Modeling and Optimization of Stochastic Scheduling for Refinery Processes

Lijie Su[a,b], Lixin Tang[a*], Ignacio E. Grossmann[c], Yiping Huang[d]

[a] National Frontiers Science Center for Industrial Intelligence and Systems Optimization, Northeastern University, Shenyang 110819, China

[b] Key Laboratory of Data Analytics and Optimization for Smart Industry (Northeastern University), Ministry of Education, China

[c] Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh 15213, PA, USA

[d] Liaoning Key Laboratory of Manufacturing System and Logistics Optimization, Shenyang, 110819, China

Abstract

This work aims at the integrated scheduling for whole refinery processes with consideration of uncertainties in demands and yields. We first define the stochastic scheduling problem for the refinery processes. The external and process uncertainties of refineries are represented through the extended scenario tree based on the history data. Using a continuous-time representation, a hybrid Mixed Integer Nonlinear Programming (MINLP) and Generalized Disjunctive Programming model is formulated for the stochastic scheduling of whole refinery processes. To solve the proposed large-scale stochastic mathematical model with scenarios, we develop an Outer Approximation method combined with the Fix-and-Relax strategy to efficiently solve the real scheduling instances. Computational results demonstrate the validity of the proposed stochastic model and the efficiency of the proposed solution method compared with the MINLP solver DICOPT.

Keywords: Refinery Processes; Stochastic Scheduling; Scenario Tree; Mixed Integer Nonlinear Programming (MINLP); Outer Approximation (OA)

1. Introduction

Refinery production processes separate a crude oil or a mixed crude oil into different intermediate products, which are then blended as components into final oil products in order to satisfy the market demand. Different processing schemes determine the types and amounts of final oil output and processing costs, which directly impact refinery production profits. There are several uncertain parameters within the entire production processes of a refinery, such as properties of crude oil,

* Corresponding author. *E-mail address*: lixintang@mail.neu.edu.cn (Lixin Tang)

processing conditions, volatility of component oils, or demand of final oil products. In addition, the supply of crude oil is affected by international political and economic factors, and the demand and prices of oil products vary also with seasons, and the economy. Refinery scheduling requires determining the processing sequence, amounts and properties of crude oil, intermediate flows, component products and final oil products, according to a production plan while optimizing the net production profits. Compared to refinery planning, scheduling is aimed at realizing the production plan in real process conditions with dynamic changes. A good schedule is able to guarantee accurate execution, production profit and flexible response to uncertainties. Therefore, scheduling of refinery processes with consideration of uncertainties is a problem of great practical significance.

Refinery processes are usually a continuous production network with storage tanks of crude oils, intermediate products and final oil products. Refinery scheduling involves determining the timetable of processed materials, inventory of each intermediate product, blending ratios, and output of final oil products over scheduling horizon. Refinery scheduling requires satisfying operation rules, production capacities of units, materials and properties balance, inventory capacity, supply of crude oil, demand of oil product, and energy consumption constraints. The target is to maximize the net profit of production, which is equal to the total product revenue minus cost of raw material and production processes.

Joly et al. (2002) studied the planning and scheduling problems in petroleum refinery, which is decomposed into three optimization problems, and proposed a mathematical model based on continuous and discrete time representations. Jia and Ierapetritou (2004) also decomposed the refinery scheduling for the whole processes into three parts, then formulated three scheduling models. Shah et al. (2009) proposed a centralized-decentralized optimization method aiming at the integrated scheduling for the whole refinery processes. Shah and Ierapetritou (2011) addressed refinery scheduling for a real refinery plant. They proposed a continuous-time formulation for the short-term refinery scheduling. Few references have focused on the whole refinery scheduling compared with many studies on batch scheduling in processes system engineering (Méndez et al., 2006). Using processing tasks and continuous-time representation for refinery scheduling was shown to be a general feasible method for continuous processes (Ierapetritou and Floudas, 1998).

The uncertainties of external markets and production processes directly affect the economic benefits of a refinery (Guillén et al., 2006). The scheduling of refinery production under uncertain demand considers production scheduling and fluctuations in demand together, so that the profit loss due to market changes is kept as small as possible. This is similar for the uncertain of production yields. The stochastic scheduling for refinery processes is able to more effectively guarantee the production profit of the refinery, whether facing demand or yield uncertainties.

Stochastic programming, especially two-stage stochastic programming, has been used to address scheduling problems with uncertainty (e.g. Goel and Grossmann, 2006). The uncertainty is represented through a scenario tree on which the stochastic scheduling problem is defined. The two-stage stochastic programming decomposes the scheduling decision process into two stages: the first stage ("here and now") without information on the realization of the uncertainties, while the second stage ("wait and see") has access to the full information on the realization of the uncertainties. Engell et al. (2004) proposed a two-stage stochastic programming model for the batch scheduling of a multiproduct plant. Aiming at the large-scale instances, they used a Lagrangean relaxation method to obtain the lower bounds, and decomposed the subproblems to obtain the upper bounds. Pinto et al. (2009) addressed a design and scheduling problem of multipurpose batch plant under uncertainties. They used a two-stage stochastic Mixed Integer Linear Programming (MILP) model, which determines the design variables in the first stage and scheduling variables in the second stages. Apap and Grossmann (2017) proposed a general modeling and solution method for multi-stage stochastic programming.

There are certain difficulties in the modeling and solution of stochastic scheduling, especially short-term stochastic scheduling. One needs to combine general two-stage stochastic programming with scheduling models, which gives rise to large-scale problems.

To our knowledge, there have been few reported research works on scheduling for whole refinery processes under uncertainties of demand orders and product yields. This paper studies the production scheduling problem with consideration of uncertainties in refinery processes based on the deterministic model of Shah and Ierapetritou (2011). Firstly, the uncertainties of the refinery plant are represented through a scenario tree with values and probabilities of the corresponding uncertain parameters. This scenario tree is used for the two-stage stochastic programming model for the scheduling of refinery production. A hybrid Mixed Integer Nonlinear Programming (MINLP) and Generalized Disjunctive Programming (GDP) model is formulated for the stochastic scheduling problem based on unit-specific continuous-time representation. To address a real large-scale instance, an Outer Approximation method combining with Fix-and-Relax strategy is developed. Finally, the computational experiments based on a real refinery plant are implemented, which show the benefits of the stochastic scheduling method. The computational comparison with the DICOPT solver shows the effectiveness of the proposed solution algorithm.

This paper is organized as follows. Section 2 defines the optimal scheduling problem for refinery processes with consideration of uncertainties. Section 3 formulates a mathematical model for the scheduling problem. Section 4 develops an Outer-Approximation method with Fix-and-Relax strategy to solve the proposed model. Section 5 implements the computational experiments based on the data of a real refinery plant. Finally, we give some conclusions on the proposed model for refinery stochastic scheduling and the solution method.

2. Problem statement

We first give a superstructure flowchart for a typical refinery plant (Shah and Ierapetritou, 2011), as shown in Figure 1. In the super-structure flowchart, the circles denote the materials, like raw materials, intermediate products, component oils, and final oil products, while the squares denote the processing units, like crude distillation units (CDU), fluid catalytic cracking (FCC), and blending tanks. For refining units, such as the CDU, there are different side product streams that are continuously exiting from the CDU.

The stochastic scheduling for the whole refinery defined on the super-structure is as follows:

Given:

1) process topology network of a refinery plant

2) processing modes and production capacity of the units

3) data of raw material, including types, properties, initial inventory, supply capacity

4) yields of side products for each unit under uncertainties

5) initial inventory of each intermediate product and final oil products

6) blending plan

7) demand orders for oil products with consideration of uncertainties

8) utility consumption and capacity

9) inventory capacity for intermediate and final oil products

10) scheduling time horizon.

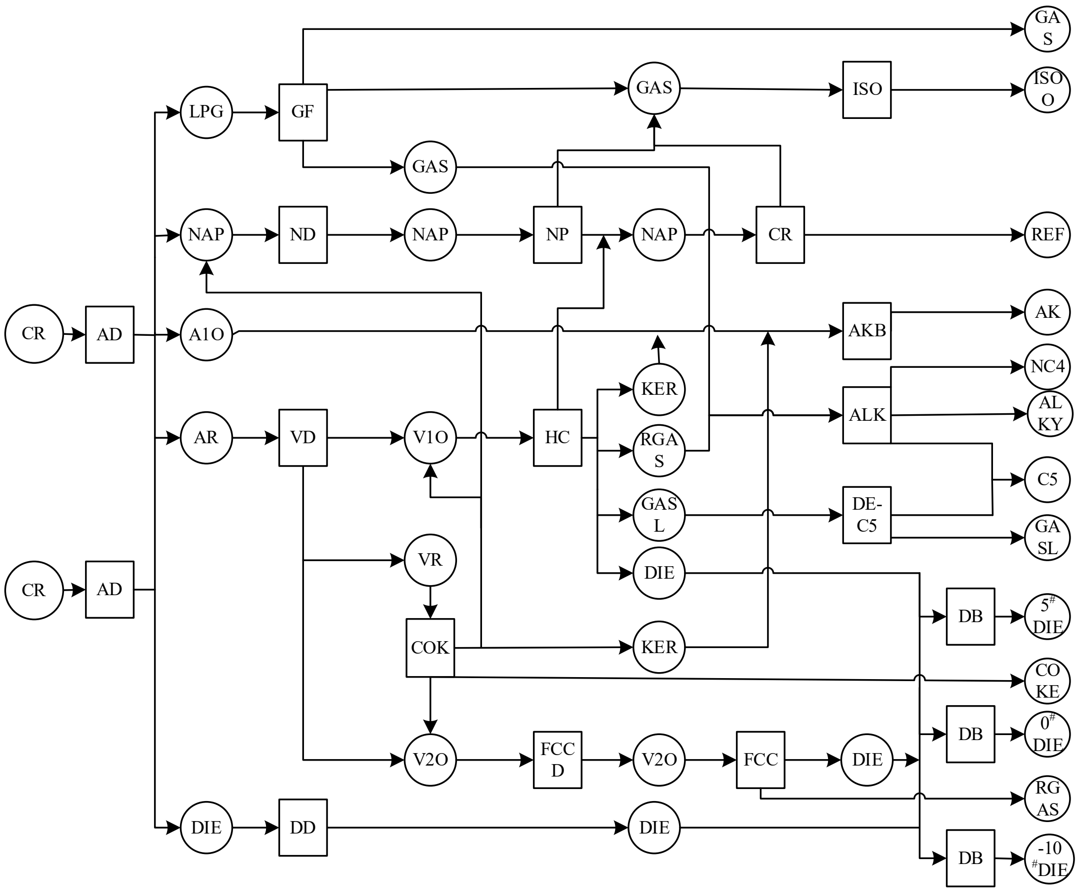

Figure 1. Superstructure flowchart for whole refinery processes

The objective of the stochastic scheduling model for refinery production is to maximize the expected production profit, which equals to the total value of the output product minus the cost of crude oil and consuming energy, the penalties of excess product properties, unsatisfied demand and due date.

Considering the property changes of the material flow over time and the different product yields of the refining units under different working conditions, we first define the production scheduling tasks for the refinery units. The scheduling tasks are then calculated from a production plan. The stochastic scheduling problem of whole-process refinery production requires determining the following items in each scenario:

1) the processing time and processing volume of each processing task in each processing unit

2) the output and properties of intermediate products

3) the types and amounts of oil products

4) the blending ratios of each component for each product

5) the energy consumption of each unit.

The refinery production scheduling must satisfy the following constraints: (1) the material balance constraints, (2) the property balance equations of the processing units and the blending unit, (3) the standard quality indicators for oil products, (4) the market demand for the oil products, (5) the due dates for each order of the oil product, (6) the operation restrictions, and (7) the restrictions for comprehensive energy consumption indicator of each refinery equipment. We assume in this paper that there is no limit to the supply of crude oil, and that there is no material loss during the production processes.

3. Mathematical formulation

We first provide the scenario-tree representation method for the uncertainties. The continuous-time representation method is then introduced to depict the scheduling horizon. Finally, we formulate the mathematical model of the stochastic scheduling for whole refinery processes.

3.1 Scenario tree

Scenario trees are commonly used to represent the uncertainty of production processes or market demand. The scenario tree for refinery uncertainties can be obtained based on the historical data of refinery processes and markets (Calfa et al., 2014). The moment matching and distribution matching methods are generally used to obtain the scenario tree of uncertain data (Høyland and Wallace, 2001; Calfa et al., 2014). A general scenario tree is shown as Figure 2 (a). Each path from the root to the leaves in the tree represents one scenario, and each node contains the value and probability of the uncertain parameter. As there are shared nodes for some scenarios, the scenarios are not directly decomposable for scheduling decisions.

The scenario tree is equivalently reformulated as shown in Figure 2 (b) (Ruszczyński, 1997) with nonanticipativity constraints. In Figure 2 (b), the nodes connected with dotted lines are the nodes with same values and probabilities. The advantage of the reformulated scenario tree is that it is decomposable by scenarios. In this paper, we use the exact moment and distribution method to generate the scenario tree (Calfa et al., 2014). The stochastic scheduling for whole refinery processes is then defined on the reformulated scenarios set as Figure 2 (b).

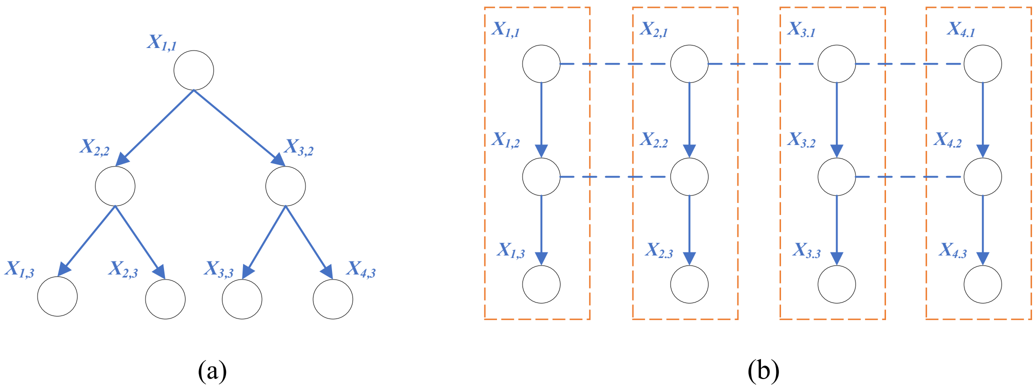

Figure 2. Scenario tree for uncertainties. a) Primal scenario tree, b) Reformulated scenario tree

3.2 Continuous-time representation

Aiming at the continuous processes of refinery production, the unit-specific event-based continuous time representation method is used to represent the starting time and finishing time of the scheduled events for the processing units and storage tanks in Figure 3 (Ierapetritou and Floudas, 1998). The production tasks of the processing units and the flow tasks of the materials are regarded as the scheduling events. The flow tasks include the flow tasks from the processing units to the storage tanks, the flow tasks from the storage tanks to the units, and the flow tasks from the storage tanks to the customer orders. We select a fixed number of event points N_{max} for all the units and tanks, which represent beginnings of tasks. The detailed times of these event points need to be determined.

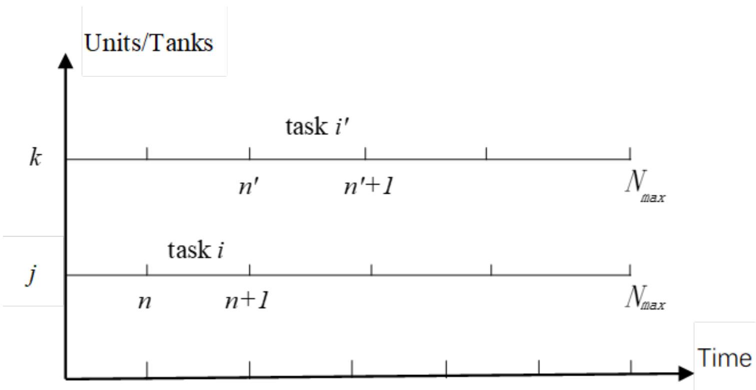

Figure 3. Continuous-time representation for whole refinery processes

3.3 Mathematical model for the refinery stochastic scheduling

Nomenclature	
Indices	
i	tasks

j	units
k	tanks
n	events
o	orders
p	properties
q	scenarios
s	states
Sets	
I_j	set of tasks i that are processed in unit j
I_s^I	set of tasks i that consume state s as feed
I_s^O	set of tasks i that produce material s
J_i	set of units j that processes task i
JC_s	set of units j that take material s as feed
JF_j	set of units that directly followed unit j
JI_k	set of units that are fed from the storage tank k
JO_k	set of units that output materials and storage in the oil storage tank k
JP_s	set of units j that produce out material s
K_s	set of oil tanks that store material s
KD_o	customer order o that correspond oil storage tank k
KF_s	set of oil storage tanks that store final oil product s
KI_j	set of oil storage tanks that store the feed of unit j
KO_j	set of oil storage tank that store the produced material from unit j
O_s	customer order that corresponds the final product s
Q	set of scenarios q
SC_j	set of materials s that are used as feed for unit j
SD_s	set of final products s that flow directly to the market
SI_i	set of materials s that are used as feeds in task i
SJ_j	set of materials s that are produced out from unit j
SO_i	set of materials s that are produced in task i

SP_s	set of final products s that are stored in oil storage tank
ST_k	set of materials s that are stored in the oil storage tank k
Parameters	
CA_j	processing capacity of unit j
CC_s	unit cost of crude oil s
CD_o	unit penalty cost of demand deviation for order o
CM_s	unit penalty cost of market demand deviation for product s
CP_s	unit penalty cost of property deviation for product s
CT_o	unit penalty cost of delivery date deviation for order o
CU_j	unit cost of energy utility for unit j
$D_{o,q}$	demand for final products for orders o under scenario q
$D_{o,s}^{min}$	minimum demand of final product s for customer order o
$D_{o,s}^{max}$	maximum demand of final product s for customer order o
D_s^{mark}	demand for final oil product s flowing directly to the market
$F_{i,j}^{min}$	minimum processing rate of unit j processing task i
$F_{i,j}^{max}$	maximum processing rate of unit j processing task i
G_q	probability of scenario q
H	time horizon of scheduling
N_{max}	total number of event points
PM_s	unit price of final product s flowing to market
PO_o	unit price of submitting order o from tank
$PP_{p,s}^{min}$	lower limitation of property p of material s
$PP_{p,s}^{max}$	upper limitation of property attribute p of material s
$R_{q,s}$	product yield of material s under scenario q
$RI_{i,s}^{min}$	lower limit of proportion of feedstock s in task i
$RI_{i,s}^{max}$	upper limit of proportion of feedstock s in task i
RK_k^{min}	minimum rate of output of final oil product from tank k
RK_k^{max}	maximum rate of output of final oil product from tank k
$RO_{i,s}^{min}$	lower limit of yield of production material s in task i

$RO_{i,s}^{max}$	upper limit of yield of production material s in task i
T_i^{min}	minimum processing time of task i
T_o^{start}	starting time of customer order o
T_o^{end}	finishing time of customer order o
$\bar{U}_j$	unit production energy consumption of unit j
$V_{k,s}^{ini}$	initial inventory of material s in oil storage tank k
V_k^{max}	maximum storage capacity of oil tank k
θ_j	fixed energy consumption fraction of unit j
Binary / Boolean Variables	
$x_{i,j,n,q}$	1 when unit j executes task i at event point n under scenario q, otherwise 0
$X_{i,j,n,q}$	Boolean variable, true when unit j executes task i at event point n under scenario q, otherwise false
$y_{j,k,n,q,s}$	1 when material s flows from unit j to tank k at event point n under scenario q, otherwise 0
$Y_{j,k,n,q,s}$	Boolean variable, true when material s flows from unit j to tank k at event point n under scenario q, otherwise false
$z_{j,k,n,q,s}$	1 when material s flows from tank k to unit j at event point n under scenario q, otherwise 0
$Z_{j,k,n,q,s}$	Boolean variable, true when material s flows from tank k to unit j at event point n under scenario q, otherwise false
$u_{k,n,o,q}$	1 when material s flows from tank k to order o at event point n under scenario q, otherwise 0
$U_{k,n,o,q}$	Boolean variable, true when material s flows from tank k to order o at event point n under scenario q, otherwise false
$w_{k,n,q,s}$	1 when material s stores in tank k at event point n under scenario q, otherwise 0
Continuous Variables	
$BM_{q,s}^{L}$	penalty fraction of final product s flowing directly to market less than demand under scenario q

$BP_{n,p,q,s}^{L}$	penalty fraction of property p of material s less than lower limit of property attribute under scenario q
$BP_{n,p,q,s}^{U}$	penalty fraction of property attribute p of material s greater than upper limit of property attribution under scenario q
$BT_{o,q}^{E}$	penalty fraction for start time of customer order o less than the earliest start time under scenario q
$BT_{o,q}^{L}$	penalty fraction for finishing time of customer order o greater than the latest finishing time under scenario q
$FC_{j,n,q,s}$	flow rate of crude oil s to unit j at event n in scenario q
$FD_{j,j',n,q,s}$	flow of material s directly from unit j to unit j' at event point n under scenario q
$FJ_{j,k,n,q,s}$	flow of material s from unit j to storage tank k at event n under scenario q
$FK_{j,k,n,q,s}$	flow rate of material s from storage tank k to unit j at event point n under scenario q
$FM_{j,n,q,s}$	flow of final product s directly to the market produced by unit j at event point n under scenario q
$FO_{k,n,o,q}$	flow of final product from oil storage tank k to customer order o at point n under scenario q
$FU_{k,n,q,s}$	flow of crude oil s to oil storage tank k at event point n in scenario q
$RP_{p,q,s}$	property p value of material s under scenario q
$RW_{j,q}$	work load of unit j under scenario q
$TI_{i,j,n,q}^{B}$	time that task i starts in unit j at event n under scenario q
$TI_{i,j,n,q}^{E}$	time that task i finishes in unit j at event point n under scenario q
$TJ_{j,k,n,q}^{B}$	time that material flow starts from unit j to oil storage tank k at event n under scenario q
$TJ_{j,k,n,q}^{E}$	time that material flow finishes from unit j to oil storage tank k at event point n under scenario q
$TK_{j,k,n,q}^{B}$	time that material flow starts from the oil storage tank k to unit j at point n under scenario q

$TK^E_{j,k,n,q}$	time that material flow finishes from tank k to unit j at point n under scenario q
$TO^B_{k,n,o,q}$	time that final oil product flow starts flowing from storage tank k to order o at event point n under scenario q
$TO^E_{k,n,o,q}$	time that final oil product flow finishes from storage tank k to order o at event point n under scenario q
$UC_{j,q}$	synthesis energy consumption of unit j under scenario q
$V_{k,n,q,s}$	storage amount of material s in oil storage tank k at event point n under scenario q
$WI_{i,j,n,q,s}$	feeding amount of material s when unit j executes task i at event point n under scenario q
$WO_{i,j,n,q,s}$	production amount of material s when unit j executes task i at event point n under the scenario q

1) Assignment constraints

In each scenario q, each task i is only processed in one unit j at time event n, as stated in Eq. (1).

$$\sum_{j \in J_i} x_{i,j,n,q} \le 1, \qquad \forall\, i \in I, n \in N, q \in Q. \tag{1}$$

The material produced from the unit j can only flow to one storage tank, as stated in Eq. (2). The loading operations and unloading operations of the final oil-product storage tanks are not allowed to occur simultaneously in order to keep the consistence of the oil product qualities. Hence, Eq. (3) indicates that the material flowing from the unit to the storage tank and from the storage tank to the order are not simultaneously performed at the same time event.

$$\sum_{k \in K_s \cap KO_j} y_{j,k,n,q,s} \le 1, \qquad \forall j \in JP_s, n \in N, q \in Q, s \in SP_s \tag{2}$$

$$\sum_{s \in ST_k} y_{j,k,n,q,s} + \sum_{o \in O_s} u_{k,n,o,q} \le 1, \qquad \forall j \in JO_k, k \in K, n \in N, q \in Q. \tag{3}$$

2) Material flow constraints

The mass balance constraints are satisfied for each processing unit, for which the total amount of feeding materials must be equal to the total amount of produced materials under each scenario, as given by Eq. (4). Considering the limitations of the refining technology, the feeding ratio of one material for unit j need be constrained between the upper ratio and lower ratio of the total feeding materials under each scenario, as shown in Eq. (5). Similarly, the product yield of one material is also limited between the upper and lower bounds, as stated in Eq. (6).

$$\sum_{s \in SI_i} WI_{i,j,n,q,s} = \sum_{s \in SO_i} WO_{i,j,n,q,s}, \qquad \forall\, i \in I_j, j \in J, n \in N, q \in Q \tag{4}$$

$$RI_{i,s}^{min} \sum_{s' \in SI_i} WI_{i,j,n,q,s'} \le WI_{i,j,n,q,s} \le RI_{i,s}^{min} \sum_{s' \in SI_i} WI_{i,j,n,q,s'}, \qquad \forall\, i \in IC_s, j \in J_i, n \in N, q \in Q, s \in S \tag{5}$$

$$RO_{i,s}^{min} \sum_{s' \in SI_i} WI_{i,j,n,q,s'} \le WO_{i,j,n,q,s} \le RO_{i,s}^{max} \sum_{s' \in SI_i} WI_{i,j,n,q,s'}, \qquad \forall\, i \in IP_s, j \in J_i, n \in N, q \in Q, s \in S. \tag{6}$$

Eq. (7) describes that the feeding flow of each refinery unit consists the materials feed from the oil storage tank, the materials produced from the last directly connected unit, and the supply of crude oil. Eq. (8) represents that the output material of each processing unit flows to the oil storage tank, the next directly connected unit as feeding stock, and the supply market as the final oil product.

$$\sum_{i \in I_j \cap I_s^C} WI_{i,j,n,q,s} = \sum_{k \in KI_j \cap K_s} FK_{j,k,n,q,s} + \sum_{j' \in JF_j \cap JP_s} FD_{j',j,n,q,s} + FC_{j,n,q,s}, \qquad \forall\, j \in JC_s, n \in N, q \in Q, s \in S \tag{7}$$

$$\sum_{i \in I_j \cap I_s^P} WO_{i,j,n,q,s} = \sum_{k \in KO_j \cap K_s} FJ_{j,k,n,q,s} + \sum_{j' \in JF_j \cap JC_s} FD_{j,j',n,q,s} + FM_{j,n,q,s}, \qquad \forall\, j \in JP_s, n \in N, q \in Q, s \in S. \tag{8}$$

For the first event of the scheduling horizon, the storage amounts of materials in the oil tank k is equal to the initial inventory, plus the sum of the material flow produced from the processing units, and the flow of crude oil to the oil storage tank k, minus the material flow to the following units and the flow to the customer order, as stated in Eq. (9). For the events $n \ge 2$ in the scheduling horizon, the storage amounts of materials in the oil tank k is equal to the storage amount of the event $n-1$, plus the material flow produced from the processing unit and the crude oil flow to the oil storage tank k, minus the material flow from the oil storage tank to the following units and the material flow from the oil storage tank to the customer order, as shown in Eq. (10).

$$V_{k,n,q,s} = V_{k,s}^{ini} + \sum_{j \in JO_k \cap JP_s} FJ_{j,k,n,q,s} + FU_{k,n,q,s} - \sum_{j \in JO_k} FK_{j,k,n,q,s} - \sum_{o \in O_s} FO_{k,n,o,q}, \qquad \forall\, k \in K_s, n = 1, q \in Q, s \in S \tag{9}$$

$$V_{k,n,q,s} = V_{k,n-1,q,s} + \sum_{j \in JO_k \cap JP_s} FJ_{j,k,n,q,s} + FU_{k,n,q,s} - \sum_{j \in JO_k} FK_{j,k,n,q,s} - \sum_{o \in O_s} FO_{k,n,o,q}\,, \quad \forall\, k \in K_s, n \geq 2, n \in N, q \in Q, s \in S\,. \tag{10}$$

3) Processing capacity constraints

The capacities of the processing units and the blending units must also be considered. The total amount of produced materials should lie between the maximum and minimum production capacities, which equal to the flow limitations times the processing durations in Eq. (11). The total amount of produced materials should be less than the maximum flow rate products the scheduling horizon under scenario q if the corresponding task is processed as shown in Eq. (12).

$$F_{i,j}^{min}(TI_{i,j,n,q}^{E} - TI_{i,j,n,q}^{B}) \leq \sum_{s' \in SO_i} WO_{i,j,n,q,s'} \leq F_{i,j}^{max}(TI_{i,j,n,q}^{E} - TI_{i,j,n,q}^{B})\,, \quad \forall\, i \in I, j \in J_i, n \in N, q \in Q \tag{11}$$

$$\sum_{s' \in SO_i} WO_{i,j,n,q,s'} \leq HF_{i,j}^{max}\, x_{i,j,n,q}, \qquad \forall\, i \in I, j \in J_i, n \in N, q \in Q\,. \tag{12}$$

4) Capacity constraints for storage tank

The amount of materials in the storage tank is less than the maximum storage capacity of the oil storage tank, as stated in Eq. (13). Similarly, the material flows from the unit j to the oil storage tank k, the flows from the oil storage tank k to the unit j, and the final oil products from the oil storage tank k to the order o are also be less than the maximum storage capacity of the oil storage tank, shown in Eqs. (14)-(16).

$$V_{k,n,q,s} \leq V_k^{max}\, w_{k,n,q,s}\,, \qquad \forall\, k \in K_s, n \in N, q \in Q, s \in S \tag{13}$$

$$FJ_{j,k,n,q,s} \leq V_k^{max} y_{j,k,n,q,s}\,, \qquad \forall\, j \in J, k \in KO_j, n \in N, q \in Q, s \in ST_k \tag{14}$$

$$FK_{j,k,n,q,s} \leq V_k^{max} z_{j,k,n,q,s}\,, \qquad \forall\, j \in J, k \in KI_j, n \in N, q \in Q, s \in ST_k \tag{15}$$

$$FO_{k,n,o,q} \leq V_k^{max} u_{k,n,o,q}\,, \qquad \forall\, k \in K_s, n \in N, o \in O_s, q \in Q\,. \tag{16}$$

5) Output constraint of final oil products

The delivery amount of the final oil products to one customer order o is less than the maximum delivery rate products the delivery duration, and greater than the minimum delivery rate times the delivery duration, as stated in Eq. (17).

$$RK_k^{min}(TO_{k,n,o,q}^{E} - TO_{k,n,o,q}^{B}) \leq FO_{k,n,o,q} \leq RK_k^{max}(TO_{k,n,o,q}^{E} - TO_{k,n,o,q}^{B})\,, \quad \forall\, k \in KF_s, n \in N, o \in O_s, q \in Q\,. \tag{17}$$

6) Property limitations for final oil products

The properties of the final oil products must satisfy the specified quality requirements, which are based on the standards for final oil products. Some properties of final oil products, like sulfur content, must be less than the limit of the upper bounds as shown in Eq. (18). Some properties of final oil products, like the Octene number, must be greater than the specified lower bound shown in Eq. (19). The relaxed variables are added in Eqs. (18)-(19) in order to describe the amounts exceeding the quality standards, which also are the profit margins for refinery production. Here, the property balance equation of the final products is simplified as that the property of final oil product times the amount of oil products equals the summation of the component properties times the component amount. Since the property and the processing amount of final oil products are all variables, that give rise to the bilinear terms in both Eqs. (18)-(19).

$$\sum_{i \in I_s^O, s' \in SI_i} RP_{p,q,s'} WI_{i,j,n,q,s'} / \sum_{i \in I_s^O} WO_{i,j,n,q,s} \leq RP_{p,s}^{max}\left(1 - BP_{n,p,q,s}^U\right), \quad \forall\, j \in JP_s, n \in N, p \in P, q \in Q, s \in SP_s \tag{18}$$

$$\sum_{i \in I_s^P, s' \in SI_i} RP_{p,q,s'} WI_{i,j,n,q,s'} / \sum_{i \in I_s^O} WO_{i,j,n,q,s} \geq RP_{p,s}^{min}\left(1 + BP_{n,p,q,s}^L\right), \quad \forall\, j \in JP_s, n \in N, p \in P, q \in Q, s \in SP_s\,. \tag{19}$$

7) Demand constraints

The final oil products flowing directly to the market meet the market demand with a penalty fraction as shown in Eq. (20). Considering the uncertainty of product demand, Eq. (21) shows the total flow of the final product from the oil storage tank to the customer order o, is equal to the customer order amount for the final product under the scenario q.

$$D_s^{mark}\left(1 + BM_{q,s}^L\right) \leq \sum_{j \in JP_s, n} FM_{j,n,q,s}\,, \qquad \forall\, q \in Q, s \in SD_s \tag{20}$$

$$\sum_{k \in KD_o, n} FO_{k,n,o,q} = D_{o,q}\,, \qquad \forall\, o \in O_s, q \in Q\,. \tag{21}$$

8) Product yield constraints

The uncertainties of yield make the production amount of the side-products uncertain, which in turn make the flows to the downstream units uncertain. Hence, the uncertain yields affect the production scheduling for the entire refinery process. For each yield scenario, the production amount of a certain side-product is equal to the scenario yield of the side product products the total crude oil feed as shown in Eq. (22).

$$WO_{i,j,n,q,s} = R_{q,s} \sum_{s' \in SI_i} WI_{i,j,n,q,s'}\,, \qquad \forall\, i \in I_s^O, j \in J_i, n \in N, q \in Q, s \in S\,. \tag{22}$$

9) Timing constraints

In order to avoid frequent changeovers of the unit, the processing time of one task in a unit is greater than a minimum time limitation if the logic variable $X_{i,j,n,q}$ is true, as given by the first constraint in Eq. (23). The second constraint in Eq. (23) states that the starting time of the task i' at the $(n+1)^{th}$ event is greater than the finishing time of the task i for the processing unit j. We use a disjunctive term constraint to represent this relationship between assignment variable and timing variables (Grossmann and Trespalacios, 2013). Compared with the big-M constraints from Shah & Ierapetritou (2011), the proposed disjunctive terms constraints are potentially tighter for these time constraints. Hence, we use disjunctive constraints in the following model to depict the timing constraints.

$$\begin{bmatrix} X_{i,j,n,q} \\ TI^{E}_{i,j,n,q} - TI^{B}_{i,j,n,q} \geq T^{min}_{i} \\ TI^{B}_{i\prime,j,n+1,q} \geq TI^{E}_{i,j,n,q} \end{bmatrix} \vee \begin{bmatrix} \neg X_{i,j,n,q} \\ TI^{E}_{i,j,n,q} = TI^{B}_{i,j,n,q} \end{bmatrix}, \tag{23}$$

$$\forall\, i \neq i' \in I, j \in J_i, n \in N, q \in Q\,.$$

Eq. (24) states that the finishing time of the task processed in the unit is greater than its starting time. When the unit executes the same tasks, the starting time of the $(n+1)^{th}$ time event is greater than the finishing time of the n^{th} event point, as stated in Eq. (25).

$$TI^{E}_{i,j,n,q} \geq TI^{B}_{i,j,n,q}\,, \qquad \forall\, i \in I, j \in J_i, n \in N, q \in Q \tag{24}$$

$$TI^{B}_{i,j,n+1,q} \geq TI^{E}_{i,j,n,q}\,, \quad \forall\, i \in I, j \in J_i, n \in N, n < N_{max}, q \in Q\,. \tag{25}$$

For two consecutive processing units with no intermediate storage tank, the starting time and finishing time of the two units executing the continuous tasks are equal. Here, the disjunctive constraints are used to represent the timing relationships when the assignment logic variables are true, as shown in Eq. (26). Eq. (26) is an embedded disjunctive constraint for the timing constraints (Grossmann and Trespalacios, 2013).

$$\begin{bmatrix} X_{i,j,n,q} \\ \vee_{\forall i' \in I_{j'}, j' \in JF_j} \begin{bmatrix} X_{i',j',n,q} \\ TI^{B}_{i',j',n,q} = TI^{B}_{i,j,n,q} \\ TI^{E}_{i',j',n,q} = TI^{E}_{i,j,n,q} \end{bmatrix} \vee \begin{bmatrix} \neg X_{i',j',n,q} \\ TI^{B}_{i',j',n,q} \leq H \\ TI^{E}_{i',j',n,q} \leq H \end{bmatrix} \end{bmatrix} \vee \begin{bmatrix} \neg X_{i,j,n,q} \\ TI^{B}_{i,j,n,q} \leq H \\ TI^{E}_{i,j,n,q} \leq H \end{bmatrix}, \tag{26}$$

$$\forall\, i \in I_j, j \in J, n \in N, q \in Q\,.$$

For the storage tanks k, the finishing time of the material flow from the unit to the storage tank is greater than its starting time as shown in Eq. (27). The starting time of the material flow from the unit to the storage tank at the event point $n+1$ is greater than the finishing time of the current event point n, as stated in Eq. (28). Similarly, the timing constraints for the material flow from oil storage tank k to the unit j are shown by Eqs. (29)-(30).

$$TJ_{j,k,n,q}^{E} \geq TJ_{j,k,n,q}^{B}, \qquad \forall j \in JO_k, k \in K, n \in N, q \in Q \tag{27}$$

$$TJ_{j,k,n+1,q}^{B} \geq TJ_{j,k,n,q}^{E}, \qquad \forall j \in JO_k, k \in K, n \in N, n < N_{max}, q \in Q \tag{28}$$

$$TK_{j,k,n,q}^{E} \geq TK_{j,k,n,q}^{B}, \qquad \forall j \in JO_k, k \in K, n \in N, q \in Q \tag{29}$$

$$TK_{j,k,n+1,q}^{B} \geq TK_{j,k,n,q}^{E}, \qquad \forall j \in JO_k, k \in K, n \in N, n < N_{max}, q \in Q\ . \tag{30}$$

For the storage tank k of the intermediate products, the starting (finishing) time of the material flow from the unit to the intermediate tank is equal to the starting (finishing) time of the material from the intermediate tank flowing to the next unit when the loading and unloading assignment logic variables for tank k are both true at the same event n as shown in Eq. (31).

$$\begin{bmatrix} Y_{j,k,n,q,s} \\ \vee_{j' \in JI_k} \begin{bmatrix} U_{j',k,n,q,s} \\ TJ_{j,k,n,q}^{B} = TK_{j',k,n,q}^{B} \\ TJ_{j,k,n,q}^{E} = TK_{j',k,n,q}^{E} \end{bmatrix} \vee \begin{bmatrix} \neg U_{j',k,n,q,s} \\ TK_{j',k,n,q}^{B} \leq H \\ TK_{j',k,n,q}^{E} \leq H \end{bmatrix} \end{bmatrix} \vee \begin{bmatrix} \neg Y_{j,k,n,q,s} \\ TJ_{j,k,n,q}^{B} \leq H \\ TJ_{j,k,n,q}^{E} \leq H \end{bmatrix}, \tag{31}$$

$$\forall j \in JO_k, k \in K, n \in N, q \in Q, s \in S\ .$$

Considering the continuous processes, for the storage tank k storing the material produced from the unit j, the beginning and finishing times processing the task i in the unit j are equal to the beginning and finishing times for the material flowing from the unit j to the oil storage tank k when both assignment variables $Y_{j,k,n,q,s}$ and $X_{i,j,n,q}$ are true as stated in Eq. (32).

$$\begin{bmatrix} Y_{j,k,n,q,s} \\ \vee_{i \in I_j} \begin{bmatrix} X_{i,j,n,q} \\ TJ_{j,k,n,q}^{B} = TI_{i,j,n,q}^{B} \\ TJ_{j,k,n,q}^{E} = TI_{i,j,n,q}^{E} \end{bmatrix} \vee \begin{bmatrix} \neg X_{i,j,n,q} \\ TI_{i,j,n,q}^{B} \leq H \\ TI_{i,j,n,q}^{E} \leq H \end{bmatrix} \end{bmatrix} \vee \begin{bmatrix} \neg Y_{j,k,n,q,s} \\ TJ_{j,k,n,q}^{B} \leq H \\ TJ_{j,k,n,q}^{E} \leq H \end{bmatrix}, \tag{32}$$

$$\forall j \in JO_k, k \in K_s, n \in N, q \in Q, s \in S.$$

For the storage tank k of crude oil, the beginning and finishing times of the crude oil flows from the storage tank k to the processing unit j is same to the beginning and finishing times of the task i processed in the unit j at time event n when the assignment logic variables $Z_{j,k,n,q,s}$ and $X_{i,j,n,q}$ are both true as shown by Eq. (33).

$$\begin{bmatrix} Z_{j,k,n,q,s} \\ \vee_{i \in I_j} \begin{bmatrix} X_{i,j,n,q} \\ TK_{j,k,n,q}^{B} = TI_{i,j,n,q}^{B} \\ TK_{j,k,n,q}^{E} = TI_{i,j,n,q}^{E} \end{bmatrix} \vee \begin{bmatrix} \neg X_{i,j,n,q} \\ TI_{i,j,n,q}^{B} \leq H \\ TI_{i,j,n,q}^{E} \leq H \end{bmatrix} \end{bmatrix} \vee \begin{bmatrix} \neg Z_{j,k,n,q,s} \\ TK_{j,k,n,q}^{B} \leq H \\ TK_{j,k,n,q}^{E} \leq H \end{bmatrix}, \tag{33}$$

$$\forall j \in JO_k, k \in K_s, n \in N, q \in Q, s \in S.$$

For the storage tank k of the final oil products, the finishing time of the final product flowing from the oil storage tank k to the order o is greater than its start time, as shown in Eq. (34). Eq. (35) states that the starting time of the oil product flowing from the storage tank k to the order o at the $(n+1)^{th}$ event point is greater than the finishing time of the n^{th} event point.

$$TO_{k,n,o,q}^{E} \geq TO_{k,n,o,q}^{B}, \qquad \forall\, k \in KF_s, n \in N, o \in O_s, q \in Q \tag{34}$$

$$TO_{k,n+1,o,q}^{B} \geq TO_{k,n,o,q}^{E}, \qquad \forall\, k \in KF_s, n \in N, n < N_{max}, o \in O_s, q \in Q. \tag{35}$$

Eq. (36) requires that the starting time of the final product flowing from the storage tank k to the next order o' at the $(n+1)^{th}$ event be greater than the finishing time of the order o at the n^{th} event when the assignment logic variable $U_{k,n,o,q}$ is true.

$$\begin{bmatrix} U_{k,n,o,q} \\ TO_{k,n+1,o',q}^{B} \geq TO_{k,n,o,q}^{E} \end{bmatrix} \vee \begin{bmatrix} \neg U_{k,n,o,q} \\ TO_{k,n,o,q}^{E} \leq H \end{bmatrix}, \qquad \forall\, k \in KF_s, n \in N, n < N_{max}, o \neq o' \in O_s, q \in Q. \tag{36}$$

Eq. (37) states that the starting time of the oil product flowing from the storage tank k to the order o at the $(n+1)^{th}$ event point is greater than the finishing time of the material flow from the unit j to the oil storage tank k at the n^{th} event point.

$$\begin{bmatrix} Y_{j,k,n,q,s} \\ TO_{k,n+1,o,q}^{B} \geq TJ_{j,k,n,q}^{E} \end{bmatrix} \vee \begin{bmatrix} \neg Y_{j,k,n,q,s} \\ TJ_{j,k,n,q}^{E} \leq H \end{bmatrix}, \qquad \forall\, j \in JO_k, k \in K_s, n \in N, n < N_{max}, o \in O_s, q \in Q, s \in S. \tag{37}$$

The starting time of the material flow from the unit j to the oil storage tank k at the $(n+1)^{th}$ event point is greater than the finishing time of product flow from the storage tank k to the order o at the n^{th} event point under scenario q as shown in Eq. (38). The time of the final product flowing from the storage tank k to the order o must be within the customer's order time windows under scenario q. Here, the relaxed variables $BT_{o,q}^{B}$ and $BT_{o,q}^{E}$ of the delivery times are introduced to relax the order delivery time windows.

$$\begin{bmatrix} U_{k,n,o,q} \\ TJ_{j,k,n+1,q}^{B} \geq TO_{k,n,o,q}^{E} \\ TO_{k,n,o,q}^{B} \geq T_o^{start}\left(1 - BT_{o,q}^{B}\right) \\ TO_{k,n,o,q}^{E} \leq T_o^{end}\left(1 + BT_{o,q}^{E}\right) \end{bmatrix} \vee \begin{bmatrix} \neg U_{k,n,o,q} \\ TO_{k,n,o,q}^{B} \geq 0 \\ TO_{k,n,o,q}^{E} \leq H \end{bmatrix}, \qquad \forall\, j \in JO_k, k \in KF_s, n \in N, n < N_{max}, o \in O_s, q \in Q. \tag{38}$$

10) Energy consumption constraints

The total energy consumption of processing unit j in refinery production under each scenario q is calculated from Eqs. (39)-(40) (Guo and Xu, 2004).

$$RW_{j,q} = \frac{\sum_{i \in I_j, s \in SC_{j,n}} WI_{i,j,n,q,s}}{CA_j}, \qquad \forall\, j \in J, q \in Q \tag{39}$$

$$UC_{j,q} = \bar{U}_j \left[1 + \theta_j \left(\frac{1}{RW_{j,q}} - 1\right)\right] \sum_{i \in I_j, s \in SC_{j,n}} WI_{i,j,n,q,s}\ , \quad \forall\, j \in J, q \in Q. \tag{40}$$

11) Objective function

The objective function is to maximize the expected net profit, which is equal to the summation of the scenario probabilities times the net total profit under each scenario as shown in Eq. (41). The expected production profit is equal to the total income of final oil product minus the overall cost of crude oil, the penalties of product properties, demand, due date and energy consumption.

$$\begin{aligned} \max\ \ Z^{NETV} = \sum_{q \in Q} G_q \Big[& \sum_{k \in K_s, n, o \in O_s} PO_o FO_{k,n,o,q} + \sum_{j \in JP_s, n, s \in SD_s} PM_s FM_{j,n,q,s} \\ & - \sum_{k \in K_s, n, s} CC_s FU_{k,n,q,s} - \sum_{j,n,s} CC_s FC_{j,n,q,s} \\ & - \sum_{n,p,s \in SP_s} CP_{p,s} (BP^L_{n,p,q,s} + BP^U_{n,p,q,s}) - \sum_{s \in SD_s} CM_s BM^L_{q,s} \\ & - \sum_{o \in O_s} CT_o (BT^B_{o,q} + BT^E_{o,q}) - \sum_j CU_j UC_{j,q} \Big]. \end{aligned} \tag{41}$$

The proposed model is a hybrid MINLP and GDP formulation with integer and logic variables, continuous variables, nonconvex bilinear constraints. Furthermore, there are up to five indices in the variables making the problem size to be very large for real instances.

Since the refinery scheduling model can be decomposed by scenario q, we can represent in general form the scheduling model as follows:

$$\begin{aligned} \max\quad & Z^{NETV} = \textstyle\sum_{q \in Q} G_q \big(f(\tilde{X}_q) + C^T \tilde{Y}_q\big) \\ \text{s.t.}\quad & \mathrm{A}_q \tilde{X}_q + B_q \tilde{Y}_q \le \tilde{b}_q \quad \forall\, q \in Q \\ & \mathrm{V}_{q \in Q} \begin{bmatrix} \tilde{Y}_q = 1 \\ \mathrm{A}'_q \tilde{X} \le \tilde{b}'_q \end{bmatrix} \\ & \tilde{X}_q, \tilde{Y}_q \in \{0, 1\}, \end{aligned} \tag{42}$$

where the binary variable $\tilde{Y}_q$ is related to the realization of scenario q, $\tilde{X}_q$ corresponds to the continuous variables under the scenario q. Therefore, the proposed mathematical model can be decomposed by scenarios.

The general disjunctive constraints in Eq. (42) are reformulated into algebraic constraints through big-M and hull relaxation (Grossmann and Trespalacios, 2013). Compared with the big-M reformulation of GDP model, the hull relaxation reformulation generally leads to tighter representation for primal problem at the expense of increasing number of variables. Considering the large-scale of the proposed model, we apply the big-M reformulation for the disjunctive term constraints. In this way the proposed hybrid model is transformed into an MINLP model, which is able to be solved through standard MINLP solvers.

4. Outer Approximation based on Fix-and-Relax strategy

The transformed MINLP model can be solved through a general MINLP method, like Outer Approximation (OA), Generalized Benders Decomposition (GBD), LP-based Branch and Bound etc. (Grossmann, 2002). Duran and Grossmann (1986) proposed the OA algorithm for general MINLP models, which decomposes the MINLP into a Nonlinear Programming (NLP) subproblem and an MILP master problem. With fixed binary variables, the MINLP model is reduced to an NLP model, which predicts lower bounds of the MINLP model. Based on linearization of nonlinear functions, the MINLP is approximated into an MILP, which is solved to obtain a new value for the binary variables and an upper bound of the primal problem. The OA method solves NLP and MILP iteratively until satisfying the convergence conditions. Su et al. (2015, 2018) addressed the improved strategies for OA with multi-generation cuts, hybrid cuts, partial surrogate cuts and quadratic cuts.

The scheduling model for the whole-process refinery production problem under uncertainties contains a large number of discrete, continuous variables, linear and nonlinear equations. Moreover, by increasing the number of scenarios, the problem size of the MINLP model greatly increases. The general OA method has difficulty in efficiently solving such the large-scale MINLPs.

Escudero and Salmeron (2005) and Cadarso et al. (2018) designed the Fix-Relax method to decompose the large-scale MILP problem into multiple MILP sub-problems based on the subset-division of integer variables. By fixing or relaxing the integer variables of certain MILP sub-problems, the MILP solution method based on Fix-and-Relax strategy is able to reduce the solution time of the primal MILP.

To address the proposed large-scale decomposable MINLP model with scenarios, we develop OA based on Fix-and-Relax strategy (OA-FR). First, the scenarios are clustered into the subsets according to certain criteria. Considering that the scheduling problem is decomposable by scenario, the primal MINLP problem is decomposed into multiple MINLP sub-problems according to the subset of scenarios. Then by fixing and relaxing the integer variables of the subset problems in turn, we solve the relaxed MINLP subproblems to replace the solution of primal MINLP.

The scenario set Q is decomposed into several subset $Q_{m'}^{SUB}$, $m' = 1, \dots, M$, as shown in Figure 4. For each scenario subset m, the scenarios subsets are divided into three parts as $(1, \dots, m-1)$, m and $(m+1, \dots, M)$. For the first part of the subsets with $m' = 1, \dots, m-1$, we fix the values of the continuous and integer variables as X_q^{FIX}, Y_q^{FIX}. For the third part of the subsets with $m' = m+1, \dots, M$, we relax the integer variables into the continuous variables as $Y_q \in [0,1]$. All the variables in the second part of the subset $m' = m$ and the continuous, relaxed integer variables in the third part $m' = m+1, \dots, M$ are to be determined in this iteration.

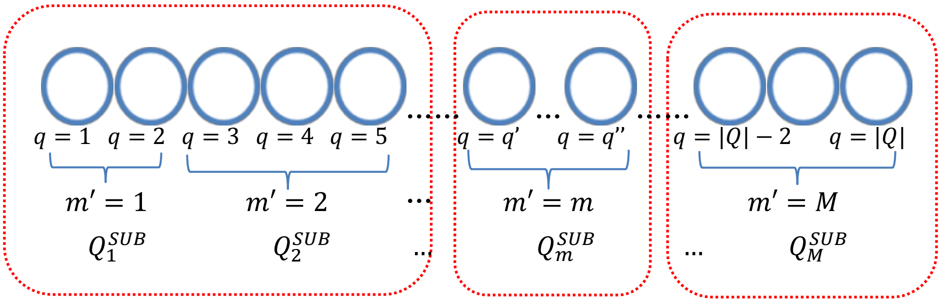

Figure 4. Illustration of clustering scenarios into M subsets and three subset parts for each m

The generated subproblem $subP_m$, $m = 1, \dots, M$ model in OA-FR is given by the following formulation:

$$
\begin{aligned}
\max \quad & Z_m^{NETV} = \sum_q G_q\left(f(X_q) + C^T Y_q\right) \\
\text{s.t.} \quad & Eqs.\ (1)-(41) \qquad \forall\, q \in Q \\
& X_q = X_q^{FIX}, Y_q = Y_q^{FIX} \quad \forall\, q \in Q_{m'=1,\dots,m-1}^{SUB} \\
& X_q, Y_q \quad \forall\, q \in Q_{m'=m}^{SUB} \\
& X_q, Y_q \in [0,1] \qquad \forall\, q \in Q_{m'=m+1,\dots,M}^{SUB} \, .
\end{aligned}
\qquad (subP_m)
$$

Here, the first subproblem, $subP_{m=1}$, is obtained by relaxing the integer variables in scenario subset $m' = 2, \dots, M$ into continuous variables between [0, 1]. The last subproblem, $subP_{m=M}$, is then obtained by fixing the variables in scenario subset $m' = 1, \dots, M-1$. The subproblems $subP_m$, $m = 1, \dots, M-1$ are the relaxed MINLP models of primal model with less integer variables. The subproblem $subP_M$ is the fixed MINLP model of primal model.

Proposition 1. If $subP_1$ is feasible and the optimal objective value is Z_1^{NETV}, the upper bound of the optimum for the primal problem is Z_1^{NETV}. The upper bounds are updated iteratively when the better bound obtained from solving $subP_m$ with $m < M$. The optimal objective value of $subP_M$ is the lower bound of the optimum for the primal problem. ☐

Proposition 1 is self-evident based on the construction of the subproblems, which are either relaxed subproblems $subP_{m'}$, $m' = 1, \ldots, M-1$, or fixed subproblem $subP_M$.

The proposed large-scale primal MINLP is decomposed into M subproblems with fewer integer and continuous variables and constraints. This OA based on Fix-and-Relax strategy sequentially solves the generated subproblem $subP_m$, as shown in Figure 5. Therefore, the solution of the primal model is replaced by the solution of M subproblems, which greatly reduces the solution difficulties. It should be pointed out that the solution obtained by the proposed OA-FR solution method is in general a suboptimal solution of the primal model since global optimality cannot be guaranteed.

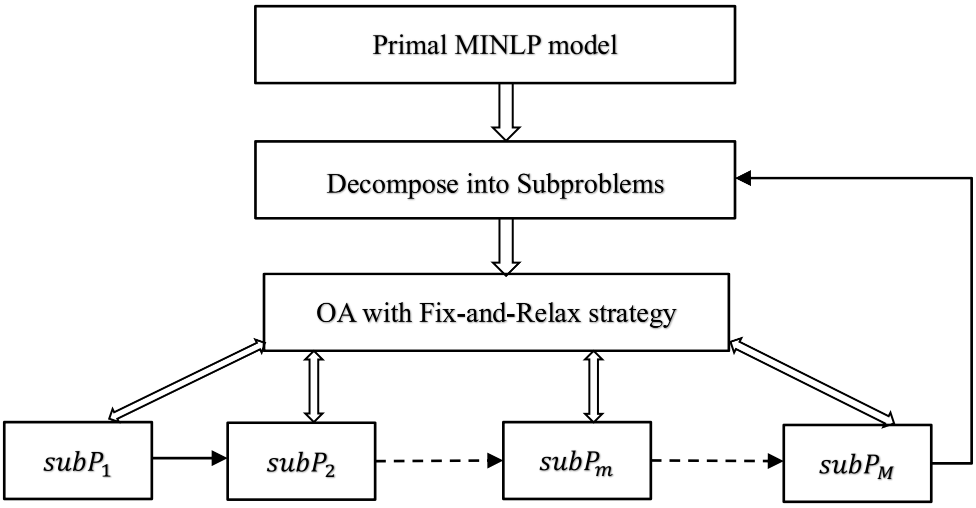

Figure 5. Illustration of Outer-Approximation with Fix-and-Relax strategy

The implementation steps of the OA-FR are described as follows:

Initialize: Cluster the scenario set Q into subsets Q_m^{SUB}, $m = 1, \ldots, M$. Set $m = 1$, the upper bound of the primal objective function as Eq. (41) $\overline{Z^{NETV}} = \infty$.

Step 1: Generate the subproblem $subP_1$ according to the fixing and relaxing strategy for the integer variables. For $subP_1$, we relax the integer variables of the scenarios $m = 2, \ldots, M$ into the continuous variables within the range [0, 1].

Step 2: Solve the subproblem $subP_1$ using the general OA algorithm. If the subproblem $subP_1$ is feasible and the optimal value is $Z_{m=1}^{*NETV}$, the upper bound of the objective function value of the primal problem is updated with $\overline{Z^{NETV}} = Z_{m=1}^{*NETV}$. Otherwise, if $subP_1$ infeasible, the primal model is also infeasible. STOP.

Step 3: If $m < M$, $m = m + 1$. Generate subproblem model $subP_m$. Then solve the subproblem $subP_m$ using the general OA algorithm. If $subP_m$ is feasible, and the value of the objective function $Z_m^{*NETV} < \overline{Z^{NETV}}$, the upper bound is updated with $\overline{Z^{NETV}} = Z_m^{*NETV}$.

***Step 4*:** When $m = M$, the optimal objective value of the subproblem $subP_M$ is the lower bound of the primal problem $\underline{Z^{NETV}}$, which also means the primal problem is feasible.

***Step 5*:** If the termination condition is satisfied, the solution of subproblem $subP_M$ yield the optimal solution of the primal problem is STOP.

Remark. When OA-FR terminates, and the relative gap $(\overline{Z^{NETV}} - \underline{Z^{NETV}})/\underline{Z^{NETV}}$ is not accepted, the scenario subsets can be clustered again. Eq. (43) states merging two adjoining scenario subsets into one subset, or dividing one scenario subset into two subsets. Then we restart the OA-FR solution procedures based on updating scenario clustering subsets.

$$Q_{m-1}^{SUB} \cup Q_m^{SUB} \Rightarrow Q_m^{SUB},\ Q_m^{SUB} \Rightarrow Q_m^{SUB} \cup Q_{m+1}^{SUB}. \quad (43)$$

We need to point out that the different clustering results of the scenarios lead to different subproblems, which may obtain different scheduling solutions with OA-FR for the same instance.

5. Numerical experiments

We implement our numerical experiments in GAMS 33.2.0 (Brook et al., 1988). The operating system is Windows 10, and Intel Core 2 Duo, CPU 2.7 GHz and 16 GB of RAM. The solutions of the MINLP problems are obtained using the solver DICOPT (Grossmann et al., 2002), and the solutions of the NLP and MILP problems are obtained using the CONOPT 3.0 (Drud, 1994) and the CPLEX 12.8 solvers (IBM CPLEX, 2018), respectively.

5.1 Data

The deterministic data of the numerical experiments is from a real refinery plant (Shah and Ierapetritou, 2011), whose superstructure is shown in Figure 1. There are 17 processing and blending units, 2 crude oil storage tanks, 5 intermediate storage tanks and 5 final oil storage tanks. There are 2 types of crude oil and13 types of final oil product with consideration of 2 types of properties. The scheduling horizon is 72 and 96 hours. The number of the scheduling tasks is 19, and the number of the material states is 35. We set the number of events for each unit to 5 and 6, respectively. There are 4 customer orders that must be satisfied during the scheduling horizon, in which the demand of Orders 1 and 2 are uncertain. The demands and time windows of the 4 customer orders are given in Table 1. The demands of the final oil products flowing directly to the market are shown in Table 2.

Table 1. Demand Information of Customer Orders

Customer Order	Order 1	Order 2	Order 3	Order 4
Product category	$5^{\#}$ DIE	$0^{\#}$ DIE	$-10^{\#}$ DIE	AK
product name	$5^{\#}$ Diesel	$0^{\#}$ Diesel	$-10^{\#}$ Diesel	Jet Fuel
Product demand for 72 hours (kbbl)	Uncertainty	Uncertainty	80	200
Time windows for 72 hours (h)	[58, 71]	[10, 20]	[28, 46]	[40, 70]
Product demand for 96 hours (kbbl)	Uncertainty	Uncertainty	100	200
Time windows for 96 hours (h)	[68, 90]	[10, 20]	[28, 46]	[40, 80]

Table 2. Demand of Final Oil Product Flowing Directly to Market

Product Category	GAS	ISOO	REF	NC4	ALKY	C5	GASL	COKE	RGAS
demand (kbbl)	0	5	5	5	20	20	20	0	5

We consider two types of uncertainties, order demand and product yield, for the production scheduling. The demand data and yield data are randomly generated within certain ranges, then obtained the scenario tree through Distribution Matching (Calfa et al., 2014). The scenario data of the demand uncertainties is given in Table 3, which is depicted through the scenario trees including 4-6 nodes. Case 1 is the deterministic instance with the given demand. Cases 2-6 are the instances for the horizon of 72 hours with uncertainties in the demands. Cases 7-10 are the instances for the horizon of 96 hours. For Cases 5 and 10, there are two types of the product demand with uncertainties, which are depicted through a scenario tree with the three-dimension node values.

Table 3. Scenario data for uncertain demand of the final oil products

Case	H(h)	Demand Uncertainties	Num. Scenes	Demand and Probabilities under Scenario Q						
				Scenarios	1	2	3	4	5	6
1	72	$5^{\#}$DIE	1	Deter. (kbbl)	40.99	/	/	/	/	/
2	72	$5^{\#}$DIE (Order1)	4	Demand (kbbl)	23.35	34.08	46.77	59.77	/	/
				Probabilities	0.14	0.27	0.26	0.33	/	/
3	72	$5^{\#}$DIE (Order1)	5	Demand (kbbl)	18.48	24.46	35.52	47.55	59.95	/
				Probabilities	0.01	0.13	0.28	0.25	0.34	/

4	72	$0^{\#}$DIE (Order2)	5	Demand (kbbl)	43.24	44.46	49.45	55.31	59.12	/
				Probabilities	0.10	0.15	0.24	0.29	0.22	/
5	72	$5^{\#}$DIE(Order1) $0^{\#}$DIE(Order2)	5	$5^{\#}$ Demand (kbbl)	18.48	24.46	35.52	47.55	59.95	/
				$0^{\#}$ Demand (kbbl)	43.24	44.46	49.45	55.31	59.12	/
				Probabilities	0.01	0.13	0.27	0.25	0.34	/
6	72	$0^{\#}$DIE(Order2)	6	$0^{\#}$ Demand (kbbl)	43.24	44.46	49.45	55.31	57.12	59.02
				Probabilities	0.002	0.15	0.24	0.29	0.22	0.10
7	96	$5^{\#}$DIE (Order1)	4	Demand (kbbl)	34.46	45.52	57.55	65.95	/	/
				Probabilities	0.15	0.24	0.29	0.22	/	/
8	96	$5^{\#}$DIE (Order1)	5	Demand (kbbl)	28.48	34.46	45.52	57.55	65.95	/
				Probabilities	0.002	0.15	0.24	0.29	0.22	/
9	96	$0^{\#}$DIE (Order2)	6	Demand (kbbl)	43.24	44.46	49.45	55.31	57.12	59.02
				Probabilities	0.002	0.15	0.24	0.29	0.22	0.10
10	96	$5^{\#}$DIE(Order1) $0^{\#}$DIE(Order2)	6	$5^{\#}$ Demand (kbbl)	28.48	34.46	45.52	57.55	65.95	69.95
				$0^{\#}$ Demand (kbbl)	43.24	44.46	49.45	55.31	57.12	59.02
				Probabilities	0.002	0.15	0.24	0.29	0.22	0.10

The data of the CDU side-product yield uncertainties is shown in Table 4. Case 11 is with deterministic side-product yields, and Cases 12-20 are with uncertain side-product yields. To represent the realization of the yield uncertainties, we use the scenario tree with 3, 4, 5 and 6 nodes to represent the yield uncertainties.

Table 4. Scenario data for uncertain yields of side-products

Case	H(h)	Yield Uncertainties	Num. Scenes	Yield and Probabilities under Scenario Q						
				Scenarios	1	2	3	4	5	6
11	72	Side-product 1	1	Yield (%)	4.64					
12	72	Side-product 1	3	Yield (%)	4.03	4.54	5.36	/	/	/
				Probabilities	0.21	0.39	0.40	/	/	/
13	72	Side-product 1	5	Yield (%)	3.80	4.16	4.48	5.09	5.53	/
				Probabilities	0.19	0.26	0.25	0.19	0.11	/

14	72	Side-product 1&2	4	Yield 1 (%)	3.91	4.57	5.00	5.35	/	/
				Yield 2 (%)	5.82	6.33	6.71	7.01	/	/
				Probabilities	0.20	0.21	0.33	0.26	/	/
15	72	Side-products 1&2	5	Yield 1 (%)	3.13	3.97	4.51	5.06	5.49	/
				Yield 2 (%)	5.09	5.99	6.45	7.10	7.36	/
				Probabilities	0.14	0.24	0.29	0.17	0.16	/
16	96	Side-products 1	4	Yield 1 (%)	4.03	4.54	5.06	5.36	/	/
				Probabilities	0.11	0.25	0.40	0.24	/	/
17	96	Side-products 1	5	Yield 1 (%)	3.80	4.16	4.48	5.09	5.53	/
				Probabilities	0.19	0.26	0.25	0.19	0.11	/
18	96	Side-products 2	5	Yield 2 (%)	5.82	6.33	6.71	7.01	7.36	/
				Probabilities	0.19	0.26	0.25	0.19	0.11	/
19	96	Side-products 1&2	5	Yield 1 (%)	3.80	4.16	4.48	5.09	5.53	/
				Yield 2 (%)	5.82	6.33	6.71	7.01	7.36	/
				Probabilities	0.19	0.26	0.25	0.19	0.11	/
20	96	Side-products 1&2	6	Yield 1 (%)	3.80	4.16	4.48	5.09	5.30	5.53
				Yield 2 (%)	5.82	6.33	6.71	7.01	7.22	7.36
				Probabilities	0.10	0.21	0.20	0.19	0.11	0.19

We also consider the mixed uncertainties of the demand and the yield, which is given in Table 5. The scheduling horizon is 72 hours. The values of the scenario nodes are also obtained through Distribution Matching (Calfa et al. 2014), in which more than one uncertain entity are considered. Here, the distributions of demand and yield are assumed consistent. Cases 21, 22 and 23 are the mixed uncertainties of the demand for Order 1 and the yield of side-product 1, which scenario trees contain 3-5 nodes. Case 24 and Case 25 are the more complex instances, with respectively two type uncertainties of orders or yields mixing with one uncertainty of yield or order.

Table 5. Data for simultaneous uncertainties of demand and yield

Case	H(h)	Simultaneous Uncertainties	Num. Scenes	Demand and Probabilities under Scenario Q					
				Scenarios	1	2	3	4	5

21	72	Order 1/ Side-product 1	3	Demand (kbbl)	43.35	54.08	59.77	/	/
				Yield (%)	4.03	4.53	5.36	/	/
				Probabilities	0.21	0.39	0.40	/	/
22	72	Order 1 / Side-product 1	4	Demand (kbbl)	43.35	54.08	56.77	59.80	/
				Yield (%)	4.03	4.53	5.36	5.50	/
				Probabilities	0.11	0.34	0.40	0.15	/
23	72	Order 1/ Side-product 1	5	Demand (kbbl)	43.35	47.07	54.08	56.77	59.80
				Yield (%)	4.03	4.27	4.54	5.36	5.50
				Probabilities	0.11	0.33	0.40	0.15	0.12
24	72	Orders 1&2/ Side-product 1	5	Order 1 (kbbl)	43.35	47.07	54.08	56.77	59.80
				Order 2 (kbbl)	43.24	44.46	49.45	55.31	59.12
				Yield 1 (%)	4.03	4.27	4.54	5.36	5.50
				Probabilities	0.11	0.33	0.40	0.15	0.12
25	72	Orders 1/ Side-products 1&2	5	Order 1 (kbbl)	43.35	47.07	54.08	56.77	59.80
				Yield 1 (%)	3.13	3.97	4.51	5.06	5.49
				Yield 2 (%)	5.09	5.99	6.45	7.10	7.36
				Probabilities	0.11	0.33	0.40	0.15	0.12

5.2 Results

We report the computational results of the instances with considerations of uncertain demand and yields, which are divided into three parts: the demand uncertainty, yield uncertainty and mixed uncertainty.

5.2.1 Demand Uncertainty

First, we solve the reformulated MINLP model with the input data Table 3 using DICOPT. The statistics of the model sizes and the results are shown in Table 6. The sizes of the deterministic equivalent stochastic model are generally 4-6 times of the original deterministic model, regardless of the number of integer variables, continuous variables and constraints. The solution time of the deterministic Case 1 is 5% of the solution time of the stochastic Cases 2-6 with the same horizon length 72 hours.

For deterministic Case 1, the net profit for the average demand is less than the net profit of the expected demand. The expected net profit of stochastic model is greater than the net profit of the deterministic model, compared Case 1 and Case 2. When the horizon length is increased to 96 hours, the solution time of the stochastic cases increases almost 10 times. When the scenario tree keeps the same structure with same numbers of nodes, like Case 4 and Case 5, the solution time is of the same magnitude order. Although there are two stochastic demand of customer orders in Case 5, compared there is only one uncertain order in Case4. The number of scenarios also partly determines the computational times in Cases 1-10.

Table 6. Statistics for Cases 1-10 with uncertainties of order demand

Case	H(h)	Num. Scenes	Bin. Var.	Cont. Var.	Const.	Iter.	Obj. Value ($)	CPU (s)
1	72	0	640	5304	6354	5	38759.04	13.57
2	72	4	2560	21513	25553	5	39013.95	226.58
3	72	5	3200	26891	31941	3	39066.14	182.49
4	72	5	3200	26891	31941	3	38812.11	269.81
5	72	5	3200	26891	31941	4	39411.86	349.26
6	72	6	3840	32269	38329	3	38914.49	335.08
7	96	4	3072	25773	30769	5	40177.09	1788.15
8	96	5	3840	32216	38461	4	40188.13	2826.75
9	96	6	4608	38659	46153	5	44770.81	2020.39
10	96	6	4608	38659	46153	3	44238.21	1127.38

To decompose the problem, we apply our designed OA with Fix-and-Relax strategy to solve stochastic Cases 2-10, whose results are shown in Table 7. Here, the statistics of the model size is one subproblem $subP_m$ in OA-FR. For Case 2, OA-FR solves 4 subproblems with the listed problem sizes. We decompose the proposed model according to the scenarios and solve the subproblems sequentially with OA-FR for Cases 2, 3, 4, 5, 7, 8. For Cases 6, 9, 10 with six scenarios, we merge the first two subproblems into one subproblem and solve 5 subproblems in the solution process of OA-FR.

Table 7. Comparison results of OA-FR algorithm with DICOPT for Cases 2-10

Case	*Bin. Var.	*Cont. Var.	*Const.	OA-FR		DICOPT		Gap OBJ	Saving CPU
				Obj. ($)	CPU (s)	Obj. ($)	CPU (s)	%	%
2	320	37173	35039	38895.77	**89.16**	39013.95	226.58	0.30	60.65
3	320	47904	44482	39056.99	**166.37**	39066.14	182.49	0.02	8.83
4	320	47904	44482	38509.62	**143.66**	38812.11	269.81	0.78	46.76
5	320	47904	44482	39094.55	**163.73**	39411.86	349.26	0.81	53.12
6	320	58295	53925	38875.64	**178.68**	38914.49	335.08	0.10	46.68
7	384	44809	42035	39566.92	**116.96**	40177.09	1788.15	**1.52**	**93.46**
8	384	57213	53378	40011.44	**296.40**	40188.13	2826.75	0.44	89.51
9	384	69617	64703	44563.11	**327.60**	44770.81	2020.39	0.46	83.79
10	384	69617	64703	43964.52	**482.39**	44238.21	1127.38	0.62	57.21

*Statistics numbers of variable and constraints are only problem size of the last subproblem in OA-FR.

Compared with the simultaneous primal model, the reduced subproblem sizes of OA-FR are mainly in the discrete variables. The numbers of the discrete variables in the primal model are almost 10 times the number of the discrete variables in the reduced subproblems of OA-FR. Therefore, the potential advantage for OA-FR is to be able to solve stochastic models with more scenarios. Using OA-FR solution method, we successfully solve all the stochastic cases. Compared with DICOPT directly solving the whole model, OA-FR greatly reduces the solution time, on average 60.00% and as high as 93.46%, as seen in Table 7. On the other hand, the objective values obtained through OA-FR are somewhat lower than the objective values of DICOPT, with an average gap of 0.56% and a maximum of 1.52%.

5.2.2 Yield Uncertainty

The problem size and solution statistics of the stochastic scheduling model with uncertain yields is given in Table 8. As for the cases of demand uncertainties, the problem sizes of the stochastic yield cases are also about 5 times of the deterministic yield case. The solution time of the stochastic models increases 30 times compared to the deterministic case. The solution times are almost within

1000 seconds except Case 8. The computational times do not increase greatly by increasing the number of scenarios. Comparing Table 8 with Table 6, the yield uncertainties seem to require lower computational time than the demand uncertainties. The case with the deterministic yield obtains larger net profit than the cases with the stochastic yields, which means uncertain yields of side-products lead to lower output of the final oil products compared with positive deterministic yields.

Table 8. Statistics for Cases 11-20 with uncertainties of yield

Case	H(h)	Num. Scenes	Bin. Var.	Cont. Var.	Const.	Iter.	Obj. Value ($)	CPU (s)
11	72	0	640	5304	6359	3	39593.71	5.24
12	72	3	1920	16135	19180	5	39526.49	111.81
13	72	5	3200	26891	31966	5	39533.96	245.37
14	72	4	2560	21513	25593	3	39485.07	127.21
15	72	5	3200	26891	31991	3	39478.59	148.46
16	96	4	3072	25773	30793	3	45336.39	553.13
17	96	5	3840	32216	38521	3	45341.70	757.22
18	96	5	3840	32216	38521	4	45315.79	1063.60
19	96	5*	3840	32216	38521	3	45392.05	732.16
20	96	6*	4608	38659	46225	3	45350.75	871.43

Table 9 shows the statistics of OA-FR solving the stochastic Cases 12-20 compared to DICOPT. Through decomposition, there is 1/10 discrete variables in the reduced subproblems of OA-FR. The results of Cases 13, 14 and15 solving by OA-FR outperform the results when the whole models are solved directly by DICOPT. The worst case of the objective function value in OA-FR is Case 16, which is 1.02% relative deviation from the objective value of the whole model. The CPU times of OA-FR are significantly lower than the CPU times solving the whole model directly, with a maximum reduction of relative ratio of 71.74%. The increased relative ratio of computation time for OA-FR is 2-4 times from scheduling horizon 72 hours to 96 hours, compared with 5-10 times of OA solving the whole models by DICOPT.

Table 9. Comparison results of OA-FR algorithm with DICOPT for Cases 12-20

Case	*Bin. Var.	*Cont. Var.	*Const.	OA-FR		DICOPT		Gap	Saving
				Obj. ($)	CPU (s)	Obj. ($)	CPU (s)	Obj. %	CPU %
12	320	27122	25608	39490.90	**68.31**	39526.49	111.81	0.09	38.91
13	320	47904	44507	39548.12	**116.20**	39533.96	245.37	0.04	52.64
14	320	37513	35079	39500.37	**75.13**	39485.07	127.21	0.04	40.94
15	320	47904	44532	39490.51	**130.36**	39478.59	148.46	0.03	12.19
16	384	44809	42077	44874.20	**156.33**	45336.39	553.13	**1.02**	**71.74**
17	384	57213	53408	45204.83	**347.63**	45341.70	757.22	0.30	54.09
18	384	57213	53438	45159.69	**357.10**	45315.79	1063.60	0.34	66.43
19	384	35289	52868	45042.33	**209.86**	45392.05	732.16	0.77	71.34
20	384	43331	64091	45270.97	**521.19**	45350.75	871.43	0.18	40.19

*Statistics numbers of variable and constraints are only problem size of the last subproblem in OA-FR.

5.2.3 Simultaneous Uncertainties of Demand and Yield

The computational results for the simultaneous demand and yield uncertainties are shown in Table 10. For the structures of the scenario tree in the simultaneous cases are same to the former cases, the solution time of the simultaneous cases are similar to the solution time of the demand and yield uncertainties.

Table 10. Statistics for Cases 21-25 with simultaneous uncertainties of demand and yield

Case	H(h)	Num. Scenes	Bin. Var.	Cont. Var.	Const.	Iter.	CPU (s)	Obj. Value ($)
21	72	3	1920	16135	19180	4	70.25	39870.22
22	72	4	2560	21513	25573	4	88.45	39865.12
23	72	5	3200	26891	31966	5	249.09	39682.01
24*	72	5	3200	26891	31966	4	304.39	39688.43
25*	72	5	3200	26891	31966	4	226.26	39692.87

*There are two orders and one yield of side-product uncertainties in Case 24. In Case 25, there is one orders and two yields of side-product uncertainties.

Table 11 displays the results of OA-FR for the simultaneous uncertain cases. For the simultaneous uncertainties of demand and product yields, OA-FR still shows advantages in solution time since 37.34% is saved on average. The relative deviations of objective function value for OA-FR are within 1.00%, compared with the primal OBJ values. For Case 22, the obtained objective function value of OA-FR is larger than the primal OBJ value. The proposed primal model is nonconvex, it is possible that the results of OA-FR method outperform the results of the primal model.

Table 11. Comparison results of OA-FR algorithm with DICOPT for Cases 21-25

Case	*Bin. Var.	*Cont. Var.	*Const.	OA-FR		DCIOPT		Gap OBJ	Save CPU
				Obj. ($)	CPU (s)	Obj. ($)	CPU (s)	%	%
21	320	27122	25608	39813.53	**49.26**	39870.22	70.25	0.14	29.88
22	320	37513	35059	39895.04	**85.21**	39865.12	88.45	0.08	3.66
23	320	29643	44032	39290.18	**118.49**	39682.01	249.09	0.99	52.43
24	320	47904	44507	39593.89	**120.51**	39688.43	304.39	**0.24**	**60.41**
25	640	43852	41799	39588.10	**135.00**	39692.87	226.26	0.26	40.33

*Statistics numbers of variable and constraints are only problem size of the last subproblem in OA-FR.

Based on all the computational results, the OA-FR solution method may be a good choice for stochastic programming models, which are decomposable by scenarios. We can consider that design OA-FR for more general stochastic programming models.

5.2.4 Stochastic scheduling results and implications

Analyzing the scheduling results obtained from the proposed stochastic model, we simultaneously obtain multiple scheduling solutions under respective scenarios and expected net profit. Compared with the deterministic method, the stochastic method is able to illustrate different scheduling solutions and marginal benefit analysis.

The scheduling solution of Scenario 2 in Case 2 is shown in Gantt chart in Figure 6. The data above the line states the number of processing task, and the data below the line states the flow amount in the Gantt chart. The makespan of the refinery scheduling is 72 hours for all processes. Except individual flow changes, most of the tasks processed on the refinery units almost keep same flow amount, which is easy to be implemented in real production and management. Moreover, the starting and finishing times for all tasks are same. The delivery scheduling for the custom orders in Case 2 of Scenario 2 is given in Figure 7. It can be seen that Orders 2, 3, 4 are delivered in the 2 or 3 batches from the storage tanks. There are the penalties for the delivery time for Orders 2, 3, 4, for the delivery time of Orders 2, 3, 4 is out of the required time windows.

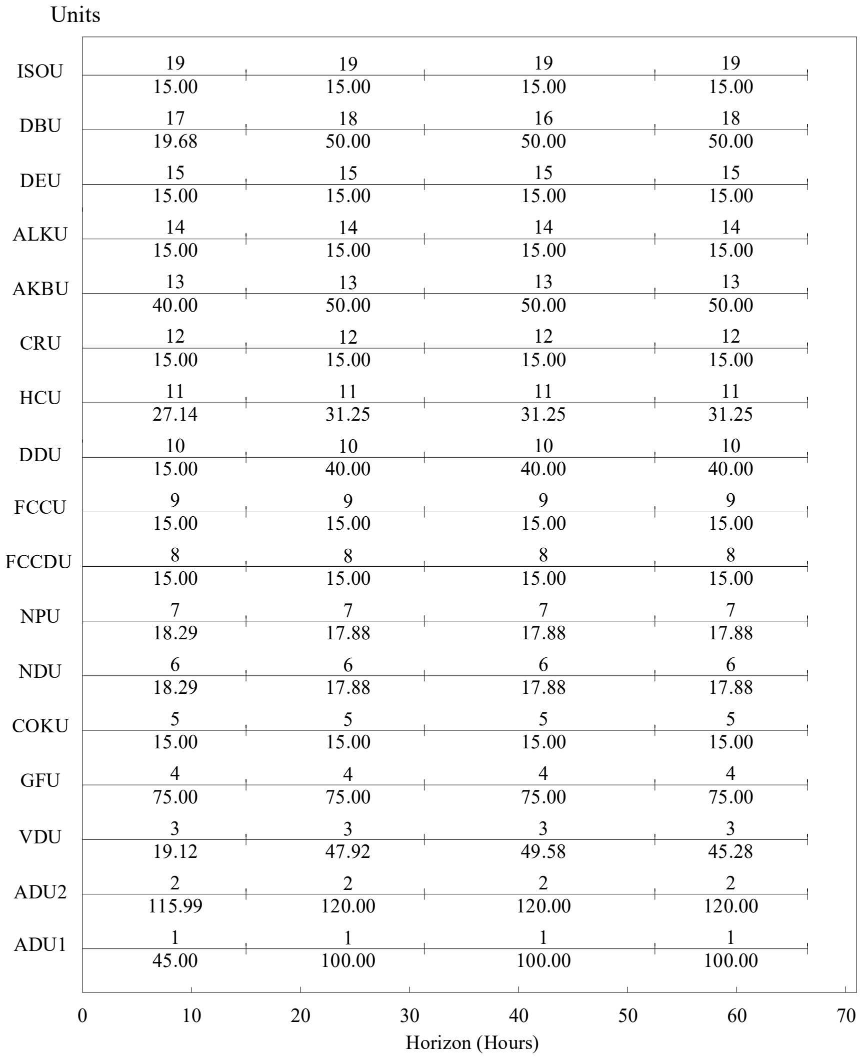

Figure 6. Scheduling solution of Gantt chart for scenario 2 in Case 2

For real applications, the stochastic scheduling solution and expected net profit can be referred as the implemented schedule and as the profit margin. The multiple combination of the stochastic scheduling solutions under different scenarios provides alternative options. Furthermore, a sensitivity analysis for the scheduling solutions can be performed to analyze more uncertainties.

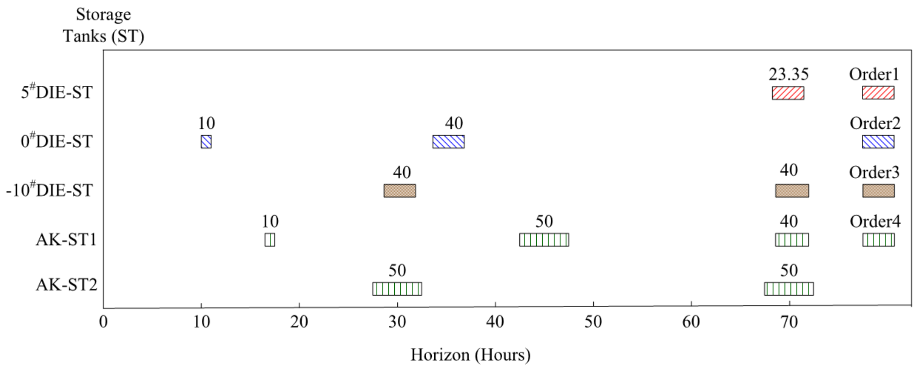

Figure 7. Gantt chart of order delivery scheduling for scenario 2 in Case 2

6. Conclusions

This paper has addressed the stochastic scheduling for the whole refinery processes with uncertainties of demand and product yields. The scenario trees are used to represent the uncertainties. We define the two-stage stochastic scheduling problem for refinery-wide processes. A hybrid mathematical model of MINLP and GDP has been proposed to describe the stochastic scheduling problem based on continuous-time representation, which is decomposable by scenarios. Aiming at solving real large-scale instances, we developed an OA solution method based on the Fix-and-Relax strategy, which decomposes the proposed model into subproblems with the smaller problem size. By fixing and relaxing the discrete variables in the sub-problems, the number of discrete variables in the generated subproblems is reduced, which makes the improved OA algorithm computationally more effective. The decomposed subproblems in OA-FR keep all the constraints in the primal model, which guarantee that the obtained solutions are globally consistent and feasible. The results of numerical experiments show the validity and efficiency of the proposed model and solution method. Future work will focus on refinery scheduling with consideration of more complex uncertainties and global optimization method.

Acknowledgments

This research was supported by the Major Program of National Natural Science Foundation of China (72192830, 72192831), the National Natural Science Foundation of China (72072029) and the 111 Project (B16009). The authors also acknowledge the Center for Advanced Process Decision-making at Carnegie Mellon University for financial support.

References

Apap, R. M., Grossmann, I. E. (2017). Models and computational strategies for multistage stochastic programming under endogenous and exogenous uncertainties. *Comput. Chem. Eng.,* https://doi.org/10.1016/j.compchemeng.2016.11.011

Brook, A., Kendrick, D., Meeraus, A. (1988). GAMS, a user's guide. ACM SIGNUM Newsl. doi:10.1145/58859.58863

Cadarso, L., Escudero, L. F., Marín, A. (2018). On strategic multistage operational two-stage stochastic 0-1 optimization for the rapid transit network design problem. *European Journal of Operational Research,* doi:10.1016/j.ejor.2018.05.041

Calfa, B. A., Agarwal, A., Grossmann, I. E., Wassick, J. M. (2014). Data-driven multi-stage scenario tree generation via statistical property and distribution matching. *Comput. Chem. Eng.,* https://doi.org/10.1016/j.compchemeng.2014.04.012

Drud, A. S. (1994). CONOPT -- A Large-Scale GRG Code. *ORSA. J. Comput.,* doi:10.1287/ijoc.6.2.207

Duran, M. A., Grossmann, I. E. (1986). An outer-approximation algorithm for a class of mixed-integer nonlinear programs. *Math. Program.,* http://dx.doi.org/10. 1007/BF02592064.

Engell, S., Maerkert, A., Sand, G., Schultz, R. (2006). Aggregated scheduling of a multiproduct batch plant by two-stage stochastic integer programming. *Optim. Eng.*, 5: 335-359.

Escudero, L. F., Salmeron, J. (2005). On a fix-and-relax framework for a class of project scheduling problems. *Annals of Operations Research,* 140(1): 163-188.

Goel, V., Grossmann, I. E. (2006). A Class of stochastic programs with decision dependent uncertainty. *Math. Program.*, Ser. B. doi:10.1007/s10107-006-0715-7

Grossmann, I. E. (2002). Review of nonlinear mixed-integer and disjunctive programming techniques. *Optim. Eng.*, 3, 227-252, http://dx.doi.org/10.1023/A:1021039126272.

Grossmann, I. E., Trespalacios, F. (2013). Systematic Modeling of Discrete-Continuous Optimization Models through Generalized Disjunctive Programming. *AIChE. J.,* https://doi.org/10.1002/ aic.14088

Grossmann, I. E., Viswanathan, J., Vecchietti, A., Raman, R., Kalvelagen, E. (2002). *GAMS/ DICOPT: A Discrete Continuous Optimization Package*.

Guillén, G., Mele, F. D., Espuña, A., Puigjaner, L. (2006). Addressing the scheduling of chemical supply chains under demand uncertainty. *Ind. Eng. Chem. Res.,* https://doi.org/10.1021 /ie051424a

Guo, W. H., Xu, J. L. (2004). Energy consumption evaluation index and comparison of refinery. *Proceedings of refining energy-saving technology for Sinopec*. (Chinese)

Høyland, K., Wallace, S. W. (2001). Generating Scenario Trees for Multistage Decision Problems. *Management Science,* 46(2):295-307.

Ierapetritou, M. G., Floudas, C. A. (1998). Effective Continuous-Time Formulation for Short-Term Scheduling. 2. Continuous and Semicontinuous Processes. *Ind. Eng. Chem. Res.,* doi: 10.1021/ie970928

IBM CPLEX Optimization Studio V12.8. http://www-01.ibm.com/support/docview. wss?uid=swg 27050618 (accessed 10.3.18).

Jia, Z., Ierapetritou, M. (2004). Efficient short-term scheduling of refinery operations based on a continuous time formulation. *Comput. Chem. Eng.,* doi:10.1016/j.compchemeng.2003.09.007

Joly, M., Moro, L .F. L., Pinto, J. M. (2002). Planning and scheduling for petroleum refineries using mathematical programming. *Braz. J. Chem. Eng*., doi: 10.1590/S0104-66322002000200008

Méndez, C. A., Cerdá, J., Grossmann, I. E., Harjunkoski, I., Fahl, M. (2006). State-of-the-art review of optimization methods for short-term scheduling of batch processes. *Comput. Chem. Eng.*, 30, 913-946. doi:10.1016/j.compchemeng.2006.02.008

Pinto, V. T., Barbosa, P. A. P. F. D., Novais, A. Q. (2009). Design and scheduling of periodic multipurpose batch plants under uncertainty. *Ind. Eng. Chem. Res.*, doi:10.1021/ie900137p

Ruszczyński, A. (1997). Decomposition methods in stochastic programming. *Math. Program.*, 79: 333-353.

Shah, N. K., Ierapetritou, M. G. (2011). Short-term scheduling of a large-scale oil-refinery operations: Incorporating logistics details. *AIChE J.*, doi:10.1002/aic.12359

Shah, N., Saharidis, G. K. D., Jia, Z. Y., Ierapetritou, M. G. (2009). Centralized–decentralized optimization for refinery scheduling. *Comput. Chem. Eng.*, doi:10.1016/j.compchemeng. 2009.06.010

Su, L. J., Tang, L. X., Grossmann, I. E. (2015). Computational Strategies for Improved MINLP Algorithms. *Comput. Chem. Eng.*, doi:10.1016/j.compchemeng.2015.01.015

Su, L. J., Tang, L. X., Bernal, E. D., Grossmann, I. E. (2018). Improved quadratic cuts for convex mixed-integer nonlinear programs. *Comput. Chem. Eng.*, doi:10.1016/j.compchemeng.2017.10.011

Zhao, H., Rong, G., Feng, Y. (2014). Multiperiod planning model for integrated optimization of a refinery production and utility system. *Ind. Eng. Chem. Res.*, doi:10.1021/ie502717e