, 2379339. doi:10.1080/00219592.2024.2379339.

Dowson, O., 2020. The policy graph decomposition of multistage stochastic programming problems. Networks 76, 3–23. doi:10.1002/net.21932.

Dowson, O., Kapelevich, L., 2021. SDDP.jl: a Julia package for stochastic dual dynamic programming. INFORMS Journal on Computing 33, 27–33. doi:10.1287/ijoc.2020.0987.

Dowson, O., Morton, D.P., Pagnoncelli, B.K., 2020. Partially observable multistage stochastic programming. Operations Research Letters 48, 505–512. doi:10.1016/j.orl.2020.06.005.

Dowson, O., Morton, D.P., Pagnoncelli, B.K., 2022. Incorporating convex risk measures into multistage stochastic programming algorithms. Annals of Operations Research doi:10.1007/s10479-022-04977-w.

Feng, W., Feng, Y., Zhang, Q., 2021. Multistage robust mixed-integer optimization under endogenous uncertainty. European Journal of Operational Research 294, 460–475. doi:10.1016/j.ejor.2021.01.048.

Föllmer, H., Schied, A., 2002. Convex measures of risk and trading constraints. Finance and Stochastics 6, 429–447. doi:10.1007/s007800200072.

Gao, X., 2014. Scheduling Optimization Methods based on Optimal Process Control. Ph.D. thesis. Tsinghua University.

Gao, X., Jiang, Y., Chen, T., Huang, D., 2015. Optimizing scheduling of refinery operations based on piecewise linear models. Computers & Chemical Engineering 75, 105–119. doi:10.1016/j.compchemeng.2015.01.022.

Gao, X., Jiang, Y., Huang, D., 2016. Process modelling based on integration of unitwide optimal process control and plantwide scheduling. CIESC Journal 67, 5105–5111. doi:10.11949/j.issn.0438-1157.20161200.

Gao, X., Shang, C., Jiang, Y., Huang, D., Chen, T., 2014. Refinery scheduling with varying crude: A deep belief network classification and multimodel approach. AIChE Journal 60, 2525–2532. doi:10.1002/aic.14455.

Ge, D., Huangfu, Q., Wang, Z., Wu, J., Ye, Y., 2022. Cardinal Optimizer (COPT) user guide. https://guide.coap.online/copt/en-doc.

Girardeau, P., Leclere, V., Philpott, A.B., 2015. On the convergence of decomposition methods for multistage stochastic convex programs. Mathematics of Operations Research 40, 130–145. doi:10.1287/moor.2014.0664.

Göthe-Lundgren, M., Lundgren, J.T., Persson, J.A., 2002. An optimization model for refinery production scheduling. International Journal of Production Economics 78, 255–270. doi:10.1016/s0925-5273(00)00162-6.

Guevara, E., Babonneau, F., Homem-de Mello, T., 2024. Uncertainty dynamics in energy planning models: An autoregressive and

markov chain modeling approach. Computers & Industrial Engineering 191, 110084. URL: https://www.semanticscholar.org/paper/a023ca1873ab023c9d735cae5819083448b2abab, doi:10.1016/j.cie.2024.110084.

Guigues, V., 2016. Convergence analysis of sampling-based decomposition methods for risk-averse multistage stochastic convex programs. SIAM Journal on Optimization 26, 2468–2494. doi:10.1137/140983136.

Han, B., 2022. Research on Several Problems of Refining Production Planning and Scheduling under Uncertainty. Ph.D. thesis. Tsinghua University.

Han, B., 2024. Multiple kernel learning-aided column-and-constraint generation method. Optimization Online URL: https://optimization-online.org/?p=28865.

Han, B., Shang, C., Huang, D., 2021. Multiple kernel learning-aided robust optimization: Learning algorithm, computational tractability, and usage in multi-stage decision-making. European Journal of Operational Research 292, 1004–1018. doi:10.1016/j.ejor.2020.11.027.

Han, B., Shang, C., Jiang, Y., Huang, D., 2022. Object-oriented refinery plant-wide scheduling optimization model and program framework. CIESC Journal 73, 1623. doi:10.11949/0438-1157.20211737.

He, W., Zhao, J., Zhao, L., Li, Z., Yang, M., Liu, T., 2023. Data-driven two-stage distributionally robust optimization for refinery planning under uncertainty. Chemical Engineering Science 269, 118466. doi:10.1016/j.ces.2023.118466.

Hou, X. (Ed.), 1991. Chinese Refining Technology. China Petrochemical Press, Beijing.

Jia, Z., Ierapetritou, M., 2004. Efficient short-term scheduling of refinery operations based on a continuous time formulation. Computers & chemical engineering 28, 1001–1019. doi:10.1016/j.compchemeng.2003.09.007.

Joly, M., 2012. Refinery production planning and scheduling: The refining core business. Brazilian Journal of Chemical Engineering 29, 371–384. doi:10.1590/s0104-66322012000200017.

Joly, M., Moro, L.F.L., Pinto, J.M., 2002. Planning and scheduling for petroleum refineries using mathematical programming. Brazilian Journal of Chemical Engineering 19, 207–228. doi:10.1590/s0104-66322002000200008.

Joly, M., Odloak, D., Miyake, M.Y., Menezes, B.C., Kelly, J.D., 2018. Refinery production scheduling toward industry 4.0. Frontiers of Engineering Management 5, 202–213. doi:10.15302/j-fem-2017024.

Kaufmann, T., Weng, P., Bengs, V., Hüllermeier, E., 2024. A survey of reinforcement learning from human feedback. arXiv preprint arXiv:2312.14925 URL: https://arxiv.org/abs/2312.14925, doi:10.48550/arXiv.2312.14925.

Khor, C.S., Varvarezos, D., 2017. Petroleum refinery optimization. Optimization and Engineering 18, 943–989. doi:10.1007/s11081-016-9338-x.

King, A.J., Rockafellar, R.T., 1993. Asymptotic theory for solutions in statistical estimation and stochastic programming. Mathematics of Operations Research 18, 148–162. doi:10.1287/moor.18.1.148.

Lambert, N., 2025. Reinforcement learning from human feedback. arXiv preprint arXiv:2504.12501 URL: https://arxiv.org/abs/2504.12501, doi:10.48550/arXiv.2504.12501.

Lappas, N.H., Gounaris, C.E., 2018. Robust optimization for decision-making under endogenous uncertainty. Computers & Chemical Engineering 111, 252–266. doi:10.1016/j.compchemeng.2018.01.006.

Larraz, R., 2021. A brief history of oil refining. Substantia 5, 129–152. doi:10.36253/substantia-1191.

Leiras, A., Ribas, G., Hamacher, S., Elkamel, A., 2011. Literature review of oil refineries planning under uncertainty. International Journal of Oil, Gas and Coal Technology 4, 156–173. doi:10.1504/ijogct.2011.038927.

Li, C., Grossmann, I.E., 2021. A review of stochastic programming methods for optimization of process systems under uncertainty. Frontiers in Chemical Engineering 2, 622241. URL: https://www.semanticscholar.org/paper/9fef5a10c644cf744bc88b3ad6b330afcdde5fd1, doi:10.3389/fceng.2020.622241.

Li, Z., Lin, R., Su, H., Xie, L., 2025. Reinforcement learning-driven plant-wide refinery planning using model decomposition. arXiv preprint arXiv:2504.08642 doi:10.48550/arXiv.2504.08642.

Liu, B., 2010. Uncertainty Theory: A Branch of Mathematics for Modeling Human Uncertainty. Studies in Computational Intelligence, Springer Berlin, Heidelberg. doi:10.1007/978-3-642-13959-8.

Liu, Y., He, W., Zhao, L., 2022. Data-driven robust optimization with multiple kernel learning for refinery planning under price uncertainty, in: 2022 Australian & New Zealand Control Conference (ANZCC), IEEE. IEEE. pp. 81–86. doi:10.1109/anzcc56036.2022.9966966.

Luo, C., Rong, G., 2007. Hierarchical approach for short-term scheduling in refineries. Industrial & Engineering Chemistry Research 46, 3656–3668. doi:10.1021/ie061354n.

Lyu, W., Zhu, Y., Huang, D., Jiang, Y., Jin, Y., 2010. A new strategy of integrated control and on-line optimization on high-purity distillation process. Chinese Journal of Chemical Engineering 18, 66–79. doi:10.1016/s1004-9541(08)60325-0.

Maggio, M., Terraneo, F., Leva, A., 2014. Task scheduling: A control-theoretical viewpoint for a general and flexible solution. ACM Transactions on Embedded Computing Systems 13, 1–22. doi:10.1145/2560015.

Homem-de Mello, T., Bayraksan, G., 2014. Monte carlo sampling-based methods for stochastic optimization. Surveys in Operations Research and Management Science 19, 56–85. doi:10.1016/j.sorms.2014.05.001.

Menezes, B.C., Kelly, J.D., Grossmann, I.E., 2013. Improved swing-cut modeling for planning and scheduling of oil-refinery distillation units. Industrial & Engineering Chemistry Research 52, 18324–18333. doi:10.1021/ie4025775.

Motamed Nasab, F., Li, Z., 2021. Multistage adaptive stochastic mixed integer optimization under endogenous and exogenous uncertainty. AIChE Journal 67, e17333. doi:10.1002/aic.17333.

Ou, W., Bi, S., 2024. Sequential decision-making under uncertainty: A robust mdps review. arXiv preprint arXiv:2404.00940 doi:10.48550/arXiv.2404.00940.

Pereira, M.V.F., Pinto, L.M.V.G., 1991. Multi-stage stochastic optimization applied to energy planning. Mathematical Programming 52, 359–375. doi:10.1007/bf01582895.

Philpott, A., Guan, Z., 2008. On the convergence of sampling-based methods for multi-stage stochastic linear programs. Operations Research Letters 36, 450–455.

Pinto, J.M., Joly, M., Moro, L.F.L., 2000. Planning and scheduling models for refinery operations. Computers & Chemical Engineering 24, 2259–2276. doi:10.1016/s0098-1354(00)00571-8.

Powell, W.B., 2011. Approximate Dynamic Programming: Solving the Curses of Dimensionality. Wiley, Hoboken, NJ. doi:10.1002/9781118029176.

Puterman, M.L., 1994. Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley & Sons, New York, NY. doi:10.1002/9780470316887.

Rockafellar, R., Uryasev, S., 2002. Conditional value-at-risk for general loss distributions. Journal of Banking & Finance 26, 1443–1471. doi:10.1016/s0378-4266(02)00271-6.

Rockafellar, R.T., 2007. Coherent approaches to risk in optimization under uncertainty, in: OR Tools and Applications: Glimpses of Future Technologies. INFORMS, pp. 38–61. doi:10.1287/educ.1073.0032.

Shah, N., Saharidis, G.K., Jia, Z., Ierapetritou, M.G., 2009. Centralized–decentralized optimization for refinery scheduling. Computers & Chemical Engineering 33, 2091–2105. doi:10.1016/j.compchemeng.2009.06.010.

Shah, N.K., Ierapetritou, M.G., 2011. Short-term scheduling of a large-scale oil-refinery operations: Incorporating logistics details. AIChE Journal 57, 1570–1584. doi:10.1002/aic.12359.

Shapiro, A., 1993. Asymptotic behavior of optimal solutions in stochastic programming. Mathematics of Operations Research 18, 829–845. doi:10.1287/moor.18.4.829.

Shapiro, A., Dentcheva, D., Ruszczynski, A., 2021. Lectures on Stochastic Programming: Modeling and Theory. SIAM.

Shi, L., Jiang, Y., Wang, L., Huang, D., 2014. Refinery production scheduling involving operational transitions of mode switching under predictive control system. Industrial & Engineering Chemistry Research 53, 8155–8170. doi:10.1021/ie500233k.

Shi, L., Jiang, Y., Wang, L., Huang, D., 2015. Efficient Lagrangian decomposition approach for solving refinery production scheduling problems involving operational transitions of mode switching. Industrial & Engineering Chemistry Research 54, 6508–6526. doi:10.1021/acs.iecr.5b00290.

Shobrys, D.E., White, D.C., 2002. Planning, scheduling and control systems: why cannot they work together. Computers & Chemical Engineering 26, 149–160. doi:10.1016/s0098-1354(01)00737-2.

Street, A., Valladão, D., Lawson, A., Velloso, A., 2020. Assessing the cost of the hazard-decision simplification in multistage stochastic hydrothermal scheduling. Applied Energy 280, 115939. doi:10.1016/j.apenergy.2020.115939.

Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning. Adaptive computation and machine learning. second edition ed., MIT Press, Cambridge, MA.

Wang, C., Peng, X., Shang, C., Fan, C., Zhao, L., Zhong, W., 2021. A deep learning-based robust optimization approach for refinery planning under uncertainty. Computers & Chemical Engineering 155, 107495. doi:10.1016/j.compchemeng.2021.107495.

Yang, Y., Barton, P.I., 2016. Integrated crude selection and refinery optimization under uncertainty. AIChE Journal 62, 1038–1053. doi:10.1002/aic.15075.

Yuan, Z., Qin, W., Zhao, J., 2017. Smart manufacturing for the oil refining and petrochemical industry. Engineering 3, 179–182. doi:10.1016/j.eng.2017.02.012.

Zhang, J., Zhu, X.X., Towler, G.P., 2001. A level-by-level debottlenecking approach in refinery operation. Industrial & Engineering Chemistry Research 40, 1528–1540. doi:10.1021/ie990854w.

Zhang, L., Jiang, Y., Huang, D., 2015. Alternative formulations and improvements with valid inequalities for the refinery production scheduling problem involving operational transitions. Industrial & Engineering Chemistry Research 54, 7871–7889. doi:10.1021/acs.iecr.5b00600.

Zhang, L., Yuan, Z., Chen, B., 2022. Adjustable robust optimization for the multi-period planning operations of an integrated refinery-petrochemical site under uncertainty. Computers & Chemical Engineering 160, 107703. doi:10.1016/j.compchemeng.2022.107703.

Zhang, Q., Feng, W., 2020. A unified framework for adjustable robust optimization with endogenous uncertainty. AIChE Journal 66, e17047. doi:10.1002/aic.17047.

Zhang, Z., Wang, D., Yang, H., Si, S., 2023. A review of sequential decision making via simulation. arXiv preprint arXiv:2312.04090 URL: https://arxiv.org/pdf/2312.04090, doi:10.48550/arXiv.2312.04090.

Zou, J., Ahmed, S., Sun, X.A., 2019. Stochastic dual dynamic integer programming. Mathematical Programming 175, 461–502. doi:10.1007/s10107-018-1249-5.