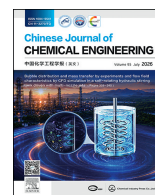




Contents lists available at ScienceDirect

Chinese Journal of Chemical Engineering

journal homepage: www.elsevier.com/locate/CJChE

Full Length Article

Data-driven robust optimization for integrated refinery planning and blending scheduling under multi-mode uncertainty

Wenhao Yu¹, Xin Dai^{1,*}, Yuanhang Yue², Xinwei Lin¹, Minglei Yang^{1,*}¹ Key Laboratory of Smart Manufacturing in Energy Chemical Process, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China² Guangdong Petrochemical Co., Ltd, Jieyang 522031, China

ARTICLE INFO

Article history:

Received 18 June 2025

Received in revised form

7 February 2026

Accepted 25 February 2026

Available online 16 April 2026

Keywords:

Process system engineering

Global optimization

MINLP

Product blending scheduling

Refinery

ABSTRACT

Integrating refinery planning and product blending scheduling under uncertainty is challenging due to multi-mode operating conditions and mismatched temporal scales. This work proposes a bilevel optimization framework that explicitly coordinates long-term refinery planning with short-term blending scheduling. To represent mode-dependent uncertainty in blending behavior, a data-driven robust optimization (DDRO) strategy is developed, in which mode-associated uncertainty sets are constructed from historical data using a multi-stage clustering (MSC) algorithm combined with weighted support vector machines (SVM). This enables uncertainty modeling that is adaptive to operating modes and less conservative than conventional holistic robust approaches. To solve the resulting bilevel robust model efficiently, a robust-buffer-based decomposition (RBD) algorithm is proposed, which iteratively introduces a specification buffer at the planning level to preserve scheduling feasibility under uncertainty. In the case study, the proposed multi-mode robust framework improves the net profit by approximately 1.0% compared to holistic robust model, while maintaining 100% feasibility in quality constraints and reducing computational time by more than 50% relative to the single-level formulation. These results demonstrate that the proposed framework effectively balances economic performance, robustness, and computational tractability in integrated refinery planning and scheduling.

© 2026 The Chemical Industry and Engineering Society of China, and Chemical Industry Press Co., Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

1. Introduction

Planning and scheduling constitute fundamental tasks in process systems engineering, operating across distinct temporal scales. Planning focuses on long-term (seasonal/monthly) allocation of production and distribution resources, whereas scheduling manages short-term (weekly/daily) execution of manufacturing operations [1]. These problems in refinery industry are typically formulated as mixed-integer programming (MIP) models, solvable through both emerging data-driven approaches and conventional mathematical programming methods. The evolution of computational capabilities has accelerated the adoption of data-driven or derivative-free techniques [2,3]. For instance, Dias and Ierapetritou [4] proposed a framework incorporating data-driven

feasibility analysis and surrogate modeling to ensure the feasibility of planning and scheduling solutions. Similarly, reinforcement learning has been utilized to solve Markov decision processes associated with planning and scheduling problems [5]. Harjunkoski *et al.* [6] employed data analytics to leverage real-time data, reducing the number of product grade changes, while also integrating reinforcement learning to influence the optimization process in explicit scheduling. Yang *et al.* [7] employed a hybrid approach using GA to expedite the calculation for large-scale MIPs. On the other hand, traditional deterministic methods remain prevalent in industrial practice. Commercial solvers such as CPLEX and Gurobi often address planning and scheduling challenges using techniques like the dual simplex algorithm, branch-and-cut methods, outer approximation, and generalized Benders decomposition. Research has also explored ways to alleviate the computational burden by applying deterministic global optimality-based approaches [8,9]. For instance, Persson and Gothe-Lundgren [10] applied column generation and valid inequalities to address shipment scheduling problems in oil

* Corresponding authors.

E-mail addresses: xindai@ecust.edu.cn (X. Dai), mlyang@ecust.edu.cn (M. Yang).

refineries. Uribe-Rodríguez *et al.* [11] used spatial decomposition methods to solve integrated scheduling problems in refinery-chemical complexes. Franzoi *et al.* [12] adopted a state-of-art optimization framework to address large-scale refinery scheduling problems for industrial use.

While standalone solutions for planning and scheduling have shown theoretical promise, their industrial integration remains challenging due to temporal scale differences and conflicting optimization objectives. A planning solution satisfying refinery-wide mass balances may violate unit-level blending kinetics during scheduling, especially when sequential operations are involved. Researchers have increasingly focused on the field of integrated planning and scheduling to mitigate these mismatches. Some studies have proposed simultaneous methods like monolithic approaches. It tackle different layers of the problem holistically, where feasibility in lower levels is satisfied via explicit incorporation of their constraints [2]. Yang *et al.* [13] addressed the integrated crude-oil scheduling and refinery planning problem using a monolithic model, which was then decomposed into subproblems via Lagrangian decomposition. Similarly, Su *et al.* [14] employed a hybrid continuous-discrete time representation to model and solve similar challenges. Another promising strategy for addressing integration challenges is bilevel programming, where planning being the upper-level while scheduling serves as the lower-level problem. In this context, the planning acts as the leader while scheduling plays the follower, resembling a Stackelberg game [15], where the planning is subject to the optimal solution of the scheduling level [2]. Chu *et al.* [16] presented a bilevel model leveraging an agent-based method to manage scheduling uncertainties. However, due to the computational intractability of bilevel programming, researchers have sought ways to reformulate the bilevel programming problem into a holistic single-level problem. For example, Pineda *et al.* [17] utilized Karush–Kuhn–Tucker (KKT) conditions to replacing the follower problems with big-M constraints, while Beykal *et al.* [18] proposed a framework where bilevel optimization problems are approximated as single-level optimization problems by collecting samples of the upper-level objective and solving the lower-level problem to global optimality at those sampling points. Li and Ierapetritou [19,20] translated bilevel model into a single level problem and further decomposed it into subproblems to alleviate intractability. For single-level problems, the modeling approach is similar to that of monolithic problems and similar solution strategies can be employed.

The integration of planning and scheduling, as discussed earlier, also faces challenges due to uncertainties that influence decision-making across both levels. When addressing these problems, enterprise-wide decision-makers must account for various uncertain factors, such as prices of raw materials, unstable crude oil quality, and fluctuating product demands [21,22]. To address these uncertainties, researchers have implemented various approaches. For example, Yue and You [23] presented a MINLP formulation using the service-guaranteed approach to handle uncertainty in inventory systems. Recent studies on refinery planning and scheduling under uncertainty have largely relied on stochastic programming [14,24]. Chen *et al.* [25] introduced a two-stage stochastic programming framework with flexibility constraints that accounts for market and operational uncertainties. Additionally, efforts have also been made to integrate robust optimization with stochastic programming to strike a balance between profitability and reliability [26,27]. While these methods have shown promise, approaches like stochastic programming and chance-constrained programming often require accurate probability distributions for uncertain parameters, which are difficult to obtain in practice due to data inexplicitness. Furthermore, traditional RO methods using convex sets often lead

to excessive conservatism by encompassing low-probability regions between data modes, while stochastic programming faces computational intractability in high-dimensional, multi-modal scenarios.

To overcome these limitations and address the practical difficulty of obtaining precise probability distributions, data-driven robust uncertainty sets constructed by data-driven robust optimization (DDRO) approaches have emerged as a viable alternative when distribution information is insufficient [28,29]. Long *et al.* [30] proposes a data-driven robust optimization framework integrating PCA-RKDE to optimize coupled energy-material systems in refineries under uncertainty, effectively reducing costs and enhancing solution robustness. Razm *et al.* [31] presents a row-and column-wise robust optimization model for weather uncertainties in biorefinery systems, and enhancing profitability through machine learning-based DDRO. Baghban *et al.* [32] proposes a DDRO approach for refined product pipeline scheduling under injection rate uncertainty, using support vector clustering to construct the uncertainty set. The uncertainty surrounding product demand and price can also be incorporated into models using DDRO approaches [33–35]. By leveraging historical data, DDRO enables the construction of uncertainty sets that are not only more representative of actual variations but also less conservative compared to classical approaches [36]. For example, Ning and You [37] presented a Bayesian machine learning algorithm called the dirichlet process mixture model (DPMM) leveraging the correlated, asymmetrical, multimode nature of the data. Ning and You [38] introduced the weighted generalized intersection kernel (WGIK) into the construction of uncertainty set for high dimensional data which can be adjusted with only one parameter. Based on these advancements, Ghiasvand and Rahmani [36] proposed a novel DDRO approach which enables the adjustment of the boundary reduction rate of uncertainty sets, effectively pulling uncertainty sets towards the predicted points.

While these advances highlight the practical value of DDRO, most existing methods construct holistic uncertainty sets that include all observed data points without distinguishing between operational modes. Such holistic sets, though comprehensive, may introduce unnecessary conservatism due to their broader coverage. For example, although forming the uncertainty set as a union of basic polytopic components [37] represent multimodal distributions, the solution framework still focuses on the worst-case mode, without associating specific decision instances with particular operating modes. In contrast, explicitly assigning mode-specific uncertainty sets at each decision point can improve both precision and robustness. This allows the model to adapt its response based on the mode to which the current operating condition belongs, thereby reducing conservatism and improving solution quality. Hence, mode-aware DDRO offers a more realistic and operationally efficient alternative for dealing with multimodal uncertainty in practical applications.

While Yu *et al.* [39] previously presented a continuous-time formulation for refinery planning and diesel blending scheduling under fluctuating diesel demand with a relaxation-based pruning technique, their approach does not consider coordination with other products or uncertainty in blending properties. This work aims to address mode-dependent uncertainty in decision-making and simultaneously integrates refinery planning and product blending scheduling, which are often challenging to coordinate due to their differing temporal and operational scales in industrial processes. Although robust optimization has been widely applied to planning and scheduling problems separately, the structure in which the upper-level (leader) problem is deterministic while the lower-level (follower) problem incorporates robust optimization has received limited attention. To the best of our knowledge, rigorous solution methods for such bilevel robust frameworks remain largely unexplored [40].

To this end, we propose a bilevel optimization model, where the upper level generates long-term refinery plans, and the lower level determines detailed sequential blending schedules under multimode uncertainty using a DDRO approach. In particular, the follower-level scheduling model captures the synergistic uncertainty among units during sequential operations, which are difficult to represent in high-level planning models. In particular, the follower-level scheduling model captures the synergistic uncertainty among units during sequential operations, which are difficult to represent in high-level planning models. To solve the resulting bilevel robust model, a robust buffer-based decomposition (RBD) approach is proposed. In this framework, a specification buffer variable is embedded into the planning level to preemptively tighten product property constraints. This robust buffer serves as a preemptive tightening margin, reserving flexibility to guard against infeasibility during the scheduling stage, where operational details and uncertainties may render a nominally feasible plan unimplementable. The robust buffer is adjusted iteratively through an optimization loop between the planning and scheduling levels, where feasibility information from the scheduling model informs constraint tightening at the planning level, enabling improved coordination and reduced conservatism.

This paper proposes a novel approach to integrate refinery planning and scheduling by combining bilevel programming with robust optimization in the lower-layer effectively addressing the multi-mode uncertainties in scheduling. The key contributions of this work are as follows:

- ① A bilevel optimization framework is proposed to integrate long-term refinery planning with short-term product blending scheduling under mode-dependent uncertainty.
- ② Mode-aware, data-driven uncertainty sets are constructed and refined by multi stage clustering (MSC) algorithm to accurately capture operational variability while reducing conservatism in robust optimization.
- ③ A DDRO model is developed by embedding the constructed uncertainty sets into the scheduling formulation to address multimodal uncertainties in sequential blending operations.

- ④ An RBD approach is proposed to enhance coordination and reduce conservatism through iterative adjustment of a preemptive robust buffer variable.

The remainder of this paper is organized as follows:

Section 2 defines the integrated problems. Section 3 introduces the deterministic formulation for both plan and product blending scheduling, and section 4 further reformulate the deterministic model into data-driven optimization model that are used for handling uncertainty. Section 5 presents integrated framework and proposed RBD computational strategy. Case study is conducted in Section 6 and we conclude our study in Section 7.

2. Problem Statement

This study focuses on refinery optimization integrated with blending scheduling under the uncertainty of blending effects on gasoline over a planning horizon. The refinery structure follows that of a previous refinery planning model study, with the simplified flowsheet shown in Fig. 1 [9,41]. The planning horizon spans from one week to 15 days, during which the crude oil recipe and cut scheme are assumed to remain fixed, as no new vessels are expected to arrive within such a short timeframe. For multi-mode units, we assume that each mode can be used only once within a single planning horizon, as frequent mode switching complicates actual operations.

The problem can be described as follows.

2.1. Given

A planning horizon of a specific length.

A set of crude oils, their TBP curves, the pseudocomponent profiles, and supply profiles (e.g., costs and available amounts).

A set of final products (e.g., LPG, gasoline, kerosene, diesel, fuel oil, coke), their quality specifications, prices and demand in the planning horizon.

A set of storage tanks and their initial, minimum, and maximum initial inventories.

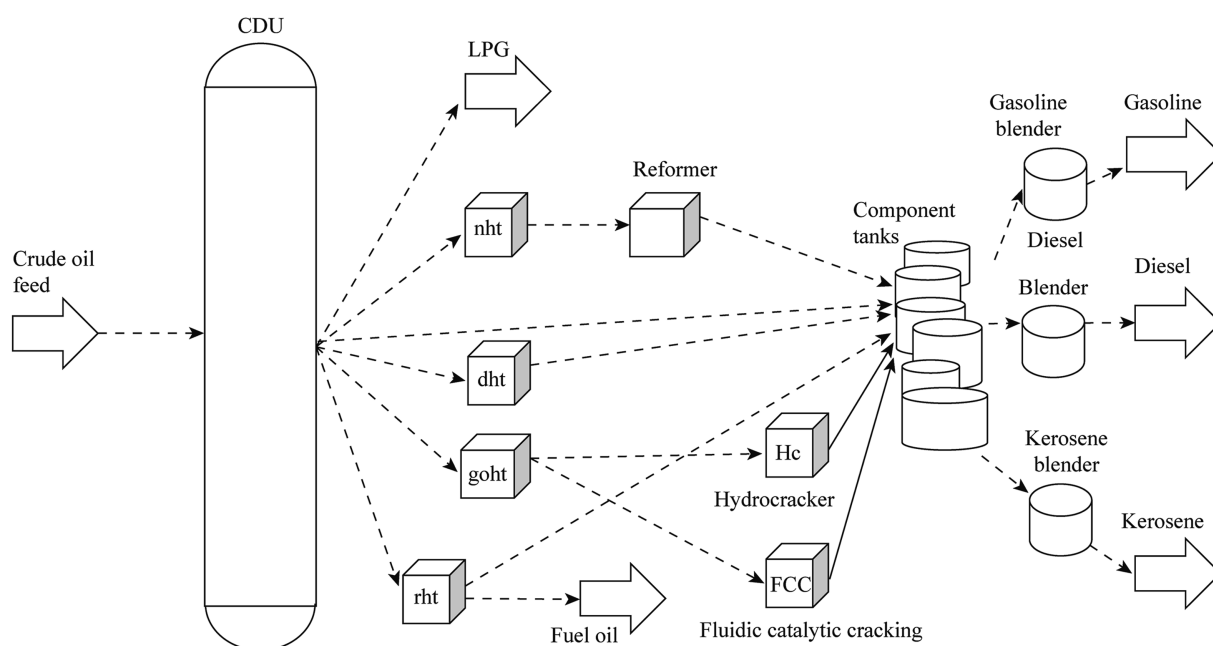


Fig. 1. Simplified flowsheet of the refinery model [9].

A set of blenders and their minimum and maximum blending rates.

Processing units (i.e., CDU, hydrotreaters, catalytic reformer, fluid catalytic cracker, etc.) with minimum and maximum throughput rates and their operating modes.

The topology of all the processing units and storage tanks.

Historical RON data of produced gasoline and corresponding recipe under different unit modes.

2.2. Determine

The amount of each crude oil feedstock to be used.

The production volumes in each processing unit, each operating mode.

The final inventories of the storage tanks.

The volume blended of final products in each blender.

(Scheduling) The volume blended of products in each blender during each time slot.

(Scheduling) The inventories of the storage tanks involved in blending at the end of each time slot.

2.3. While

Maximizing the profit (sales revenue minus total cost). (upper-level objective).

Minimizing the scheduling profit discrepancy from planning model. (lower-level objective).

3. Deterministic Optimization Model for Integrated Problem

3.1. Refinery planning model

3.1.1. Objective function

The profit in Eq. (1) serves as the objective function for the refinery planning problem. The total cost includes material costs, unit operating costs, and inventory costs. In Eqs. (2) and (3), the parameters PS_s , CS_s , CU_i , and IC_t represent price and cost values. The sets DS and SS are subsets of streams S , denoting the destination and source streams, respectively. The sets PU , CDU , and HTU are subsets of units U , representing the processing units, crude oil distillation units, and hydrotreating units. The inventory level of a tank t at the end of the planning horizon is a decision variable, while the initial inventory is denoted by the parameter.

$$\text{max profit} = \text{revenue} - \text{totalcost} \quad (1)$$

$$\text{revenue} = \sum_{s \in DS} PS_s \times V_s \quad (2)$$

$$\text{totalcost} = \sum_{s \in SS} CS_s \times V_s + \sum_{i \in PU, CDU, HTU} CU_i \times V_i + \sum_{t \in T} IC_t \times (V_t - V_t^i) \quad (3)$$

3.1.2. Scheduling models

The volumetric flow rates of the supply streams are restricted within the minimum and maximum available limits (Eq. (4)). The characteristics of these streams, including the TBP temperatures of the crude oils, are fixed at known values (Eq. (5)).

$$A_s^{\min} \leq V_s \leq A_s^{\max} \quad s \in SS \quad (4)$$

$$Q_{s,q} = Q_{s,q}^{\text{fix}} \quad s \in SS, q : (s, q) \in SQ \quad (5)$$

3.1.3. Demand constraints

The volumetric flow rates of the demand streams are constrained within the range of minimum and maximum demand requirements (Eq. (6)). Quality restrictions for the product demands are applied at the blenders, where the qualities of the blended products must stay within the specified minimum and maximum limits.

$$D_s^{\min} \leq V_s \leq D_s^{\max} \quad s \in DS \quad (6)$$

3.1.4. Blender constraints

The blending processes are governed by material balances outlined in Eqs. (7)–(16). The index s' is used as an alias for s , representing all stream indices. The stream s' denotes the outlet stream corresponding to the blended product, while the stream s is the inlet stream representing a blend component. Therefore, $V_{s,s'}$ refers to the volume of blend component s being incorporated into product s' . Eqs. (7) and (8) represent the mass balance between inlet stream and outlet product. A single blender may handle multiple product grades, each with different quality specifications. When a product is blended, the binary variable b_s is set to 1, and 0 otherwise (Eqs. (9) and (10)). This allows the optimization model to decide which products are blended. In a planning scenario, if blending (or non-blending) is mandated, the binary variables b_s are treated as parameters, either set to 1 or 0. $R_b^{\min}$, $R_b^{\max}$ denote the minimum and maximum blending rates during the planning period, and are non-negative values. Eq. (11) enforces that the upper blending capacity is not violated. Product composition constraints are described by Eq. (12), where $R_{s,s'}^{\min}$ and $R_{s,s'}^{\max}$ represent the minimum and maximum blend recipe from s to s' . The blending of quality properties is assumed to be linear, either volumetrically or by weight, as shown in Eqs. (13)–(16). Here, $Q_{s',q}^{\min}$ and $Q_{s',q}^{\max}$ are the minimum and maximum quality q specification of stream s' . The auxiliary variables $QM_{s,s',q}$ represent the bilinear product of volume flow and volumetric basis quality, while $QM_{s,s',q}$ represents the bilinear product of mass flow and weight basis quality. The sets VQ , WQ are the subsets of Q , representing the volumetric basis and weight basis qualities. The set BSQ is the subset of SQ , containing the qualities associated with the streams involved in blending.

$$V_{s'} = \sum_{s: (b,s) \in BI} V_{s,s'} \quad s' : (b,s') \in BO \quad (7)$$

$$\sum_{s': (b,s') \in BO} V_{s,s'} = V_s \quad s : (b,s) \in BI \quad (8)$$

$$V_s \geq V_s^{\min} \times b_s \quad s : (b,s) \in BO \quad (9)$$

$$V_s \leq R_b^{\max} \times L \times b_s \quad b, s : (b,s) \in BO \quad (10)$$

$$\sum_{s: (b,s) \in BO} V_s \leq R_b^{\max} \times L \quad b \in B \quad (11)$$

$$\begin{cases} R_{s,s'}^{\min} \times V_{s'} \leq V_{s,s'} \leq R_{s,s'}^{\max} \times V_{s'} \\ s : (b,s) \in BI, s' : (b,s') \in BO \end{cases} \quad (12)$$

$$\begin{cases} QV_{s,s',q} = V_{s,s'} \times Q_{s,q} \\ s : (b,s) \in BI, s' : (b,s') \in BO, \\ q \in \{q|q \in VQ \wedge q : (s,q) \in SQB\} \end{cases} \quad (13)$$

$$\begin{cases} QM_{s,s',q} = QV_{s,s'',sg'} \times Q_{s,q} \\ s : (b,s) \in BI, s' : (b,s') \in BO, \\ q \in \{q|q \in WQ \wedge q : (s,q) \in SQB\} \end{cases} \quad (14)$$

$$\begin{cases} Q_{s',q}^{\min} \times V_{s'} \leq \sum_{s:(b,s) \in BO} QV_{s,s',q} \leq Q_{s',q}^{\max} \times V_{s'} \\ s' : (b,s') \in BO, \\ q \in \{q|q \in WQ \wedge q : (s',q) \in SQB\} \end{cases} \quad (15)$$

$$\begin{cases} Q_{s',q}^{\min} \times \sum_{s:(b,s) \in BI} QV_{s,s'',sg'} \leq \sum_{s:(b,s) \in BI} QM_{s,s',q} \\ \leq Q_{s',q}^{\max} \times \sum_{s:(b,s) \in BI} QV_{s,s'',sg'} \\ s' : (b,s') \in BO \\ q \in \{q|q \in VQ \wedge q : (s',q) \in SQB\} \end{cases} \quad (16)$$

3.1.5. Storage constraints

A storage tank, initially holding inventory V_t^i is subject to constraints on its inventory levels, which are bounded by the minimum and maximum limits ($V_t^{\min}$ and $V_t^{\max}$). The total withdrawal rate is also restricted within a range, between the minimum and maximum rate ($R_t^{\min}$ and $R_t^{\max}$). Thereby Eqs. (17) and (18) set limits to the tank storage volume and inlet streams. The outlet stream's volumetric flow and quality are determined by Eqs. (19)–(21), where $Q_{t,q}^i$ is the initial quality. It is assumed that ideal mixing rules are valid and that the mixing in the tank is perfect.

$$V_t^{\min} \leq V_t \leq V_t^{\max} \quad t \in T \quad (17)$$

$$R_t^{\min} \times L \leq \sum_{s:(t,s) \in TO} V_s \leq R_t^{\max} \times L \quad t \in T \quad (18)$$

$$\sum_{s:(t,s) \in TI} V_s + V_t^i = V_t + \sum_{s':(t,s') \in TO} V_{s'} \quad t \in T \quad (19)$$

$$\begin{aligned} & \sum_{s:(t,s) \in TI} V_s \times Q_{s,q} + V_t^i \times Q_{t,q}^i \\ &= V_t \times Q_{s',q} + \sum_{s':(t,s') \in TO} V_{s'} \times Q_{s',q} \end{aligned} \quad (20)$$

$$t \in T, q \in \{q|q \in VQ \wedge q : (s,q) \in SQ\}$$

$$\begin{aligned} & \sum_{s:(t,s) \in TI} V_s \times Q_{s,q} \times Q_{s',sg'} + V_t^i \times Q_{t,q}^i \times Q_{t',sg'}^i \\ &= V_t \times Q_{s',q} \times Q_{s',sg'} + \sum_{s',(t,s') \in TO} V_{s'} \times Q_{s',q} \times Q_{s',sg'} \end{aligned} \quad (21)$$

$$t \in T, q \in \{q|q \in WQ \wedge q : (s,q) \in SQ\}$$

3.1.6. Units Constraints with a single operating mode

Certain units operating in a single mode belong to the subset SMU within the set of all units $U = \{i\}$, which encompasses the mixers, splitters and hydrotreating units, and others. These units are subject to material balance constraints as well as the capacity

limitations of each unit (Eqs. (22) and (23)), where $R_i^{\min}$ and $R_i^{\max}$ are the minimum and the maximum unit operating rate during the length of the time horizon.

$$V_i = \sum_{s:(u,s) \in UI} V_s \quad i \in SMU \quad (22)$$

$$R_i^{\min} \times L \leq V_i \leq R_i^{\max} \times L \quad i \in SMU \quad (23)$$

$$\sum_{s':(i,s') \in UO} V_{s'} = V_i \quad i \in MIX \vee SPL \quad (24)$$

$$\begin{cases} \sum_{s:(i,s) \in UI} V_s \times Q_{s,q} = \sum_{s':(i,s') \in UO} V_{s'} \times Q_{s',q} \\ q \in \{q|q \in VQ \wedge q : (s,q) \in SQ\} \end{cases} \quad (25)$$

$$\begin{cases} \sum_{s:(i,s) \in UI} V_s \times Q_{s,q} \times Q_{s',sg'} = \sum_{s':(i,s') \in UO} V_{s'} \times Q_{s',q} \times Q_{s',sg'}^i \\ q \in \{q|q \in WQ \wedge q : (s,q) \in SQ\} \end{cases} \quad (26)$$

$$V_{s'} = Y_{s',i}^{\text{fix}} \times V_i \quad s' : (i,s') \in UO, i \in SMU - (MIX \vee SPL) \quad (27)$$

$$Q_{s',q} = Q_{s',q}^{\text{fix}} \quad s' : (i,s') \in UO, q : (s,q) \in SQ \quad (28)$$

The mixers (subset MIX) and the splitters (subset SPL) are constrained by additional material balance and the quality balance, as defined in Eqs. (24)–(26) respectively. For all process units excluding mixers and splitters (and excluding the CDU), the operating mode follows a fixed yield ($Y_{s,i}^{\text{fix}}$). Thus, the outlet stream flow is determined by Eq. (27), while the quality is fixed at specified values (Eq. (28)). The CDU units (subset CDU) are the same as SMU, where the yields and the qualities of the outlet streams are computed using the tri-section CDU model.

3.1.7. Units constraints with multiple operating modes

Units with multiple distinct operating modes belong to the subset RU within the set of all units U . Each operating mode of a real unit ($RU = \{m\}$) is described as a single unit i . The set $MMU = \{(m,i)\}$ defines the relationship between the operating modes and the real units. The material balance for M follows the same form as for SMU (Eq. (29)), but the capacity is constrained by the real unit, which is the summation of all the operation modes (Eq. (30)). The binary variables b_m control the activation of the unit operating modes (Eq. (31)), which means the operating modes are determined by optimization. The outlet flowrate and quality are given by Eqs. (32) and (33). In a planning problem, if the operating modes are forced to be operated (or not), the binary variables b_m are set to be parameters as 1 (or 0). Additionally, the total number of the operating modes is restricted by Eq. (34), which means the multiple modes cannot be operated simultaneously.

$$V_m = \sum_{s:(i,s) \in UI} V_s \quad i : (m,i) \in MMU \quad (29)$$

$$R_m^{\min} \times L \leq \sum_{m:(m,i) \in MMU} V_m \leq R_m^{\max} \times L \quad m \in RU \quad (30)$$

$$R_m^{\min} \times L \times b_i \leq V_m \leq R_m^{\max} \times L \times b_m \quad i : (m,i) \in MMU \quad (31)$$

$$V_{s'} = Y^{\text{fix}} \times V_m \quad s' : (i,s') \in UO, i : (m,i) \in MMU \quad (32)$$

$$Q_{s',q} = Q_{s',q}^{fix} \quad s' : (i, s') \in UO, q : (s, q) \in SQ \quad (33)$$

$$\sum_{m:(m,i) \in MMU} b_m \leq 1 \quad i : (m, i) \in MMU \quad (34)$$

Units operating under different modes can induce multimode behavior in downstream processes and exhibit complex synergistic or competitive interactions with other units. For instance, the fluid catalytic cracker (FCC) may compete with the reformer for upstream VGO, while the hydrogen consumption by the hydrocracker and the hydrogen production by the reformer can cause imbalances in the refinery's hydrogen network. These interactions ultimately impact the blending performance and therefore must be analyzed from a multimode perspective. This work specifically focuses on the multimode influence of the hydrocracker, which will be discussed in detail in Section 6.

3.1.8. Crude oil cut constraints

$$\begin{cases} TBP_{s'} = F_u(ccw_{s'}, cut_{s'}, TBP_s, v_i) \\ s : (i, s) \in UI, s' : (i, s') \in UO, i \in CDU \end{cases} \quad (35)$$

$$Q_{s,q} = g_q(TBP_s, Q_q^{profile}) \quad q \in Q, s : (i, s) \in UO \quad (36)$$

The general mathematical form of the fractionation cutting model for the atmospheric and vacuum crude distillation unit (CDU) can be represented by Eq. (35), where F_i is a series of linear functions corresponding to different TBP points and different distillation products. The true boiling point (TBP) temperature of fractionated outlet product $TBP_{s'}$ is determined by the cumulative cutting width $ccw_{s'}$, the cutting point temperature $cut_{s'}$, the TBP temperature of the mixed crude oil feed TBP_s and the operating variables of the CDU v_i . The TBP temperatures of the distillation products and quality profiles $Q_q^{profile}$ are used to compute the quality of the product by function g_q (Eq. (36)).

3.1.9. Units durations constraints

$$\frac{V_s}{R_m^{max}} \leq TU_{i,m} \leq \frac{V_s}{R_m^{min}} \quad (i, m, s) \in UMO, m \in RU \quad (37)$$

$$\frac{V_s}{R_b^{max}} \leq TB_s \leq \frac{V_s}{R_b^{min}} \quad (b, s) \in BO \quad (38)$$

Eqs. (37) and (38) ensure that the unit or blender operates within the allowable rate limits. These constraints become particularly important when the planning model is integrated with the scheduling model, where the durations of the units are key parameters.

In this way, the leader's problem in our proposed bilevel optimization problem can be written as such:

$$(\text{Leader's problem}) \max \text{ Profit} = \text{Revenue} - \text{Cost}$$

$$\text{s.t. Eqs. (2) - (38)}$$

3.2. Product blending scheduling model

The product blending scheduling model functions as the lower problem within the bilevel optimization framework. This model uses a continuous-time representation, dividing the planning horizon into multiple time slots, each denoted as the ts^{th} time slot. The total number of time slots is $n = |TS|$. Each time slot is defined by the starting time point TP_{ts} and the ending point of this

time slot TP_{ts+1} , resulting in totally $n+1$ time points TP_{ts} , which are continuous variables to be optimized.

In the scheduling model, the planning outcome of the upstream to the blender (including the outlet stream of secondary processing unit) are transferred to the lower layer scheduling model, serving as the parameter V_s^* and Q_s^* . Similarly, the operating time of secondary processing units $TU_{i,m}^*$ or blenders TB_s^* are also passed to the lower-level model.

3.2.1. Objective function

The objective function of the product blending scheduling model quantifies the deviation from the revenue of the blended product, reflecting how closely the scheduling plan aligns with the upper-level planning. The deviation is zero when there is no gap between the plan and the scheduling outcome. Eq. (39) defines the profit loss from the upper planning layer, which is calculated as the product of the product price and the quantity loss (Eq. (39)). $\sum_{ts \in TS} V_s^{ts}$ represents the actual blended product based on the scheduling scheme, which is the sum of the products over all time slots. The term $|V_s^* - \sum_{ts \in TS} V_s^{ts}|$ thus captures the deviation in the blended product between the upper and lower layers, as shown in Eq. (40).

$$\min \text{ lossprofit} = \sum_{s \in DS} PS_s \times \text{loss}_s \quad (39)$$

$$\text{loss}_s = \left| V_s^* - \sum_{ts \in TS} V_s^{ts} \right|, s : (b, s) \in BO \quad (40)$$

Strictly speaking, Eq. (6) is enforced as a hard constraint within the upper-level planning model, ensuring that any generated plan strictly adheres to contract limits ($D_s^{\min}, D_s^{\max}$). However, the overall integrated system exhibits “soft” behavior regarding demand satisfaction. This is because the lower-level scheduling model treats the planned volume (V_s^*) as a reference target to be matched via the objective function (Eq. (39)), rather than a rigid lower bound. This structure allows the system to quantify potential shortfalls (“lost sales”) under uncertainty without rendering the problem infeasible.

3.2.2. Mode transition constraints

In the case where unit i shifts from mode m to m' at time point TP_{ts} , the binary variable $z_{imm',ts}$ is set to 1. Conversely, it is assigned a value of 0 if no such transition takes place.

Each unit can undergo at most one transition at any given time point (Eq. (41)):

$$\sum_{(m,m') \in TR_i} z_{imm',ts} \leq 1 \quad \forall i \in MMU, ts \in TS \quad (41)$$

A binary variable $y_{im,ts}$ is introduced to indicate whether unit i is operating in mode m during the ts^{th} time slot (from time point TP_{ts} to time point TP_{ts+1}). If both $y_{im,ts-1}$ and $y_{im',ts}$ are equal to 1, it implies that a mode switch has occurred at time point TP_{ts} .

The logical relationship governing this is as follows: $y_{im,ts-1} \wedge y_{im',ts} \leftrightarrow z_{imm',ts}$. This can be formulated into a linear equation as follows (Eq. (42)):

$$\sum_{m' \in \overline{TR}_{im}} z_{im'm,ts} - \sum_{m' \in \overline{TR}_{im}} z_{imm',ts} = y_{im,ts} - y_{im,ts-1} \quad \forall (i, m) \in MMU, ts \in TS, ts \neq 1 \quad (42)$$

where the set $\overline{TR}_{im}$ stands for the modes which can change to mode m , while the set $\overline{TR}_{im}$ denotes the modes to which mode m can change.

Eqs. (43)–(45) are formulated to control the duration of each mode for every unit. TR_i denotes the set of modes that can transition between each other.

Eq. (43) enforces that the time difference between the next transition point is not less than the duration of mode m' if a transition to this mode has occurred.

$$TP_{ts} - TP_{ts'} \geq z_{imm',ts'} \times TU_{im'}^* - TU_{im'}^* \times \left(1 - \sum_{(n,n') \in TR_i} z_{inn',ts}\right) \quad \forall i \in MMU, (m, m') \in TR_i, \forall ts', ts \in TS, ts' < ts \quad (43)$$

For units with two modes, Eq. (44) guarantees that the time difference between the next transition point and the current time point does not surpass the running time of the current mode, U_{im} . If no transition occurs at time point TP_{ts} , Eq. (44) simplifies to $TP_{ts} - TP_{ts'} \leq L$ for all $ts' < ts$.

$$TP_{ts} - TP_{ts'} \leq z_{imm',ts'} \times TU_{im'}^* + L \times (1 - z_{imm',ts'}) \quad \forall i \in MMU, (m, m') \in TR_i, \forall ts', ts \in TS, ts' < ts \quad (44)$$

For unit with three modes, Eq. (45) is similarly formulated to enforce that if a mode lasts from TP_{ts} to TP_{ts+1} , the duration must be no more than $TU_{im''}^*$.

$$TP_{ts+1} - TP_{ts} \leq (z_{im'm'',ts} + z_{im''m,ts+1} - 1) \times TU_{im''}^* + L \times (2 - z_{im'm'',ts} - z_{im''m,ts+1}) \quad \forall i \in I, (m, m', m'') \in SEQ_i, \forall ts \in TS, ts < |TS| - 1 \quad (45)$$

(Note that $|TS| = n$.)

3.2.3. Multi-mode Secondary unit constraints

For units with multiple operating modes, the binary variable $y_{s,ts}$ is set to 1 when the corresponding mode $m : (i, m, s) \in UMO$ is currently active. To simplify the introduction, it is assumed that the quality of a stream during a time slot remains consistent with the value determined at the planning stage. Eq. (46) enforces that only one mode can be active at a time. This leads to the formulation of Eq. (47a). Since variations in flow rate are permissible in the scheduling process, Eq. (47a) can be relaxed further into Eq. (47b), where η is a coefficient less than 1 (e.g., 0.85). Eq. (48) regulates that the steam quality during each time slot in the scheduling layer is consistent with that in the planning layer.

$$\sum_{m:(i,m,s) \in UMO} y_{s,ts} = 1 \quad \forall ts \in TS, i \in MMU \quad (46)$$

$$V_s^{ts} = \frac{V_s^* \times y_{s,ts}}{TU_{im}^*} \times (TP_{ts+1} - TP_{ts}) \quad \forall (i, m, s) \in UMO, i \in MMU \quad (47a)$$

$$V_s^{ts} \geq \eta \frac{V_s^* \times y_{s,ts}}{TU_{im}^*} \times (TP_{ts+1} - TP_{ts}) \quad \forall (i, m, s) \in UMO, i \in MMU \quad (47b)$$

$$Q_{s,q}^{ts} = Q_{s,q}^* \quad \forall (i, m, s) \in UMO, ts \in TS, q \in Q \quad (48)$$

3.2.4. Blender constraints

The binary variable b_s^{ts} indicates whether product s is blended under the ts^{th} time slot. Eq. (49) ensures that only one grade of

diesel product can be blended at any given time slot. Eqs. (50), (51), (54) and (55) are designed to guarantee that the blending rate stays within the allowable range and that the total blending time at the scheduling level is generally consistent with the planning results. Notably, β is a slack parameter introduced for computational convenience, permitting small deviation in the blending time from the leader problem. Other blending constraints follow the same principles as those in the planning model, but applied on a finer time scale. These include mass balance constraints (52) and (53), recipe constraint (54), blending quality constraints (57)–(60), and others.

$$\sum_{ts} b_s^{ts} \leq 1 \quad \forall ts \in TS, s : (b, s) \in BO \quad (49)$$

$$(1 - \beta) \cdot TB_s^* \leq \sum_{ts} b_s^{ts} \cdot (TP_{ts+1} - TP_{ts}) \leq (1 + \beta) \cdot TB_s^* \quad \forall ts \in TS, s : (b, s) \in BO \quad (50)$$

$$\frac{V_s^{ts}}{R_b^{\max}} \leq b_s^{ts} \cdot (TP_{ts+1} - TP_{ts}) \leq \frac{V_s^{ts}}{R_b^{\min}} \quad \forall ts \in TS, s : (b, s) \in BO \quad (51)$$

$$V_{s'}^{ts} = \sum_{s:(b,s) \in BI} V_{s,s'}^{ts} \quad \forall ts \in TS, s' : (b, s') \in BO \quad (52)$$

$$\sum_{s':(b,s') \in BI} V_{s,s'}^{ts} = V_s^{ts} \quad \forall ts \in TS, s : (b, s) \in BI, s' : (b, s') \in BO \quad (53)$$

$$R_b^{\min} \times (TP_{ts+1} - TP_{ts}) \leq \sum_{s:(b,s) \in BO} V_s^{ts} \quad \forall ts \in TS \quad (54)$$

$$\sum_{s:(b,s) \in BO} V_s^{ts} \leq R_b^{\max} \times (TP_{ts+1} - TP_{ts}) \quad \forall ts \in TS, b \in B \quad (55)$$

$$R_{s,s'}^{\min} \times V_{s'}^{ts} \leq V_{s,s'}^{ts} \leq R_{s,s'}^{\max} \times V_{s'}^{ts} \quad \forall ts \in TS, s : (b, s) \in BI, s' : (b, s') \in BO \quad (56)$$

$$QV_{s,s',q}^{ts} = V_{s,s'}^{ts} \times Q_{s,q}^{ts} \quad \forall ts \in TS, s : (b, s) \in BI, s' : (b, s') \in BO, q \in \{q | q \in VQ \wedge q : (s, q) \in SQB\} \quad (57)$$

$$QM_{s,s',q}^{ts} = QV_{s,s',sg'}^{ts} \times Q_{s,q}^{ts} \quad \forall ts \in TS, s : (b, s) \in BI, s' : (b, s') \in BI, q \in \{q | q \in WQ \wedge q : (s, q) \in SQB\} \quad (58)$$

$$Q_{s',q}^{\min} \times V_{s'}^{ts} \leq \sum_{s:(b,s) \in BI} QV_{s,s',q}^{ts} \leq Q_{s',q}^{\max} \times V_{s'}^{ts} \quad (59)$$

$$\forall ts \in TS, s' : (b, s') \in BO, q \in \{q | q \in WQ \wedge q : (s', q) \in SQB\}$$

$$Q_{s',q}^{\min} \times \sum_{s:(b,s) \in BI} QV_{s,s',sg'}^{ts} \leq \sum_{s:(b,s) \in BI} QM_{s,s',q}^{ts} \leq Q_{s',q}^{\max} \times \sum_{s:(b,s) \in BI} QV_{s,s',sg'}^{ts} \quad \forall ts \in TS, s' : (b, s') \in BO, q \in \{q | q \in VQ \wedge q : (s', q) \in SQB\} \quad (60)$$

3.2.5. Storage tank constraints

The constraints are similar to those in the planning model, assuming ideal mixing conditions and perfect mixing within the tank. Eqs. (61) and (62) impose limitations on the flow rate and tank level during each time slot. Eq. (63) ensures that the tank level at the start of

the first time slot matches the initial tank level from the planning layer. Mass and quality conservation are enforced by Eqs. (64)–(66).

$$R_t^{\min} \times (Tp_{ts+1} - Tp_{ts}) \leq \sum_{s:(t,s) \in TO} V_s^{ts} \leq R_t^{\max} \times (Tp_{ts+1} - Tp_{ts}) \quad \forall ts \in TS, t \in T \quad (61)$$

$$V_t^{\min} \leq V_t^{ts} \leq V_t^{\max} \quad \forall ts \in TS, t \in T \quad (62)$$

$$V_t^1 = V_t^i \quad \forall ts = 1, t \in T \quad (63)$$

$$\sum_{s:(t,s) \in T_{diesel}} V_s^{ts} + V_t^{ts} = V_t^{ts+1} + \sum_{s':(t,s') \in TO_{diesel}} V_{s'}^{ts} \quad \forall ts \in TS, t \in T \quad (64)$$

$$\sum_{s:(t,s) \in T} V_s^{ts} \cdot Q_{s,q}^{ts} + V_t^{ts} \cdot Q_{t,q}^{ts} = V_t^{ts+1} \cdot Q_{t,q}^{ts+1} + \sum_{s':(t,s') \in TO_{diesel}} V_{s'}^{ts} \cdot Q_{s',q}^{ts} \quad \forall ts \in TS, t \in T, q \in \{q | q \in VQ \wedge q : (s', q) \in SQB\} \quad (65)$$

$$\sum_{s:(t,s) \in T_{diesel}} V_s^{ts} \cdot Q_{s,q}^{ts} \cdot Q_{s',sg'}^{ts} + V_t^{ts} \cdot Q_{t,q}^{ts} \cdot Q_{t',sg'}^{ts} = V_t^{ts+1} \cdot Q_{t,q}^{ts+1} \cdot Q_{t',sg'}^{ts+1} + \sum_{s':(t,s') \in TO_{diesel}} V_{s'}^{ts} \cdot Q_{s',q}^{ts} \cdot Q_{s',sg'}^{ts} \quad \forall ts \in TS, t \in T, q \in \{q | q \in WQ \wedge q : (s', q) \in SQB\} \quad (66)$$

Therefore, the follower's problem can be formulated as such

$$(\text{Follower's problem}) \min \text{lossProfit} = \sum_{s \in DS} PS_s \times \text{loss}_s$$

s.t. Eqs.(40) – (66)

The entire bilevel optimization problem is then:

$$(\text{Leaser's problem}) \max \text{Profit} = \text{Revenue} - \text{Cost}$$

s.t. Eqs.(2) – (38)

$$(\text{Follower's problem}) \min \text{lossProfit} = \sum_{s \in DS} PS_s \times \text{loss}_s$$

s.t. Eqs.(40) – (66),

where the optimal solution of the follower's problem serves as the constraint for the leader's problem, thus forming the bilevel structure.

Taking uncertainty into account in the scheduling layer, it will be reformulated as a robust optimization model, which will be discussed in detail in Section 4. Furthermore, due to the computational challenges associated with the bilevel structure, the model is reformulated into a single-level optimization problem in Section 5.1 and decomposed into two subproblems in Section 5.2, both of which will be addressed later.

4. Data-driven Robust Optimization Model for Integrated Problem

4.1. Blending effect model

In the deterministic model, a linear blending model is employed (as seen in Eqs. (13)–(16) and Eqs. (57)–(60)). However, there is an additional component that contributes nonlinearly to the quality of

the blended product. To capture this, the blending effect model [35] is introduced. This model quantifies the blending effects of different components based on blending data and adjusts the component properties accordingly. An uncertain parameter, ub , is introduced to represent the blending effect value. As a result, the blending quality constraints can now be formulated as follows:

$$Q_{s',q}^{ts} = \frac{\sum_{s:(b,s) \in BI} V_{s,s'}^{ts} \times (Q_{s,q}^{ts} + ub_{s,q})}{V_{s'}^{ts}} = \sum_{s:(b,s) \in BI} r_{s,s'}^{ts} \times (Q_{s,q}^{ts} + ub_{s,q}) \quad \forall ts \in TS, s' : (b, s') \in BO, q \in \{q | q \in WQ \wedge q : (s', q) \in SQB\} \quad (67)$$

where $r_{s,s'}^{ts} = \frac{\sum_{s:(b,s) \in BI} V_{s,s'}^{ts}}{V_{s'}^{ts}}$ stands for the recipe of component s in the product s' in the ts^{th} time slot. It can be observed that this model is equivalent to a linear blending model plus a biased term.

This work focuses on the nonlinear blending effect of the Research Octane Number (RON) in the gasoline blending process. RON is a key indicator of a fuel's resistance to compression, and it directly influences the overall quality of the gasoline.

4.2. Multi-mode uncertainty set induced by weighted one-class supporting vector machine

In the field of uncertainty modeling, one-class support vector machines (OC-SVM) represent an important data-driven approach for constructing uncertainty sets. The core idea is to utilize historical data points to form a geometric set—typically a hypersphere or a polyhedron—referred to as the uncertainty set, which maximally encloses these data points. By incorporating the uncertainty set into a deterministic optimization model, a data-driven robust optimization formulation can be obtained. When a commercial solver is used to solve this model, it ensures that all constraints are satisfied for every point within the uncertainty set. In other words, the resulting solution remains feasible even under the worst-case realization of uncertain data.

Traditional OC-SVM approaches assume equal importance for all data points, determining their inclusion in the uncertainty set solely based on their distance to the boundary. However, in real-world industrial settings, data points often differ in importance. For instance, some data points may occur more frequently, align better with first-principles models, or be more representative. To address this, a weighting mechanism can be introduced so that higher-weighted data points are more likely to be included in the uncertainty set. This enhances the representativeness of the uncertainty set with respect to practical scenarios, thereby improving the practical relevance of the optimization results. In Section 4.2.1, the uncertainty set is further refined toward the goal point using the multi-stage clustering (MSC) algorithm proposed by Ghiasvand and Rahmani. By repeatedly running the weighted one-class support vector machine (WOC-SVM), this approach uncovers the clustering patterns within the uncertain data and improves the characterization of the uncertainty set.

In this section, an $N \times 3$ matrix is introduced, composed of three uncertainty vectors: $ub_{s',RON} = [ub_{\text{ref},RON}, ub_{\text{fcc},RON}, ub_{\text{c5},RON}]^T$.

Here, ref, fcc, and c5 represent reformate, fluid catalytic cracking (fcc) outlet, and C5, respectively—key components influencing blending nonlinearity. This matrix is derived from historical data using the recursive least squares (RLS) method.

Compared to conventional OC-SVM, a weight matrix $[\omega_{\text{ref}}, \omega_{\text{fcc}}, \omega_{\text{c5}}] N \times 3$ is introduced to account for the importance of each data

point. Ghiasvand and Rahmani combined the Weighted-OC-SVM presented by Bicego & Figueiredo [42] with the multi kernel-based OC-SVM (MKL-based OC-SVM) model developed by Han et al. [43], and the resulting MKL-based WOC-SVM model is formulated as follows:

$$\min_{\{W_m\}, \pi, p, \xi} \frac{1}{2} \sum_{m=1}^M \frac{1}{\pi_m} \|W_m\|^2 - p + \frac{1}{N \cdot \nu} \sum_{i=1}^N \omega_i \cdot \xi_i$$

$$\text{s.t. } \sum_{m=1}^M W_m^T \phi_m(ub_{i, \text{RON}}) \geq p - \xi_i, \forall i = 1, \dots, N$$

$$\xi_i \geq 0 \quad \forall i = 1, \dots, N$$

$$\sum_{m=1}^M \pi_m = 1, 0 \leq \pi_m \leq \frac{1}{M \cdot \mu}, m = 1, 2, 3$$

Where ϕ_m projects the input data samples into a higher dimension feature space. By evaluating the kernel function, the inner product of each mapping function can be calculated as $K_m(ub_{i, \text{RON}}, ub_{j, \text{RON}}) = \phi_m(ub_{i, \text{RON}}) \cdot \phi_m(ub_{j, \text{RON}})$. π_m is the coefficient of each kernel function, in this work, they are all set as 1/3. Also, ξ_i is a slack variable which turn some data points to outliers.

The above problem can be reformulated according to KKT conditions as follow, which is the dual problem of the original problem:

$$\min_{\alpha, \gamma, \zeta} -\gamma + \frac{1}{M \mu} \sum_{m=1}^M \zeta_m$$

$$\text{s.t. } -\frac{1}{2} \sum_{i,j=1}^N \alpha_i \alpha_j K_m(ub_{i, \text{RON}}, ub_{j, \text{RON}}) \geq \gamma - \zeta_m, \forall m = 1, 2, 3$$

$$\zeta_m \geq 0 \quad m = 1, 2, 3$$

$$\sum_{i=1}^N \alpha_i = 1, 0 \leq \alpha_i \leq \frac{\omega_i}{N \cdot \nu}, \forall i = 1, \dots, N$$

By solving this problem, we can determine whether or not the i th data sample is placed precisely on the boundary of the uncertainty set, which is known as Boundary Support Vector (BSVs) [36]. For those $\alpha_i = 0$, the corresponding data sample lies inside the uncertainty set and $\alpha_i = \frac{\omega_i}{N \cdot \nu}$ signifies the outliers.

$$SV = \{i | \alpha_i > 0, \forall i = 1, \dots, N\}$$

$$BSV = \{i | 0 < \alpha_i < \frac{\omega_i}{N \cdot \nu}, \forall i = 1, \dots, N\}$$

A new piecewise linear concave kernel function is introduced by Han et al. [43]

$$K_m(ub_{i, \text{RON}}, ub_{j, \text{RON}}) = 1 - \left| \frac{qb_m^T(ub_{i, \text{RON}} - ub_{j, \text{RON}})}{c_m \cdot k} \right|, \forall i, j = 1, \dots, N, \text{ and } \forall m = 1, 2, 3$$

where qb_m are the eigenvectors of the covariance matrix of the dataset and are extracted as direction vectors. And the parameters c_m, k are calculated as follow:

$c_m = \max_i^e \{qb_m^T(ub_{i, \text{RON}})\} - \max_i^{1-e} \{qb_m^T(ub_{i, \text{RON}})\}, \forall m = 1, 2, 3$, where $\max_i^e \{\cdot\}$ stands for the e -upper quantile of the set $\{\cdot\}$. Kernel function is positive-definite if the parameter k is large enough, therefore the parameter k as such:

$$k = \max_{m=1,2,3} \frac{S_m}{C_m}, S_m = \max_{ub_{i, \text{RON}}, ub_{j, \text{RON}} \in ub_{s, \text{RON}}} |qb_m^T(ub_{i, \text{RON}} - ub_{j, \text{RON}})|, \forall m = 1, 2, 3$$

Thereby, the uncertainty set is as follow:

$$U(ub_{s, \text{RON}}) = \left\{ u \left| \sum_{i \in SV} \alpha_i \sum_{m \in SK} \pi_m \left| \frac{qb_m^T(ub_{i, \text{RON}} - ub_{j, \text{RON}})}{c_m \cdot k} \right| \leq 1 - p \right. \right\} \quad (68a)$$

Where p is defined as

$$p = \frac{1}{|BSV|} \sum_{j \in BSV} \sum_{i \in SV} \alpha_i \sum_{m \in SK} \pi_m \cdot K_m(ub_{i, \text{RON}}, ub_{j, \text{RON}})$$

4.2.1. Setting the weight parameter and trimming uncertainty set

This section introduces an iterative pruning strategy based on the MSC algorithm [36]. In each iteration, an uncertainty set is constructed using the WOC-SVM method. Samples outside the current set (outliers) are removed, and the retained internal points are used to retrain the WOC-SVM, refining the uncertainty set. This process repeats until the number of retained samples meets a predefined threshold. Through this mechanism, the uncertainty set gradually contracts toward a specified goal point, guided by the discriminative boundary of WOC-SVM. Pruning distant outliers reduces the complexity of the subsequent optimization model, balancing solution accuracy with computational efficiency.

In the WOC-SVM, the weight parameter ω_i for each sample i must be specified. This parameter determines the penalty imposed when a sample is classified as an outlier. A larger weight encourages the model to retain the sample within the uncertainty set, while a smaller weight increases the likelihood of the sample being excluded. In this work, we define weights based on the Euclidean distance between each sample and a specified target point G : samples closer to the target point are considered more valuable for retention and thus assigned higher weights. Specifically, the inverse distance is defined as: $d_i^{\text{inv}} = \frac{1}{\|G - ub_{i, \text{RON}}\|^2 + e}$, where e is a minimum value like 0.01. In Ghiasvand & Rahmani's work, the weight is calculated by $\omega_i = \frac{\bar{d}_i \cdot N}{\sum_{i=1}^N \bar{d}_i}$, where $\bar{d}_i = (\lambda^* (d_i^{\text{inv}} - d_{\min}^{\text{inv}})) + (1 - \hat{b})$, $\lambda = \frac{2\hat{b}}{d_{\max}^{\text{inv}} - d_{\min}^{\text{inv}}}$. Here, $d_{\max}^{\text{inv}}$ and $d_{\min}^{\text{inv}}$ denotes the maximum and minimum values of inverse distance among all samples, and $\hat{b} \in (0, 1)$ is a tunable parameter for controlling the degree of proximity of the uncertainty set.

Fig. 2 shows the scatter of historical blending data. These points are widely distributed in property space. Traditional models often average stream properties over long periods, overlooking mode switches in units like hydrocrackers (e.g., between GM and KM). As a result, uncertainty sets may contain irrelevant samples, leading to mismatches between planning and scheduling. The proposed method enables mode-specific uncertainty sets. For instance, in GM mode, a goal point can be selected based on expert judgment, and samples farther from it are pruned by WOC-SVM due to lower weights—enhancing relevance and robustness.

In the case study, the MSC algorithm processes 300 historical samples. Goal points are selected via expert knowledge. In each iteration, outliers identified by WOC-SVM are removed, and the

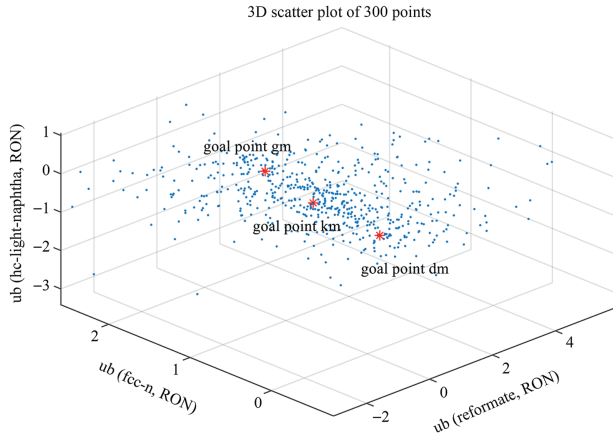


Fig. 2. Scatter plot of historical points of blending effect.

model is retrained until the desired coverage level $1-\lambda$ is achieved. The pseudocode of the MSC algorithm is provided below.

Algorithm 1. Pseudocode of multi-stage clustering [36].

Input: Dataset D , Goal point G , Regularization parameter ν , Inliers ratio $(1-\lambda)$, Tolerance ε

1. Set N equal to the number of samples in D
2. $C = (1-\lambda) \times N$
3. While $N > C + \varepsilon$ do:
4. Compute the covariance matrix of dataset D ; set principal eigenvector as q_m
5. Compute the new piecewise quadratic kernel function K_m (refer to Section 4.2)
6. Set $\pi_m = \frac{1}{M}$
7. Compute the average kernel matrix: $K = \sum_m \pi_m \times k_m$
8. Compute the weight ω_i for each data point
9. Solve the WOC-SVM problem using ν , K and ω_i ; and obtain dual variables α_i
10. Remove outliers, retain inliers to form updated dataset D
11. Update $N \leftarrow$ number of remaining data points in D
12. End while
13. Update ν ; $\nu \leftarrow 0.01$
14. Recompute covariance matrix of D , and set eigenvector q_m
15. Recompute K_m
16. Set $\pi_m = \frac{1}{M}$
17. Average kernel matrix $K = \sum_m \pi_m \times k_m$
18. Set uniform weights ω_i ; $\omega_i = 1, \forall i = 1, \dots, N$
19. Solve single kernel WOC-SVM with ν , K and ω_i ; and obtain final α_i

4.2.2. Robust counterpart reformulation

The uncertainty set defined above involves the L1 norm (i.e., absolute values), which is not computationally efficient for commercial solvers and is also inconvenient for direct solver input. Therefore, this section reformulates the uncertainty set and associated constraints into a robust counterpart to facilitate efficient computation.

By introducing dummy variables $\Theta = \{\theta_{im} | \forall i \in SV, \forall m \in SK\}$, the L1 norm in Eq. (68a) are removed:

$$\bar{U}_{\mu, \nu}(ub_{s, RON}) = \left\{ (ub, \Theta) \left| \begin{array}{l} -\theta_{im} \leq \frac{q_m^T(ub - ub_i)}{c_m \cdot k} \leq \theta_{im}, \forall i \in SV, m \in SK \\ \sum_{i \in SV} \alpha_i \sum_{m \in SK} \pi_m \cdot \theta_{im} \leq 1 - p \end{array} \right. \right\} \quad (68b)$$

For the quality of RON, the main concentration is focused to meet the lower requirement. Consequently we have

$$\min_{ub} \frac{\sum_{s: (b, s) \in BI} V_{s, s'}^{ts} \times (Q_{s, q}^{ts} + ub_{s, q})}{V_{s'}^{ts}} = \min_{ub} \sum_{s: (b, s) \in BI} r_{s, s'}^{ts} \times (Q_{s, RON}^{ts} + ub_{s, q}) \geq Q_{s', RON}^{\min} \quad (69)$$

By introducing Lagrangian multipliers for Eq. (68b), the problem can be formulated as below:

$$\begin{aligned} & \sum_{i \in SV} \sum_{m \in SK} (\mu_{im} - \lambda_{im}) \times \frac{qb_m^T \cdot ub_i}{c_m \cdot k} + \eta \times (p - 1) \geq Q_{s', RON}^{\min} - \\ & \sum_{s: (b, s) \in BI} r_{s, s'}^{ts} \times Q_{s, RON}^{ts} \\ & \sum_{i \in SV} \sum_{m \in SK} (\mu_{im} - \lambda_{im}) \times \frac{qb_m}{c_m \cdot k} = r_{s, s'}^{ts} \\ & \mu_{im} + \lambda_{im} = \eta \times \alpha_i \times \pi_m \quad \forall i \in SV, m \in SK \\ & \mu_{im}, \lambda_{im} \geq 0 \quad \forall i \in SV, m \in SK \\ & \eta \geq 0 \end{aligned} \quad (70a)$$

Further, since the hydrocracker has three modes: GM, KM, DM mode, three different uncertainty set targeting different goal points are constructed accordingly. Therefore the dual variables are induced accordingly (e.g., λ_{im}^{GM} , λ_{im}^{KM} , λ_{im}^{DM}). The counterpart reformulation is finally derived:

$$\begin{aligned} & \sum_{\{GM, KM, DM\}} \sum_{i \in SV} \sum_{m \in SK} (\mu_{im} - \lambda_{im}) \times \frac{qb_m^T \cdot ub_i}{c_m \cdot k} + \eta \times (p - 1) \geq \\ & Q_{s', RON}^{\min} - \sum_{s: (b, s) \in BI} r_{s, s'}^{ts} \times Q_{s, RON}^{ts} \\ & \sum_{\{GM, KM, DM\}} \sum_{i \in SV} \sum_{m \in SK} (\mu_{im} - \lambda_{im}) \times \frac{qb_m}{c_m \cdot k} = r_{s, s'}^{ts} \\ & \mu_{im} + \lambda_{im} = \eta \times \alpha_i \times \pi_m \quad \forall i \in SV, m \in SK \\ & \mu_{im}, \lambda_{im} \geq 0 \quad \forall i \in SV, m \in SK \\ & \eta \geq 0 \end{aligned} \quad (70b)$$

The different operating modes of a hydrocracker (gasoline mode, kerosene mode, diesel mode) influence the yield and composition of light and heavy naphtha, thereby affecting blending behavior. Light naphtha is primarily used for direct blending, while heavy naphtha typically requires reforming to enhance its octane rating, leading to variations in linear blending deviations (positive or negative). Additionally, hydrocracker products impact the blending characteristics of FCC gasoline and reformate, further influencing the final fuel quality and yield.

As shown in Fig. 2, the GM mode is assumed to have the most positive effect on all three blending components in terms of RON contribution, while the KM mode is expected to have the most negative effect. This assumption is reflected in the predefined goal points for this study.

Therefore we have the robust scheduling model:

$$(\text{Follower's problem}) \min \text{lossProfit} = \sum_{s \in DS} PS_s \times \text{loss}_s$$

s.t. Eqs.(40) – (66)

5. Integrated Framework & Computational Strategy

5.1. Single-level reformulation

It can be observed that the difference between the objective functions of the two problems represents the actual or “net profit” of the refinery. In this study, our primary focus is on the negative impact of nonlinearity in gasoline blending, which leads to a reduction in overall production. Additionally, since V_s^* was a lower-level parameter that is passed from the upper level in the bilevel approach, it is now a variable when considered in a single-level perspective, therefore we have:

$$\begin{aligned} \text{NetProfit} &= \text{Profit} - \text{lossProfit} = \\ &= -\text{totalCost} + \sum_{s \in \text{DS}} \text{PS}_s \times V_s - \sum_{s \in \text{DS}} \text{PS}_s \times \text{loss}_s \\ &= -\text{totalCost} + \sum_{s \in \text{DS}} \text{PS}_s \times \left(V_s - \left| V_s^* - \sum_{ts \in \text{TS}} V_s^{ts} \right| \right) = \\ &= -\text{totalCost} + \sum_{s \in \text{DS}} \text{PS}_s \times \left(V_s - V_s^* + \sum_{ts \in \text{TS}} V_s^{ts} \right) \end{aligned} \quad (71)$$

Therefore, the previous bilevel problem can be rewritten as a single-level problem:

$$\min \text{NetProfit} = -\text{totalCost} + \sum_{s \in \text{DS}} \text{PS}_s \times \sum_{ts \in \text{TS}} V_s^{ts}$$

s.t. Eqs.(2) – (38), (40) – (66), (70b)

However, as the scale of the single-level integrated model increases, the computational burden becomes significant. Therefore, it is necessary to employ a decomposition approach to break down the holistic problem into smaller, more manageable subproblems.

5.2. Calculation scheme: robust buffer-based decomposition

Due to the aforementioned reasons, the planning and scheduling models are now solved separately. In our proposed computational framework, the planning problem is solved first, followed by the scheduling problem in a sequential manner. However, since the planning model treats the blending process as linear, its solution may lead to infeasible parameters for the scheduling model, where nonlinear blending effects are explicitly considered. This discrepancy can result in infeasibility when the scheduling model attempts to implement the planned solution.

To address this issue without significantly increasing computational complexity, we introduce a robust buffer variable, ΔRON_s , in the planning model's quality constraints. The rationale is that a slight increase in the RON specification at the upper planning level can provide a safeguard against the adverse effects of uncertainty at the lower scheduling level. This robust buffer serves as a positive margin to accommodate nonlinear blending behaviors, thereby ensuring that when the planned inlet streams are passed to the scheduling stage, sufficient blending flexibility remains for feasible gasoline production. It also ensures the robustness of the upper level's plan. Accordingly, we define (72):

$$\begin{aligned} \sum_{s: (b, s') \in \text{BI}} QV_{s, s'}^{ts, \text{RON}} &\geq (Q_{s, \text{RON}}^{\min} + \Delta \text{RON}_s) \times V_s^{ts} \\ \forall ts \in \text{TS}, s' : (b, s') &\in \text{BO} \end{aligned} \quad (72)$$

After solving the planning model with Eq. (72), results of quality, stream flowrate, operating times are transferred as the parameter of scheduling model, then the NetProfit is calculated if the scheduling problem is feasible.

$$\text{NetProfit}^{K+1} = \text{Profit}^K - \text{lossProfit}^K$$

If the NetProfit does not change significantly compared to the previous iteration, we assume that at least a local maximum has been reached, and the iteration process is stopped to save computational time. However, if the scheduling model remains feasible under a given planning result, it suggests that the robust buffer ΔRON_s may not be large enough. In this case, ΔRON_s is increased, and a new iteration begins.

It can be observed that once ΔRON_s is determined, the planning model's result is fixed, which in turn determines the scheduling outcome. This implies that the objective function can be expressed as $\text{NetProfit}(\Delta \text{RON}_s)$. Inspired by this observation, a heuristic approach resembling gradient-based search methods is proposed to iteratively adjust ΔRON_s and obtain a sufficiently good solution.

In this work, it is assumed that as the solution approaches the global optimum, the lossProfit or deviation variable loss_s should approach zero. Therefore, the ratio $\frac{\text{loss}_s}{\sum_s \text{loss}_s}$ is adopted as an indicator for step size adjustment. The lower bound LB^K stores the current best-found objective NetProfit and is updated whenever a new, higher objective value is found (Fig. 3). The difference $\text{NetProfit}^K - \text{LB}^K$ serves as a pseudo-gradient. If $\text{NetProfit}^K \geq \text{LB}^K$, the increase in objective suggests that the previous step direction was beneficial, so the step size is increased by $\alpha(\text{NetProfit}^K -$

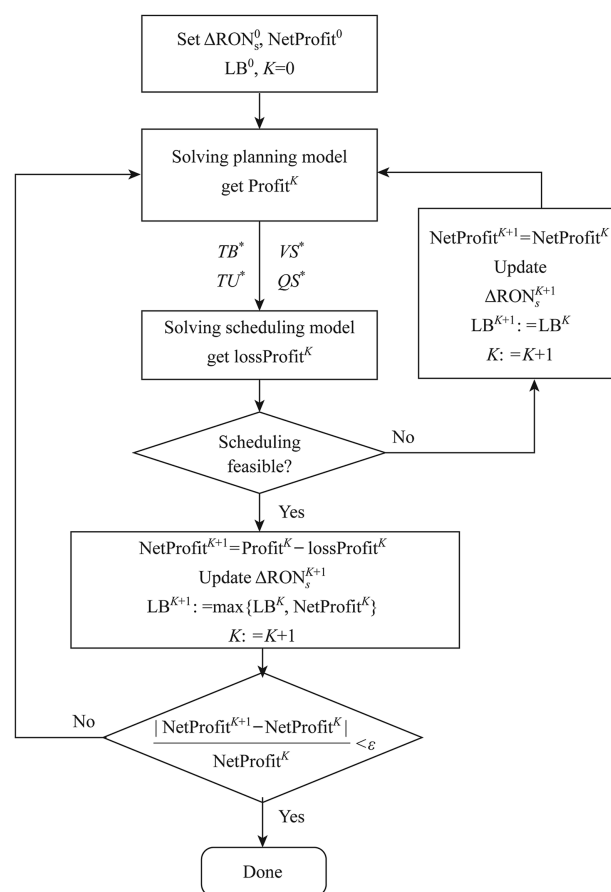


Fig. 3. Proposed bi-level decomposed framework: RBD.

$LB^K) \frac{\text{loss}_s}{\sum_s \text{loss}_s^2}$, where a greater loss_s indicates more room for optimization. Conversely, if $\text{NetProfit}^K \leq LB^K$, the search direction needs to be adjusted. In this case, a higher loss_s suggests a more constrained problem, so robust buffer variable ΔRON_s^K should not be decreased too much to avoid making scheduling infeasible. Instead, adjustments should be made more aggressively for products with smaller loss_s , as they have less impact on operability but may improve the objective function. To achieve this, the term $\alpha(\text{NetProfit}^K - LB^K) \exp\left(-\frac{\text{loss}_s}{\sum_s \text{loss}_s^2}\right)$ is introduced to ensure a negative correlation between the adjustment and loss_s when $\text{NetProfit}^K \leq LB^K$.

If $\text{loss}_s = 0$ in a given iteration, it suggests that ΔRON_s^K may have reached an optimal value. However, it is more likely that the robust buffer has been overestimated, which can unnecessarily reduce the objective function. Therefore, in this case, ΔRON_s^K is slightly reduced. The heuristic robust buffer adjustment strategy is formulated as follows:

$$\Delta \text{RON}_s^{K+1} = \begin{cases} \max\left(0, \Delta \text{RON}_s^K + \alpha(\text{NetProfit}^K - LB^K) \frac{\text{loss}_s}{\sum_s \text{loss}_s^2}\right), \\ (\text{loss}_s > 0) \text{ and } (\text{NetProfit}^K \geq LB^K) \\ \max\left(0, \Delta \text{RON}_s^K + \alpha(\text{NetProfit}^K - LB^K) \exp\left(-\frac{\text{loss}_s}{\sum_s \text{loss}_s^2}\right)\right), \\ (\text{loss}_s > 0) \text{ and } (\text{NetProfit}^K < LB^K) \\ \max(0, \Delta \text{RON}_s^K - \beta), \\ \text{loss}_s = 0 \end{cases}$$

If the scheduling model is infeasible, the infeasibility indicates ΔRON_s^K is not large enough to hedge against the lower-level uncertainty, and thus it is adjusted accordingly:

$$\Delta \text{RON}_s^{K+1} = \Delta \text{RON}_s^K + \kappa$$

The overall computational approach is illustrated in Fig. 3. In this way, by introducing a robust buffer variable at the upper level, the originally inseparable problem can be decomposed into two subproblems, making it easier to iteratively search for the optimum with reduced computational burden.

Remarks: The relatively small improvement of the objective value in each iteration is an intentional design choice rather than a limitation of the RBD scheme. In integrated planning and scheduling, the feasibility of the scheduling layer is highly sensitive to the targets generated at the planning level. Overly aggressive updates of the robust buffer may therefore induce feasibility oscillations between overly conservative and infeasible solutions. The proposed feedback-based update adopts incremental adjustments guided by the observed scheduling loss, which trades off convergence speed for stability and robustness.

6. Case Study

To better demonstrate the advantages of the proposed solution strategy, a base-case refinery planning and blending scheduling problem is solved using three representative approaches commonly adopted in the literature:

- a deterministic sequential planning–scheduling model.
- a robust model with a holistic uncertainty set.
- the proposed bilevel framework with multi-mode data-driven uncertainty sets.

All approaches are evaluated using the same data and performance metrics to ensure a fair comparison. The deterministic model corresponds to the conventional approach widely used in refinery planning studies, where uncertainty is neglected. The

Table 1
Results of computation of different model by single-level/bilevel decomposed approach.

Model	Inlier ratio (coverage)	Single-level (0.1%)		Bilevel (subproblem model 0.1%, iteration gap 1%)	
		Objective	Total time /s	Objective	Total time /s
Deterministic		1636980	436	1668466	343
Multi-mode	(0.3,0.3,0.3)	1381736	5899	1647271	1292
Holistic	0.9	1042260	675	1631188	571

Deterministic model is solved with single-level/sequential approach.

The one-level model exhibits unstable numerical behavior, often leading to infeasibility or suboptimality unless some hydrocracker mode durations are fixed. This instability is highly sensitive to small changes—e.g., a 0.001% adjustment in mode length may significantly alter the objective or cause infeasibility. Therefore, representative results are reported after several trials. In the deterministic model, the hydrocracker DM mode is fixed at 0.67607 of the total horizon; in the multi-set and holistic-set models, it is fixed at 0.67611.

Table 2
The model scale of single-level formulation.

Model	Objective decrease compared to bilevel method	Number of variables (Continuous/Binary)	Number of equations (linear/Nonconvex nonlinear)	CPU Time/s (gap = 0.01%)
Deterministic	31486	1247/604	2307/619	436
Multi-mode	265535	2357/604	2773/681	5899
Holistic	588928	1687/604	2443/1101	675

holistic uncertainty set–based robust model represents typical data-driven robust optimization methods reported in the literature, in which a single uncertainty set is used to hedge against all operating conditions.

To illustrate the capability of the proposed approach and to enable a fair base-case comparison with commonly used methods in the literature, the example is solved using both single-level and bilevel frameworks. Specifically, three types of models are considered: a deterministic model, a robust model with a holistic uncertainty set, and the proposed robust model with multi-mode uncertainty sets. All models are implemented in GAMS 33.1.0 and solved using ANTIGONE 1.1. The MILP solver and NLP solver used by ANTIGONE are CPLEX and CONOPT. The computing platform was Windows 11 Platform with i7-12700 2.30 GHz CPU and 32 GB memory. An optimality gap of 0.01% and a maximum solving time of 36000s have been set to be the termination criteria for the solver. For the bilevel framework, the stopping criterion is set to a 0.1% gap between iterations. This means that the iteration process continues until the relative change in the objective function between consecutive iterations is less than 0.1%.

Table 1 presents the results for different modes solved using either the one-level reformulation approach or the bilevel approach. It can be observed that the objective value generally decreases as the uncertainty set is introduced and expanded. In the deterministic model, where uncertainty is not considered, the objective value is the most optimistic. In contrast, the holistic model yields the lowest objective value, reflecting the largest and most conservative uncertainty set. The multi-mode model produces an intermediate objective value, as the mode-aware uncertainty sets are tighter than those in the holistic model. Therefore, the multi-mode model offers a balanced solution: less idealized than the deterministic model, but more realistic and accurate than the holistic model.

The deterministic bilevel calculation follows a sequential approach, where the planning model is solved first, and the unit length, stream rate, and quality parameters are transferred to the deterministic scheduling model before it is solved. Despite the absence of iterations in this deterministic calculation, the sequential solution provides the best possible value, indicating no deviation between the two layers (Table 1). On the other hand, when the deterministic model is solved using the single-level reformulation approach, the objective decreases by a certain amount (1636980 vs. 1668466) due to the increased computational burden. A similar decrease is observed for the robust model with multi-mode or holistic uncertainty sets (1381736 vs. 1647271 for multi-mode, 1042260 vs. 1631188 for holistic). It is also evident that the increase in the number of uncertainty sets in the multi-mode approach results in a surge in the number of continuous variables and, consequently, longer computational times (Table 2). However, when the bilevel model is solved using the bilevel approach, the subproblems become more tractable (as shown in Table 3), leading to higher objectives and reduced computational time (as illustrated in Fig. 4, convergence plot with stars indicating bilevel iteration results and lines representing single-level solutions.). Because the objective does not change significantly, all calculation in bilevel

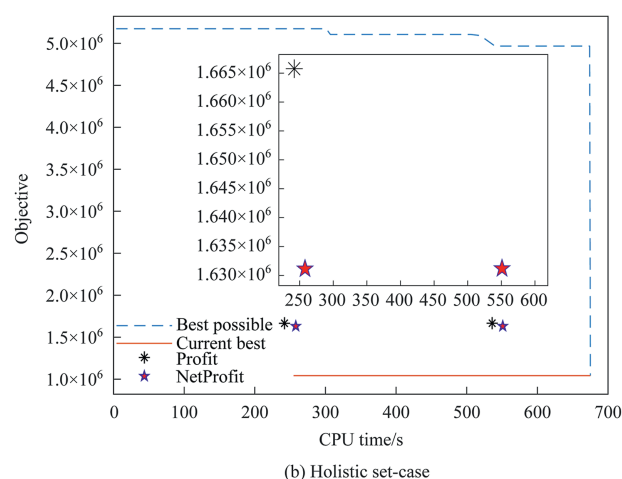
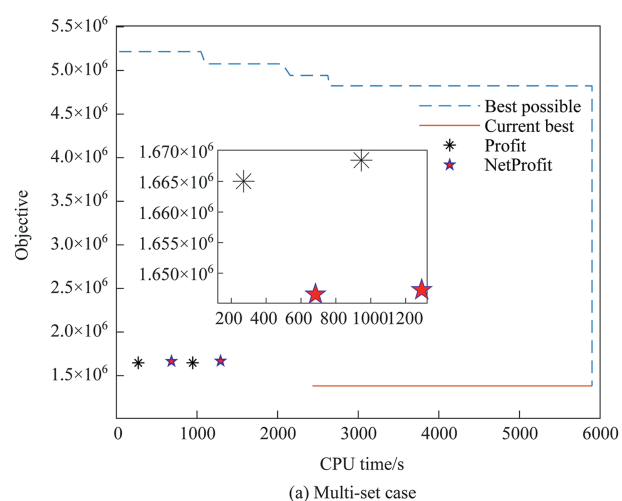


Fig. 4. Plot of converge curve (star denotes the value of Profit/NetProfit).

iterative approach ended in the second iteration. The detailed scheduling schemes and blending formula can be seen in Fig. 5.

From the results of the bilevel methods (Table 4), the gasoline product recipe (Regular Gasoline/Premium Gasoline) shows differences between the deterministic and robust models. Across all schemes, the solver prioritizes the production of premium gasoline, only meeting the minimum demand for regular gasoline. This indicates that the marginal benefits of premium gasoline are higher. However, in the robust model, the blending effect is influenced by negative deviations. The minimum demand of regular gasoline in the upper level in this case is 16400, which is not strictly met in lower level under uncertainty. In Table 4, the multi-mode uncertainty set leads to a decrease in regular gasoline production (16161.8 vs 16400), and the larger holistic uncertainty set causes even more significant reductions (16003.8 vs 16400), with a

Table 3

The model scale of planning model/scheduling model in bilevel framework.

Model		Number of variables (Continuous/Binary)	Number of equations (linear/Nonconvex nonlinear)
Planning model		836/476	1255/163
Scheduling	Deterministic	1074/491	2764/636
Model	Multi-mode	3262/491	3790/537
	Holistic	1924/491	3004/1556

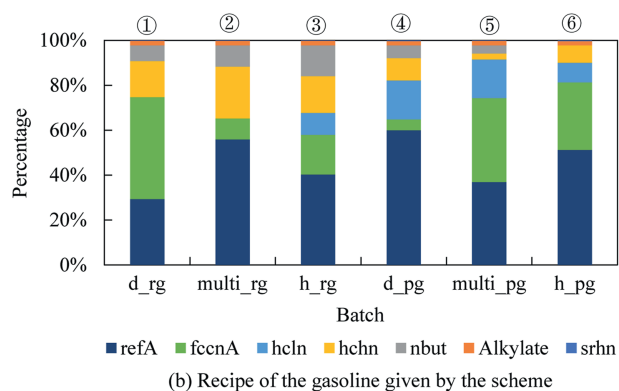
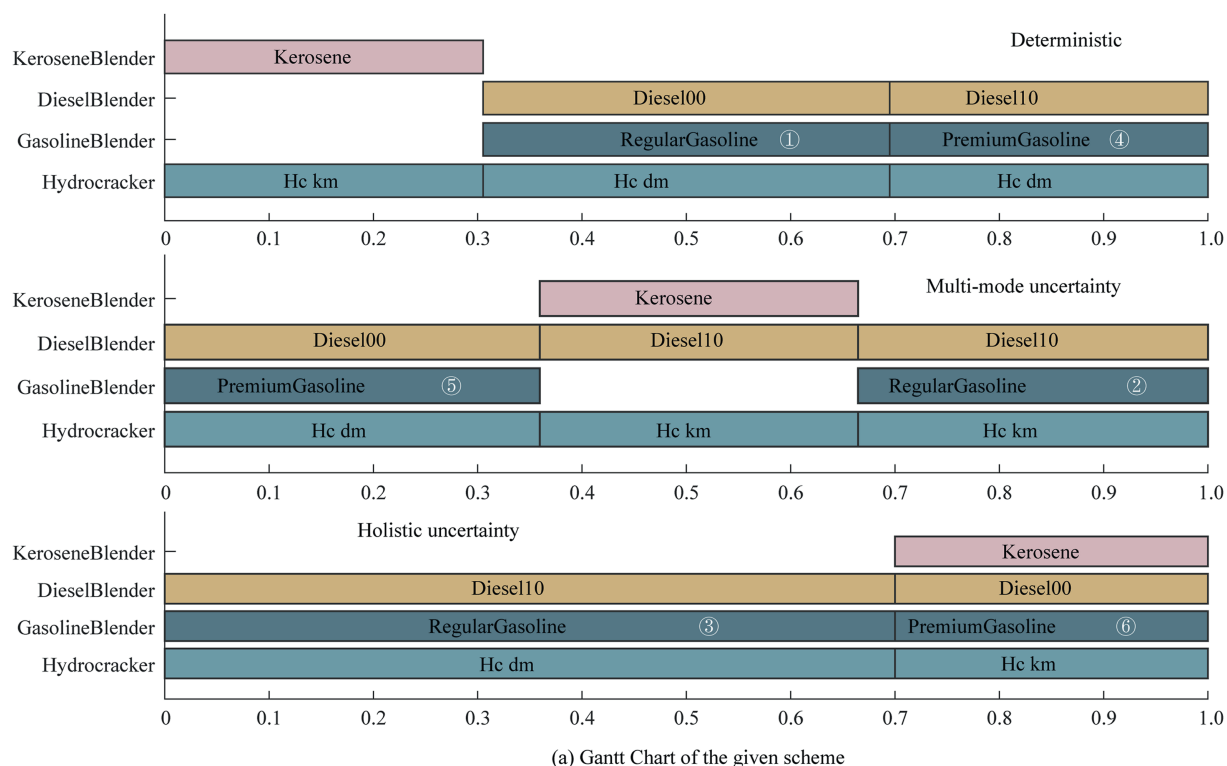


Fig. 5. Bilevel approach's results: d_rg: deterministic model, regular gasoline; multi_rg: multi-mode, regular gasoline; h_rg: holistic set, regular gasoline; d_pg: deterministic, premium gasoline; multi_pg: multi-mode, premium gasoline; h_pg: holistic set, premium gasoline.

slight decrease in premium gasoline as well (18178.7 vs 18197.7). This is because the scheduling now operates under stricter quality conditions. From the feasibility test (Fig. 6), it is evident that regular gasoline is more likely to fall short of the RON requirement. Accordingly, both the multi-mode and holistic models utilize more reformat (ref), which has the highest RON value, to increase the RON. Under uncertainty, reformat is prioritized over regular gasoline, while premium gasoline consumes more fluid catalytic

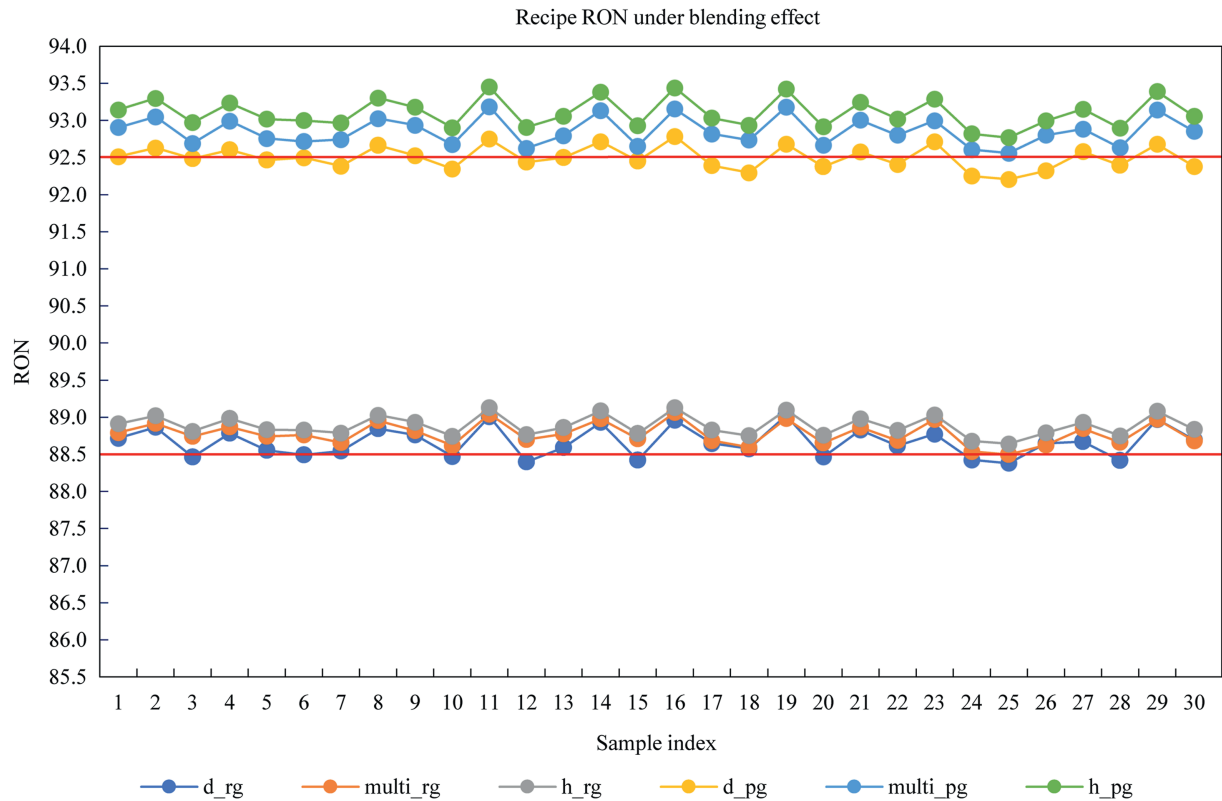
cracking naphtha (fccn), another high-RON component. However, it is important to note that the ratio of regular to premium gasoline differs between the multi-mode and holistic models. One possible explanation is that the goal point in Fig. 2 for $ub_{fccn, RON}$, RON exhibits a relatively large difference, which results in varying boundaries for the uncertainty sets across different modes. This variation in boundaries causes significant changes in the fccn recipe for both premium and regular gasoline under different

Table 4
Production amount of different product by bilevel decomposed method.

	Regular gasoline	Premium gasoline	Kerosene	Diesel100	Diesel110
Deterministic	16400	18197.7	1667.01	6000	9925.82
Multit	16161.8	18197.7	1667.01	6000	9925.82
Holistic	16003.8	18178.7	1667.01	6000	9925.82

Table 5
Recipe of gasoline product.

	RON	d_rg	multi_rg	h_rg	d_pg	multi_pg	h_pg
Amount	—	16400	16161.8	16003.8	18197.7	18197.7	18178.7
srhn	44.68	0.002466	0.002744	0.002596	0.002971	0.002757	0.002913
Alkylate	95	0.020000	0.020295	0.020495	0.020000	0.020000	0.020000
nbut	93.8	0.070054	0.093996	0.136486	0.057006	0.036659	0
hchn	53.59	0.160393	0.231026	0.163719	0.099143	0.025422	0.077102
hcln	82.12	0	0	0.097551	0.172522	0.172522	0.086822
fccnA	91.18	0.453372	0.093354	0.177079	0.048358	0.374032	0.301526
refA	102	0.293716	0.558586	0.402072	0.600000	0.368607	0.511636

**Fig. 6.** Feasibility test of RON quality by historical blending effect data.

modes. Specifically, due to the larger negative blending effect under the KM mode, regular gasoline uses little fccn. In comparison, in the holistic model, since the same uncertainty set is applied across all modes, the difference in fccnN usage between regular and premium gasoline is smaller. Furthermore, due to the small scale of the hcln dimension (as shown in $ub_{hc-light-naphtha,RON}$ Fig. 2) under the multi-mode set, the recipe for hcln is almost identical to that of the deterministic model (see Table 5).

Thirty points on the periphery of the uncertainty set (and at the intersection of different mode uncertainty sets for convenience of comparison, since the two grades of gasoline are blended under different hydrocracker modes) were selected to test the feasibility of meeting the minimum gasoline RON requirement. These intersection regions are chosen as they represent the most critical space where different operational modes must be harmonized to satisfy global blending constraints. Ensuring feasibility within these overlapping areas, rather than just within mode-specific subsets, provides a more rigorous validation of the model's robust “blending flexibility” across the entire uncertainty spectrum.

From Fig. 6, it can be observed that the blending recipe provided by the deterministic model has the poorest robustness (with only 73.33% feasible recipe for regular gasoline and 50% for premium gasoline, Table 6). In contrast, the holistic uncertainty set, with its largest coverage, demonstrates the highest robustness, reflected by the 100% feasible rate. Although the RON in the multi-mode approach is generally lower than in the holistic approach, it still satisfies the blending requirements within the given mode, indicating that the multi-mode methods hedge against the uncertainty while avoid over-conservatism. Relatively, the multi-mode method has a 0.986% increment in the objective NetProfit compared to holistic uncertainty method.

Table 6
Results of feasibility tests.

	Deterministic	Multi-mode	Holistic
Regular gasoline	73.33%	100%	100%
Premium gasoline	50%	100%	100%

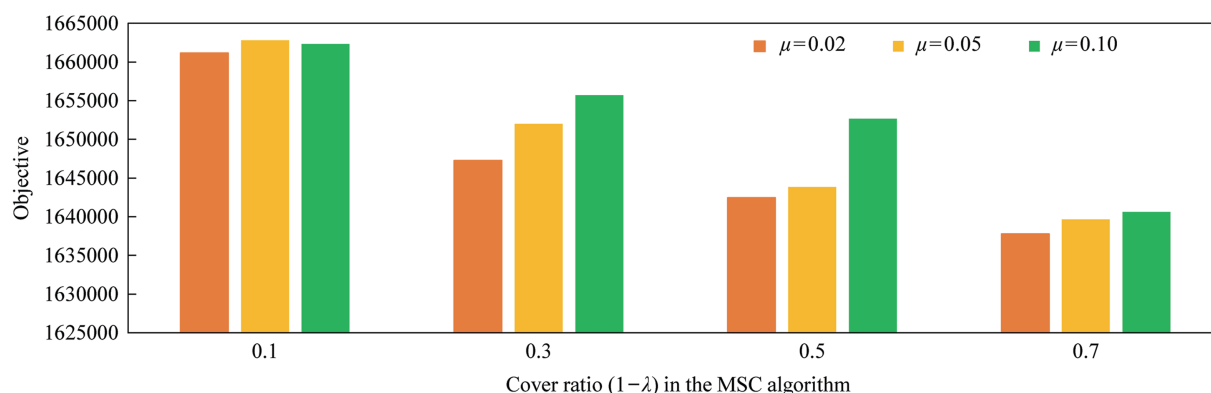


Fig. 7. Effect of parameters MSC on the result of objective.

Table 7
Sensitivity analysis of parameters in MSC algorithm.

parameters (1-λ) of 3 modes	(0.1,0.1,0.1)			(0.3,0.3,0.3)			(0.5,0.5,0.5)			(0.7,0.7,0.7)		
μ	0.02	0.05	0.1	0.02	0.05	0.1	0.02	0.05	0.1	0.02	0.05	0.1
Continuous/discrete variables	3262/491	3622/491	3622/491	3262/491	3622/491	3622/491	3642/491	3622/491	3622/491	398/491	3622/491	3622/491
Equations (linear/nonlinear)	3790/537	3790/537	3790/537	3790/537	3790/537	3790/537	3970/537	3970/537	3970/537	4150/537	3970/537	3970/537
CPU time/s	1739	1916	1757	1292	1474	1280	1490	2262	1400	1951	1396	1716
NetProfit	1661177	1662831	1662345	1647271	1651986	1655714	1642475	1643839	1652650	1637804	1639630	1640635
lossProfit	4744	1776	5140	21200	14983	11370	25991	23254	14699	30662	27438	27831

The robust calculations are performed using the uncertainty set constructed by the Multi-Stage Clustering algorithm (as discussed in Section 4.2) with varying parameters, including the coverage ratio and μ . The coverage ratio, which affects the data coverage of the uncertainty set after all iterations, is set to 0.1, 0.3, 0.5, and 0.7 for each mode, respectively. This parameter is crucial in determining the size of the uncertainty set. The parameter μ (as defined in Section 4.2), which influences the number of outlier data points selected during each iteration, is set to 0.02, 0.05, and 0.10.

The optimization results and model performance under different parameters are shown in Fig. 7 and Table 7, respectively. From Fig. 7, we observe that, for a fixed μ , the objective function (NetProfit) decreases as the coverage ratio increases. This aligns with the growth of the uncertainty set, as more conservatism is incorporated into the solution. On the other hand, when the coverage ratio is fixed, the objective function increases with μ for cover ratios of 0.3, 0.5, and 0.7. This is reasonable, as increasing μ results in the selection of more outlier data points during each iteration, leading to a smaller uncertainty set by the end of the MSC process. However, when the coverage ratio is set to 0.1, the objective function for $\mu = 0.10$ is slightly lower than for $\mu = 0.05$. This can be explained by the fact that when $\mu = 0.10$, more outliers are trimmed in each iteration. As a result, convergence occurs more quickly, but the trimmed dataset may deviate more significantly from the intended target point, potentially including more adverse data points. These remaining worst-case scenarios can negatively impact the objective value when $\mu = 0.10$. In contrast, when the coverage ratio is large enough (0.3, 0.5 or 0.7), this deviation has minimal influence on the final outcome. In practical use, the parameters can be tuned according to specific scenarios and demands.

To conclude, the results presented in this section demonstrate the effectiveness and flexibility of the proposed approach for

solving complex blending and scheduling problems under uncertainty. By comparing the proposed method with commonly adopted approaches in the literature, including deterministic sequential planning–scheduling models and robust models with holistic uncertainty sets, the advantages of the proposed framework are clearly illustrated. The comparison between single-level and bilevel frameworks highlights the significant improvements in both robustness and computational efficiency achieved through the bilevel formulation. Moreover, the proposed multi-mode uncertainty modeling strategy provides a balanced trade-off between conservatism and performance, achieving a higher NetProfit than the holistic robust model while still effectively hedging against uncertainty. Robustness is further validated through feasibility tests on minimum RON requirements, where the deterministic model exhibits the poorest robustness, whereas both robust models achieve full feasibility. Compared with the holistic uncertainty approach, the proposed multi-mode method avoids unnecessary conservatism by exploiting mode-dependent uncertainty structures. These results confirm that the proposed approach outperforms conventional methods in the literature in terms of robustness, economic performance, and computational tractability.

Overall, this case study confirms the potential of using multi-stage clustering-based uncertainty sets combined with bilevel optimization frameworks to solve real-world blending and scheduling problems in the presence of significant uncertainty, offering both improved robustness and computational tractability.

7. Conclusions

This study addressed the challenge of coordinating refinery planning and product blending scheduling under uncertainty by developing a bilevel, data-driven robust optimization framework

that explicitly accounts for multi-mode operating conditions. Unlike conventional deterministic and holistic robust formulations, the proposed approach exploits historical operational data to construct mode-dependent uncertainty sets, enabling more flexible and less conservative decision-making. The numerical results from the refinery case study confirm that incorporating multi-mode uncertainty modeling substantially improves solution robustness without sacrificing economic performance. In particular, the proposed framework consistently preserves feasibility of quality constraints under uncertain blending behavior, while achieving 1% higher net profit than holistic robust models. Moreover, the bilevel formulation combined with the decomposition-based solution strategy significantly enhances computational efficiency, making the approach suitable for practical, large-scale refinery applications.

In summary, the proposed framework provides a systematic and scalable methodology for integrating planning and scheduling decisions in uncertain refinery operations. Future research will focus on extending the approach to additional quality attributes and sources of uncertainty, improving data-driven uncertainty learning techniques, and applying the framework to larger industrial-scale systems with more complex operating modes, or multi-refinery networks.

CRedit Authorship Contribution Statement

Wenhao Yu: Writing – original draft, Software, Methodology, Formal analysis, Conceptualization. Xin Dai: Writing – review & editing, Supervision, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. Yuanhang Yue: Validation, Conceptualization. Xinwei Lin: Software. Minglei Yang: Supervision, Funding acquisition, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by National Key Research and Development Program of China (2023YFB3307801), National Natural Science Foundation of China (62293501), Shanghai Special Project for Promoting High-quality Industrial Development (250504), the Programme of Introducing Talents of Discipline to Universities (the 111 Project) under Grant B17017, Fundamental Research Funds for the Central Universities (222202517006) and Shanghai Pujiang Program (24PJJD025).

Nomenclature

Indices

m, m', m''	operating modes in the process model
m	index for the kernel in WOC-SVM model (different from the m in the process model)
i	index for the data points in WOC-SVM model (different from the i in the process model)

Sets and indices

B	$\{b\}$ blenders b
I	$\{i\}$ units i without blenders and tank s

Q	$\{q\}$ quality properties q
S	$\{s\}$ streams s (s' is an alias of index s)
T	$\{t\}$ tanks t
TS	$\{ts\}$ event points ts (ts' is an alias of index ts)

Subsets and indices

B_{diesel}	$\{b\}$ blender b involved in diesel blending
BI	$\{(b,s)\}$ inlet streams s of blender b
BO	$\{(b,s)\}$ outlet streams s of blender b
B_{diesel}	$\{(b,s)\}$ inlet streams s of blender b involved in diesel blending
B_{diesel}	$\{(b,s)\}$ outlet streams s of blender b involved in diesel blending
DS	$\{s\}$ demanded product stream s
M_i	$\{m\}$ modes m of multiple-mode units i
MIX	$\{m\}$ mixer
MMU	$\{i\}$ units i with multiple operating modes
SEQ_i	$\{(m,m',m'')\}$ predefined sequences of mode transitions in multi-mode units i
SPL	$\{m\}$ splitter
SQ	$\{(s,q)\}$ quality properties q of streams s
SQB	$\{(s,q)\}$ quality properties q of streams s included in blending calculations
SQB_{diesel}	$\{(s,q)\}$ quality properties q of streams s included in blending calculations for diesel blending
SS	$\{s\}$ supplied raw material stream s
T_{diesel}	$\{t\}$ tanks t involved in diesel blending
TI	$\{(t,s)\}$ inlet streams s of tanks t
TO	$\{(t,s)\}$ outlet streams s of tanks t
TR_i	$\{(m,m')\}$ modes transitions in multi-mode unit i
$\overline{TR}_{im}$	$\{m'\}$ modes of multi-mode unit i which can be directly reached from mode m
$\widehat{TR}_{im}$	$\{m'\}$ modes of multi-mode unit i from which mode m can be directly reached
UMO	$\{(i,m,s)\}$ outlet stream s from multi-mode unit i under mode m
VQ	$\{q\}$ quality properties blended linearly on a volumetric basis
WQ	$\{q\}$ quality properties blended linearly on a weight basis

Parameters

$A_s^{\max}$	the maximum availability of raw material stream s
$A_s^{\min}$	the minimum availability of raw material stream s
BSV	basic supporting vectors in the WOC-SVM model
b_s	if equal to 1, the product stream s is blended over the whole planning period
b_s^{ts}	if equal to 1, the product stream s is blended during ts^{th} time slot
$D_s^{\max}$	the maximum demand of product stream s
$D_s^{\min}$	the minimum demand of product stream s
L	length of the planning horizon
loss_s	production deviation between upper and lower levels
PS_s	unit price of the product stream s
$Q_{s,q}$	quality of a property q in stream s in planning
$Q_{s,q}^{\max}$	maximum quality specification of a property q in stream s
$Q_{s,q}^{\min}$	minimum quality specification of a property q in stream s
$Q_{s,q}^{ts}$	quality of a property q in stream s during ts^{th} time slot
$Q_{t,q}^i$	initial quality value of a property q in tank t
$R_b^{\max}$	maximum blending rate of a blender b
$R_b^{\min}$	minimum blending rate of a blender b

$R_{s,s'}^{\max}$	maximum recipe (volume fraction) allowed of blend component s in blending product s'
$R_{s,s'}^{\min}$	minimum recipe (volume fraction) allowed of blend component s in blending product s'
$R_t^{\max}$	maximum withdrawal rate of a tank t
$R_t^{\min}$	minimum withdrawal rate of a tank t
SV	supporting vectors in the WOC-SVM model
Tp_t	exact time of event point t in time horizon
TU_{im}	the running length of mode m of multi-mode unit i
TU_{im}^*	operating time for mode m as a parameter transferred from the upper level
V_s	volume flow of a stream s in planning
$V_{s,s'}$	volume flow from a blend component s to a blending product s' in planning
$V_s^{\min}$	minimum volume amount of a stream s
V_s^{ts}	volume flow of a stream s during ts^{th} time slot
$V_{s,s'}^{ts}$	volume flow from a blend component s to a blending product s' during ts^{th} time slot
V_t	inventory level in tank t in planning
V_t^i	initial inventory of a tank t
$V_t^{\max}$	maximum capacity of a tank t
$V_t^{\min}$	minimum inventory of a tank t
V_t^{ts}	inventory level in tank t during ts^{th} time slot
VS^*	production of stream s as a parameter transferred from the upper level
$ub_{s,q}$	uncertain parameter for blending effect model for quality q in product s
$y_{im,ts}$	1 if multi-mode unit i is in mode m during ts^{th} time slot
$z_{imm',ts}$	1 if multi-mode unit i switches from mode m to m' at the starting point of ts^{th} time slot (event point ts)
α_i	the dual variable for the i th datapoint in the WOC-SVM model
β, η	slacking parameters for lower level scheduling
$\mu_{im}, \lambda_{im}, \eta$	the Lagrangian dual in the reformulation of constraints involved uncertainty parameters
ω_i	the weight of the datapoint in the WOC-SVM model

References

- [1] B. Brunaud, I.E. Grossmann, Perspectives in multilevel decision-making in the process industry, *Front. Eng.* 4 (3) (2017) 256.
- [2] D. van de Berg, N. Shah, E.A. del Rio-Chanona, Hierarchical planning-scheduling-control: optimality surrogates and derivative-free optimization, *Comput. Chem. Eng.* 188 (2024) 108726.
- [3] J. Morlet-Espinosa, A. Flores-Tlacuahuac, A Bayesian optimization approach for data-driven mixed-integer nonlinear programming problems, *AIChE J.* 70 (9) (2024) e18448.
- [4] L.S. Dias, M.G. Ierapetritou, Integration of planning, scheduling and control problems using data-driven feasibility analysis and surrogate models, *Comput. Chem. Eng.* 134 (2020) 106714.
- [5] J.H. Lee, J. Shin, M.J. Realff, Machine learning: overview of the recent progresses and implications for the process systems engineering field, *Comput. Chem. Eng.* 114 (2018) 111–121.
- [6] I. Harjunkoski, T. Ikonen, H. Mostafaei, T. Deneke, K. Heljanko, Synergistic and intelligent process optimization: first results and open challenges, *Ind. Eng. Chem. Res.* 59 (38) (2020) 16684–16694.
- [7] L. Yang, J.Z. Ren, M. Eden, C.L. Chang, W.F. Shen, A penalty-free hybrid algorithm framework based on feasible stream matching principle for large-scale heat exchanger networks synthesis, *Chem. Eng. Sci.* 298 (2024) 120419.
- [8] S. Terrazas-Moreno, P.A. Trotter, I.E. Grossmann, Temporal and spatial Lagrangean decompositions in multi-site, multi-period production planning problems with sequence-dependent changeovers, *Comput. Chem. Eng.* 35 (12) (2011) 2913–2928.
- [9] P. Castillo Castillo, P.M. Castro, V. Mahalec, Global optimization algorithm for large-scale refinery planning models with bilinear terms, *Ind. Eng. Chem. Res.* 56 (2) (2017) 530–548.
- [10] J.A. Persson, M. Göthe-Lundgren, Shipment planning at oil refineries using column generation and valid inequalities, *Eur. J. Oper. Res.* 163 (3) (2005) 631–652.
- [11] A. Uribe-Rodríguez, P.M. Castro, G. Guillén-Gosálbez, B. Chachuat, Assessment of lagrangean decomposition for short-term planning of integrated refinery-petrochemical operations, *Comput. Chem. Eng.* 174 (2023) 108229.
- [12] R.E. Franzoi, B.C. Menezes, J.D. Kelly, J.A.W. Gut, I.E. Grossmann, Large-scale optimization of nonconvex MINLP refinery scheduling, *Comput. Chem. Eng.* 186 (2024) 108678.
- [13] H.K. Yang, D.E. Bernal, R.E. Franzoi, F.G. Engineer, K. Kwon, S. Lee, I.E. Grossmann, Integration of crude-oil scheduling and refinery planning by Lagrangean decomposition, *Comput. Chem. Eng.* 138 (2020) 106812.
- [14] L.J. Su, D.E. Bernal, I.E. Grossmann, L.X. Tang, Modeling for integrated refinery planning with crude-oil scheduling, *Chem. Eng. Res. Des.* 192 (2023) 141–157.
- [15] B. Colson, P. Marcotte, G. Savard, An overview of bilevel optimization, *Ann. Oper. Res.* 153 (1) (2007) 235–256.
- [16] Y.F. Chu, F.Q. You, J.M. Wassick, A. Agarwal, Integrated planning and scheduling under production uncertainties: Bi-level model formulation and hybrid solution method, *Comput. Chem. Eng.* 72 (2015) 255–272.
- [17] M.A. Muñoz, S. Pineda, J.M. Morales, A bilevel framework for decision-making under uncertainty with contextual information, *Omega* 108 (2022) 102575.
- [18] B. Beykal, S. Avraamidou, I.P.E. Pistikopoulos, M. Onel, E.N. Pistikopoulos, DOMINO: data-driven optimization of bi-level mixed-integer nonlinear problems, *J. Glob. Optim.* 78 (1) (2020) 1–36.
- [19] Z.K. Li, M.G. Ierapetritou, Integrated production planning and scheduling using a decomposition framework, *Chem. Eng. Sci.* 64 (16) (2009) 3585–3597.
- [20] Z.K. Li, M.G. Ierapetritou, Production planning and scheduling integration through augmented Lagrangian optimization, *Comput. Chem. Eng.* 34 (6) (2010) 996–1006.
- [21] F.P. Li, M.L. Yang, W.L. Du, X. Dai, Development and challenges of planning and scheduling for petroleum and petrochemical production, *Front. Eng. Manag.* 7 (3) (2020) 373–383.
- [22] P.M. Verderame, J.A. Elia, J. Li, C.A. Floudas, Planning and scheduling under uncertainty: a review across multiple sectors, *Ind. Eng. Chem. Res.* 49 (9) (2010) 3993–4017.
- [23] D.J. Yue, F.Q. You, Planning and scheduling of flexible process networks under uncertainty with stochastic inventory: MINLP models and algorithm, *AIChE J.* 59 (5) (2013) 1511–1532.
- [24] F.P. Li, F. Qian, W.L. Du, M.L. Yang, J. Long, V. Mahalec, Refinery production planning optimization under crude oil quality uncertainty, *Comput. Chem. Eng.* 151 (2021) 107361.
- [25] Y. Chen, M. Lin, H. Jiang, Z.H. Yuan, B.Z. Chen, Optimal design and operation of refinery hydrogen systems under multi-scale uncertainties, *Comput. Chem. Eng.* 138 (2020) 106822.
- [26] S. Amaran, T. Zhang, N.V. Sahinidis, B. Sharda, S.J. Bury, Medium-term maintenance turnaround planning under uncertainty for integrated chemical sites, *Comput. Chem. Eng.* 84 (2016) 422–433.
- [27] D.J. Yue, F.Q. You, Optimal supply chain design and operations under multi-scale uncertainties: nested stochastic robust optimization modeling framework and solution algorithm, *AIChE J.* 62 (9) (2016) 3041–3055.
- [28] P. Mohajerin Esfahani, D. Kuhn, Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations, *Math. Program.* 171 (1) (2018) 115–166.
- [29] E. Delage, Y.Y. Ye, Distributionally robust optimization under moment uncertainty with application to data-driven problems, *Oper. Res.* 58 (3) (2010) 595–612.
- [30] J. Long, J.W. Zhu, N. Wang, J.Z. Zhai, T.T. Xu, C. Liang, L. Zhao, Data-driven robust optimization of energy-material coupled system in refineries under uncertainty, *Expert Syst. Appl.* 267 (2025) 126184.
- [31] S. Razm, N. Brahimi, R. Hammami, A. Dolgui, Row and column-wise robust optimization model for biorefineries storing perishable biomass under weather uncertainty: boosted by machine learning, *Comput. Ind. Eng.* 200 (2025) 110823.
- [32] A. Baghban, P.M. Castro, F. Oliveira, Data-driven robust optimization for pipeline scheduling under flow rate uncertainty, *Comput. Chem. Eng.* 193 (2025) 108924.
- [33] W.L. He, J.M. Zhao, L. Zhao, Z. Li, M.L. Yang, T.B. Liu, Data-driven two-stage distributionally robust optimization for refinery planning under uncertainty, *Chem. Eng. Sci.* 269 (2023) 118466.
- [34] C. Wang, X. Peng, C. Shang, C. Fan, L. Zhao, W.M. Zhong, A deep learning-based robust optimization approach for refinery planning under uncertainty, *Comput. Chem. Eng.* 155 (2021) 107495.
- [35] X. Dai, X.Q. Wang, R.C. He, W.L. Du, W.M. Zhong, L. Zhao, F. Qian, Data-driven robust optimization for crude oil blending under uncertainty, *Comput. Chem. Eng.* 136 (2020) 106595.
- [36] M.R. Ghiasvand, D. Rahmani, A novel weighted data-driven robust optimization approach for creating adjustable uncertainty sets, *Comput. Chem. Eng.* 178 (2023) 108390.
- [37] C. Ning, F.Q. You, Data-driven adaptive nested robust optimization: general modeling framework and efficient computational algorithm for decision making under uncertainty, *AIChE J.* 63 (9) (2017) 3790–3817.
- [38] C. Ning, F.Q. You, Data-driven decision making under uncertainty integrating robust optimization with principal component analysis and kernel smoothing methods, *Comput. Chem. Eng.* 112 (2018) 190–210.
- [39] W.H. Yu, F.P. Li, X. Dai, W.L. Du, M.L. Yang, Continuous-time formulation for integrated refinery planning and diesel blending scheduling problems, *Ind. Eng. Chem. Res.* 64 (7) (2025) 3918–3936.

- [40] Y. Beck, I. Ljubić, M. Schmidt, A survey on bilevel optimization under uncertainty, *Eur. J. Oper. Res.* 311 (2) (2023) 401–426.
- [41] F.P. Li, F. Qian, M.L. Yang, W.L. Du, V. Mahalec, Product tri-section based crude distillation unit model for refinery production planning and refinery optimization, *AIChE J.* 67 (2) (2021) e17115.
- [42] M. Bicego, M.A.T. Figueiredo, Soft clustering using weighted one-class support vector machines, *Pattern Recognit* 42 (1) (2009) 27–32.
- [43] B. Han, C. Shang, D.X. Huang, Multiple kernel learning-aided robust optimization: learning algorithm, computational tractability, and usage in multi-stage decision-making, *Eur. J. Oper. Res.* 292 (3) (2021) 1004–1018.