Chinese Journal of Chemical Engineering 69 (2024) 152–166

Contents lists available at ScienceDirect

Chinese Journal of Chemical Engineering

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

Full Length Article

Data-driven Wasserstein distributionally robust chance-constrained optimization for crude oil scheduling under uncertainty

Xin Dai [1], Liang Zhao [1], Renchu He [1,*], Wenli Du [1], Weimin Zhong [1], Zhi Li [1,2], Feng Qian [1,*]

[1] Key Laboratory of Smart Manufacturing in Energy Chemical Process, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China
[2] Engineering Research Center of Process System Engineering, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China

ARTICLE INFO

Article history:
Received 18 January 2023
Received in revised form
8 December 2023
Accepted 12 December 2023
Available online 6 January 2024

Keywords:
Distributions
Model
Optimization
Crude oil scheduling
Wasserstein distance
Distributionally robust chance constraints

ABSTRACT

Crude oil scheduling optimization is an effective method to enhance the economic benefits of oil refining. But uncertainties, including uncertain demands of crude distillation units (CDUs), might make the production plans made by the traditional deterministic optimization models infeasible. A data-driven Wasserstein distributionally robust chance-constrained (WDRCC) optimization approach is proposed in this paper to deal with demand uncertainty in crude oil scheduling. First, a new deterministic crude oil scheduling optimization model is developed as the basis of this approach. The Wasserstein distance is then used to build ambiguity sets from historical data to describe the possible realizations of probability distributions of uncertain demands. A cross-validation method is advanced to choose suitable radii for these ambiguity sets. The deterministic model is reformulated as a WDRCC optimization model for crude oil scheduling to guarantee the demand constraints hold with a desired high probability even in the worst situation in ambiguity sets. The proposed WDRCC model is transferred into an equivalent conditional value-at-risk representation and further derived as a mixed-integer nonlinear programming counterpart. Industrial case studies from a real-world refinery are conducted to show the effectiveness of the proposed method. Out-of-sample tests demonstrate that the solution of the WDRCC model is more robust than those of the deterministic model and the chance-constrained model.

1. Introduction

Crude oil scheduling optimization can significantly reduce the production cost of a refinery by coordinating crude oil unloading, storage, blending, and processing, which is essential in the context of rising oil prices [1,2]. Hence, the field has attracted considerable research attention. A discrete-time mixed-integer linear programming (MILP) model has been developed by Lee *et al.* [3] to minimize crude oil operation costs. Pinto *et al.* [4] have also constructed a MILP model to reduce the number of used tanks while pursuing economic efficacy. A discrete-time mixed-integer nonlinear programming (MINLP) model has been proposed by Li *et al.* [5] with a more accurate description of crude oil mixing in storage tanks. The MINLP model of Reddy *et al.* [6] has considered more elements in real-world refineries, such as one single-buoy mooring (SBM) station, jetties, and very large crude carriers (VLCCs). Continuous-time scheduling models have also been proposed to reduce problem sizes with variable-length time slots [7,8]. Zimberg *et al.* [9] have extended the traditional continuous time optimization method to handle crude oil scheduling optimization problems in a refinery with a main pipeline. An efficient rolling horizon algorithm with variable fixing strategy and constraint pruning strategy have also been proposed by Yang *et al.* [10] to solve large size crude oil scheduling optimization problem. Su *et al.* [11] have further studied the integrated optimization problem between crude oil scheduling and refinery planning. They have designed a novel hierarchical hybrid continuous-discrete time representation for that problem. The above models are formulated based on deterministic parameters. But uncertainties like demand fluctuations are unavoidable in real-world practice and will affect the effectiveness and feasibility of the solutions.

* Corresponding authors.
E-mail addresses: renchuhe@ecust.edu.cn (R. He), fqian@ecust.edu.cn (F. Qian).

https://doi.org/10.1016/j.cjche.2023.12.007

Various approaches have been developed for the optimization problems under uncertainty [12]. Gupta and Zhang [13] have proposed a robust optimization model for crude oil inventory management under ship arrival uncertainty, where the uncertainty has been assumed to be subject to the normal distribution. A production plan optimization model under uncertainty has been built by Li *et al.* [14] with the uniform distribution assumption. Li *et al.* [15] have considered more symmetrical distribution types of uncertainties and derived different robust counterparts. Luo and Rong [16] have proposed an integrated robust optimization method to handle uncertain parameters with both continuous and discrete probability distributions. A preventive optimization method based on structure adapted genetic algorithm has been proposed by Panda and Ramteke [17] to deal with crude demand uncertainty with the assumption about the fluctuation range of uncertainties. Dai *et al.* [18] have proposed a data-driven robust optimization method, in which the information of uncertainties can be automatically extracted by the support vector clustering method and be used to guild optimization under product yield uncertainty. However, the above methods lack the ability to control the feasibility probability of optimal solutions. The chance-constrained programming method provides an opportunity to directly tune the minimum probability that the stochastic constraint is satisfied [19]. Wang *et al.* [20] have proposed a chance-constrained model under demand uncertainty. The model has been extended by Cao *et al.*[21] using fuzzy programming technology to handle uncertain demands with more general continuous probability distributions. Cao *et al.* [22] have also constructed a chance-constrained model based on the random simulation method where uncertain parameters have been believed to follow the discrete probability distributions generated by sampled data. Assumptions about probability distributions of uncertain parameters are necessary for these chance-constrained methods, while the true probability distributions may not be available a priori [23]. Deviations between the assumed and true probability distributions may reduce the reliability of the models.

In recent years, data-driven distributionally robust optimization (DRO) has been an appealing technology [24]. Instead of one single presumed probability distribution, a family of possible probability distributions, which are incorporated in the data-driven ambiguity sets, will be considered in the DRO models to optimize the worst expected performance [25]. Moment-based ambiguity sets have been established under given mean and covariance and applied in the reserve scheduling problem in the wind power system [26,27]. A DRO framework has been put forward by Shang and You [28] with ambiguity sets, which have been constructed by the principal component analysis and first-order deviation moment functions. Ambiguity sets based on the Wasserstein distance have been proved to cover the true probability distribution of uncertainty with high probability [24,29]. Ning and You [30] have proposed a DRO model for agricultural waste-to-energy network design under uncertainty with Wasserstein ambiguity sets. A Wasserstein DRO strategic energy planning optimization has been built by Guevara *et al.* [31], in which a machine learning method has been used to select important uncertain parameters. Liu *et al.* [32] have proposed a two-stage DRO model for maritime inventory routing.

Distributionally robust chance-constrained (DRCC) method has been developed by treating a chance-constrained problem as an expectation one with an indicator function [33]. The method allows the chance constraints to hold under the worst-case probability distributions within the ambiguity sets [34]. A DRCC model has been built by Zhou *et al.* [35] with a moment-based box-type ambiguity set. The model has been reformulated as a standard second-order cone programming problem *via* the dual transformation. Some strategies have been proposed to approximately tackle DRCC models with Wasserstein ambiguity sets. A bicriteria outer approximation approach has been advanced by Xie and Ahmed [36] to build a convex relaxation of the DRCCs. An inner approximation has been developed by Ordoudis *et al.* [37,38] using a conditional value-at-risk (CVaR) approximation and the Bonferroni inequality. Recently, Xie [29] has derived an exact reformulation during which an equivalent CVaR representation of the Wasserstein distributionally robust chance-constrained (WDRCC) problem has been obtained to get a more effective result. DRCC approaches have been widely applied in the optimization of power systems [39–41]. But there is no relevant work about DRCC crude oil scheduling optimization under uncertainty, to the best of our knowledge.

In this study, a data-driven WDRCC optimization method for crude oil scheduling is put forward to hedge against demand uncertainty. The ambiguity sets are constructed by demand data of CDUs on the basis of the Wasserstein distance without assumptions about the probability distribution type. The radii of these sets are chosen through a cross-validation method. A new deterministic optimization model is presented. The WDRCC model is further built with the consideration of the worst possible probability distributions in the Wasserstein ambiguity sets. It is then reformulated as an MINLP model by introducing the conditional value-at-risk representation. The benefits of the proposed method are demonstrated by real-world case studies.

The major contributions of this paper are as follows.

(1) A novel data-driven WDRCC optimization method is developed from a new deterministic optimization model to schedule crude operations under demand uncertainty.
(2) Historical data of crude demands of CDUs are employed to establish ambiguity sets based on the Wasserstein distance.
(3) A cross-validation approach is utilized to select the radii of these ambiguity sets.
(4) An equivalent CVaR representation is derived to solve the DRCC optimization model.

The rest of this paper is organized as follows. The crude oil scheduling problem is stated in the next section. Then, the deterministic optimization model of crude oil scheduling is formulated in Section 3, followed by the chance-constrained model in Section 4. The WDRCC model will be derived in Section 5. In Section 6, case studies are provided. Finally, concluding remarks in Section 7.

2. Problem Statement

Fig. 1 exhibits the schematic of a typical marine-access refinery. The refinery receives crude oil from single-parcel vessels and multiple-parcel VLCCs. VLCCs can only unload oil into tanks *via* the SBM station. The single-parcel vessels are relatively small and unload oil at the onshore jetties. Tanks serve both as storage tanks and charging tanks. It means that different types of crude oil will be blended inside tanks and fed directly into CDUs for processing. There are flow rate constraints on the crude inputs of CDUs. And the total supply of crude oil during the whole scheduling horizon should also meet the demands of CDUs.

The optimization problem of crude oil scheduling in this work is stated as follows:

Given information:

(1) Shipping information: Crude parcel types and parcel sizes in vessels, their estimated arrival times, the unloading sequence of parcels and the number of jetties.
(2) Tank information: Number of tanks, their capacity limits, and the initial crude weights and compositions in tanks.
(3) CDU information: Number of CDUs, the lower and upper bounds of amounts, compositions and properties of the

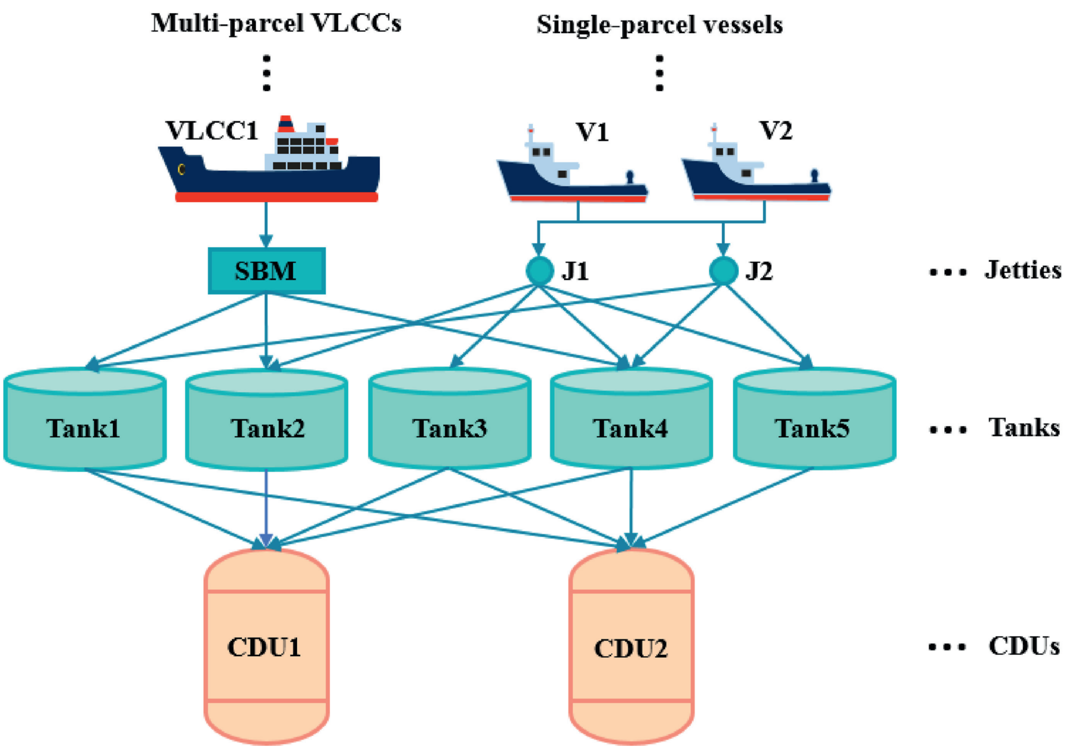

Fig. 1. Schematic of crude oil scheduling.

crude inputs of each CDU per period, the lower and upper bounds of amounts of products produced by each CDU per period, the demand data of CDUs.
(4) Connection information: The allowable connections among parcels, tanks, and CDUs and the transfer rate limits on each connection.
(5) Crude data: Number of crude types, properties, and product yields of each crude.
(6) Economic data: Crude prices, the demurrage cost of each vessel, and the cost of each changeover.

Determine:

(1) An unloading schedule of passels such as the receiving tanks, unloading times, and transfer amounts;
(2) Inventory and composition profiles of tanks at the end of each time period;
(3) Feed tanks, the amounts of oil they provide for CDUs and the charging times.

Assumptions and operating practices:

(1) There is only one SBM station, for which only one VLCC can unload oil at any moment.
(2) Crude holdup of the SBM/jetty line is negligible.
(3) For any tank, simultaneous feeding and receiving oil is not permitted.
(4) After receiving crude, a tank will settle for at least one period to remove brine.
(5) A tank can feed maximum two CDUs in any period.
(6) Operations of CDUs are continuous.
(7) The uncertain demand parameters of CDUs are considered in this study and assumed to be independent of each other.

The scheduling objective is the minimization of the total production cost.

3. Deterministic Mathematical Formulation

The model of Reddy *et al.* [6] takes into account the crude oil scheduling rules in a comprehensive manner. The model is further modified and improved in this section. The new deterministic optimization model serves as the basis of the proposed data-driven DRCC optimization model. The following constraints are included.

3.1. Unloading of parcels

Each parcel should unload oil continuously as shown in Eq. (1).

$$XP_{p,t} = XP_{p,t-1} + XF_{p,t} - XL_{p,t-1} \quad \forall (p,t) \in PT \tag{1}$$

where 0–1 continuous variable $XP_{p,t} \in [0,1]$ is equal to 1 if parcel p connects to the SBM/jetty discharge line in period t and 0 otherwise, 0–1 continuous variable $XF_{p,t} \in [0,1]$ is equal to 1 if parcel p first connects to a discharge line in period t and 0 otherwise, 0–1 continuous variable $XL_{p,t} \in [0,1]$ is equal to 1 if the parcel p disconnects the unloading facilities in period t and 0 otherwise, and $PT = \{(p,t)|$ parcel p can connect to a discharge line in period $t\}$.

Eq. (2) means that parcel p supplies a tank in period t if it finishes unloading in that period. Eq. (3) illustrates that a parcel connects the unloading facilities only once during the whole scheduling horizon.

$$XP_{p,t} \geq XL_{p,t} \quad \forall (p,t) \in PT \tag{2}$$

$$\sum_{t:(p,t) \in PT} XF_{p,t} = 1 \quad \forall p \tag{3a}$$

$$\sum_{t:(p,t)\in PT} XL_{p,t} = 1 \quad \forall p \tag{3b}$$

The first connecting time TF_p and the disconnecting time TL_p of parcel p can be obtained by $XF_{p,t}$ and $XL_{p,t}$ using Eqs. (4) and (5). ETA_p is the estimated arrival time of parcel p. Parcel p cannot unload oil until it arrives at the refinery as shown in Eq. (6).

$$TF_p = \sum_{t:(p,t)\in PT} (t-1)\cdot XF_{p,t} \quad \forall p \tag{4}$$

$$TL_p = \sum_{t:(p,t)\in PT} t\cdot XL_{p,t} \quad \forall p \tag{5}$$

$$TF_p \geq ETA_p \quad \forall p \tag{6}$$

SP is the set of VLCC parcels and JP is the set of jetty parcels. VLCC parcels should dock at the SBM station one by one as shown in Eq. (7). At most two VLCC parcels can unload oil in one period, which is realized by Eq. (8). Eq. (9) requires that at most NJ single-parcel vessels can berth at the same time if there are NJ jetties.

$$TF_{p+1} \geq TL_p - 1 \quad p\in SP, \quad p+1\in SP \tag{7}$$

$$\sum_{\substack{p:(p,t)\in PT,\\ p\in SP}} XP_{p,t} \leq 2 \quad \forall t \tag{8}$$

$$\sum_{\substack{p:(p,t)\in PT,\\ p\in JP}} XP_{p,t} \leq NJ \quad \forall t \tag{9}$$

In Reddy's work, if parcel p and tank i are connected to the SBM line in period t, then parcel p must have unloaded oil to tank i in that period [6]. The logic does not make sense because they may not be connected to the SBM line at the same time within the period t. Thus, binary variable $X_{p,i,t}\in\{0,1\}$ is used in this work to accurately denote the connections between parcels and tanks. $X_{p,i,t}$ will be 1 if parcel p unloads oil into tank i in period t and 0 otherwise. 0–1 continuous variable $XT_{i,t}\in[0,1]$ will be 1 if oil is transmitted into tank i in period t and 0 otherwise. They have the following constraints Eqs. (10) and (13), where $PI = \{(p,i)|$ tank i can receive oil from parcel $p\}$. Although $XP_{p,t}$, $XF_{p,t}$, $XL_{p,t}$, $XT_{i,t}$ are defined as 0–1 continuous variables, they can be binary automatically according to constraints Eqs. (1)–(3) and Eqs. (9) and (10). The introduction of 0–1 continuous variables is to avoid introducing unnecessary binary variables and to reduce the computational burden for the optimization model.

$$XP_{p,t} \geq X_{p,i,t} \quad \forall (p,t)\in PT, (p,i)\in PI \tag{10}$$

$$XP_{p,t} \leq \sum_{i:(p,i)\in PI} X_{p,i,t} \quad \forall (p,t)\in PT \tag{11}$$

$$XT_{i,t} \geq X_{p,i,t} \quad \forall (p,t)\in PT, (p,i)\in PI \tag{12}$$

$$XT_{i,t} \leq \sum_{\substack{p:(p,t)\in PT,\\ (p,i)\in PI}} X_{p,i,t} \quad \forall i,t \tag{13}$$

The constraints on the maximum number of connections between tanks and parcels within each period are as follows:

$$\sum_{i:(p,i)\in PI} X_{p,i,t} \leq 2 \quad \forall (p,t)\in PT \tag{14}$$

$$\sum_{\substack{p:(p,t)\in PT,\\ (p,i)\in PI}} X_{p,i,t} \leq 2 \quad \forall i,t \tag{15}$$

$$\sum_{\substack{p,i:(p,t)\in PT,\\ (p,i)\in PI, p\in SP}} X_{p,i,t} \leq 2 \quad \forall t \tag{16}$$

$$\sum_{\substack{p,i:(p,t)\in PT,\\ (p,i)\in PI, p\in JP}} X_{p,i,t} \leq 2\cdot NJ \quad \forall t \tag{17}$$

$FPT_{p,i,t}$ is defined as the unloading amount from parcel p into tank i in period t. It should meet the following requirements.

$$FPT^{\mathrm{L}}_{p,i}\cdot X_{p,i,t} \leq FPT_{p,i,t} \leq FPT^{\mathrm{U}}_{p,i}\cdot X_{p,i,t} \quad \forall (p,t)\in PT, (p,i)\in PI \tag{18}$$

$$\sum_{\substack{p:(p,t)\in PT\\ p\in SP}} \sum_{i:(p,i)\in PI} \frac{FPT_{p,i,t}}{FPT^{\mathrm{U}}_{p,i}} \leq 1 \quad \forall t \tag{19}$$

$$\sum_{i:(p,i)\in PI} \frac{FPT_{p,i,t}}{FPT^{\mathrm{U}}_{p,i}} \leq 1 \quad \forall (p,t)\in PT, p\in JP \tag{20}$$

$$\sum_{\substack{p:(p,t)\in PT,\\ (p,i)\in PI, p\in SP}} \frac{FPT_{p,i,t}}{FPT^{\mathrm{U}}_{p,i}} + \sum_{\substack{p:(p,t)\in PT,\\ (p,i)\in PI, p\in JP}} \frac{FPT_{p,i,t}}{FPT^{\mathrm{U}}_{p,i}} \leq 1 \quad \forall i,t \tag{21}$$

where $FPT^{\mathrm{L}}{}_{p,i}$ and $FPT^{\mathrm{U}}{}_{p,i}$ represent the lower and upper bounds of the unloading rate. All oil in parcel p should be unloaded as shown in Eq. (22), where PS_p denotes the mass of oil in parcel p.

$$\sum_{\substack{i,t:(p,t)\in PT,\\ (p,i)\in PI}} FPT_{p,i,t} = PS_p \quad \forall p \tag{22}$$

$$\begin{aligned} WCT_{i,c,t} = WCT_{i,c,t-1} + \sum_{\substack{p:(p,c)\in PC,\\ (p,t)\in PT}} FPT_{p,i,t} - \\ \sum_{d:(i,d)\in ID,} FCTU_{i,d,c,t} \quad \forall (i,c)\in IC, t \end{aligned} \tag{23}$$

3.2. Inventory balance in crude tanks

The inventory balance in a crude tank is presented as follows:

$$W_{i,t} = \sum_{c:(i,c)\in IC} WCT_{i,c,t} \quad \forall i,t \tag{24}$$

$$W^{\mathrm{L}}_i \leq W_{i,t} \leq W^{\mathrm{U}}_i \quad \forall i,t \tag{25}$$

where $WCT_{i,c,t}$ denotes the amount of crude c in tank i at the end of period t, $W_{i,t}$ denotes the total amount of crude in tank i at the end of period t, W^{L}_i and W^{U}_i are the corresponding lower and upper capacity bounds, $FCTU_{i,d,c,t}$ represents the amount of crude c

transferred to CDU d from tank i in period t, $PC = \{(p,c)|$ parcel p holds crude $c\}$, $ID = \{(i,d)|$ CDU d can receive oil from tank $i\}$, and $IC = \{(i,c)|$ crude c can be stored in tank i $\}$.

3.3. Crude oil supply for CDUs

The amount of oil pumped into CDU d from tank i during period t is represented by $FTU_{i,d,t}$, the bounds of which are denoted by $FTU^{L}_{i,d}$ and $FTU^{U}_{i,d}$. The relationships between $FTU_{i,d,t}$. and $FCTU_{i,d,c,t}$ are expressed as follows:

$$FTU_{i,d,t} = \sum_{c:(i,c)\in IC} FCTU_{i,d,c,t} \quad \forall (i,d)\in ID, t \tag{26}$$

$$Y_{i,d,t} \cdot FTU^{L}_{i,d} \le FTU_{i,d,t} \le Y_{i,d,t} \cdot FTU^{U}_{i,d} \quad \forall (i,d)\in ID, t \tag{27}$$

$$WCT_{i,c,t} = f_{i,c,t} \cdot W_{i,t} \quad \forall (i,c)\in IC, t \tag{28}$$

$$FCTU_{i,d,c,t} = f_{i,c,t-1} \cdot FTU_{i,d,t} \quad \forall (i,c)\in IC, (i,d)\in ID, t \tag{29}$$

where the binary variable $Y_{i,d,t}\in\{0,1\}$ is equal to 1 if tank i supplies CDU d in period t and 0 otherwise, and $f_{i,c,t}$ is the fraction of crude c in tank i at the end of period t. Eqs. (28) and (29) means that the composition of the mixed oil outputted from the tank i should be consistent with that in the tank. At any period, one tank can supply at most two CDUs. A tank cannot feed CDUs if it is receiving oil, after which one period is needed for brine settling. At most NT_d tanks can charge CDU d in one period.

$$\sum_{d:(i,d)\in ID} Y_{i,d,t} \le 2 \quad \forall i, t \tag{30}$$

$$2XT_{i,t} + Y_{i,d,t} + Y_{i,d,t+1} \le 2 \quad \forall (i,d)\in ID, t \tag{31}$$

$$\sum_{i:(i,d)\in ID} Y_{i,d,t} \le NT_d \quad \forall d, t \tag{32}$$

If there is no change in the combination of feed tanks for one CDU, every individual feed flow of the CDU should remain constant for the stability of production. However, the relevant constraints proposed by Reddy [6] will fail to meet the above requirement if the number of feed tanks of CDU d is less than NT_d in period t even though combinations of feed tanks for one CDU are the same in period $t-1$ and period t. Therefore, novel production stability constraints Eqs. (33)–(40) are proposed in this paper to effectively realize stable production of CDUs. The rigorous theoretical explanation of the novel production stability constraints is as follows. $YY_{i,d,t}\in[0,1]$ and $CO_{d,t}\in[0,1]$ are continuous 0–1 variables. $YY_{i,d,t} = 1$ if a change of the state of connection between tank i and CDU d is found. $CO_{d,t} = 1$ if a changeover of the feed tanks for CDU d is detected. Continuous 0–1 variable $Y_{i,d,t}$ is forced to be 1 by Eq. (33) with $Y_{i,d,t} = 1$ and $Y_{i,d,t-1} = 0$. $YY_{i,d,t}$ is forced to be 1 by Eq. (34) with $Y_{i,d,t} = 0$ and $Y_{i,d,t-1} = 1$. $YY_{i,d,t}$ is forced to be 0 by Eq. (35) with $Y_{i,d,t} = 0$ and $Y_{i,d,t-1} = 0$. $YY_{i,d,t}$ is forced to be 0 by Eq. (36) with $Y_{i,d,t} = 1$ and $Y_{i,d,t-1} = 1$. Thus, $YY_{i,d,t} = 1$ according to Eqs. (33) and (34) if a change of the state of connection between tank i and CDU d is found. $YY_{i,d,t}$ is equal to 0 according to Eqs. (35) and (36) if there is no change of the state of connection between tank i and CDU d. $CO_{d,t}$ is equal to 0 by Eq. (37) if there is no changeover of the feed tanks for CDU d in time period t ($\sum_{i:(i,d)\in ID} YY_{i,d,t} = 1$). $CO_{d,t}$ is equal to 1 by Eq. (38) if a changeover of the feed tanks for CDU d is detected in time period t (some $YY_{i,d,t} = 1$). $FTU_{i,d,t}$ need to be equal to $FTU_{i,d,t-1}$ when $CO_{d,t} = 0$ according to Eqs. (39) and (40) to keep the stable production of CDU d. Eqs. (39) and (40) are not valid constraints when $CO_{d,t} = 1$. It can be seen that the proposed constraints can make the stable production requirement met regardless of the number of feed tanks of a CDU.

$$YY_{i,d,t} \ge Y_{i,d,t} - Y_{i,d,t-1} \quad \forall (i,d)\in ID, t \tag{33}$$

$$YY_{i,d,t} \ge Y_{i,d,t-1} - Y_{i,d,t} \quad \forall (i,d)\in ID, t \tag{34}$$

$$YY_{i,d,t} \le Y_{i,d,t-1} + Y_{i,d,t} \quad \forall (i,d)\in ID, t \tag{35}$$

$$YY_{i,d,t} \le 2 - Y_{i,d,t-1} - Y_{i,d,t} \quad \forall (i,d)\in ID, t \tag{36}$$

$$CO_{d,t} \le \sum_{i:(i,d)\in ID} YY_{i,d,t} \quad \forall d, t \tag{37}$$

$$CO_{d,t} \ge YY_{i,d,t} \quad \forall (i,d)\in ID, t \tag{38}$$

$$M \cdot CO_{d,t} + FTU_{i,d,t} \ge FTU_{i,d,t-1} \quad \forall (i,d)\in ID, t \tag{39}$$

$$M \cdot CO_{d,t} + FTU_{i,d,t-1} \ge FTU_{i,d,t} \quad \forall (i,d)\in ID, t \tag{40}$$

where M is a big number. The amounts, compositions, properties and product outputs of the mixture inputs for CDUs in one period are ensured by using

$$FU_{d,t} = \sum_{i:(i,d)\in ID} FTU_{i,d,t} \quad \forall d, t \tag{41}$$

$$FU^{L}_{d} \le FU_{d,t} \le FU^{U}_{d} \quad \forall d, t \tag{42}$$

$$FCTUS_{d,c,t} = \sum_{i:(i,c)\in IC,(i,d)\in ID} FCTU_{i,d,c,t} \quad \forall d, c, t \tag{43}$$

$$FU_{d,t} \cdot xc^{L}_{c,d} \le FCTUS_{d,c,t} \le FU_{d,t} \cdot xc^{U}_{c,d} \quad \forall d, c, t \tag{44}$$

$$FU_{d,t} \cdot xk^{L}_{k,d} \le \sum_{c} FCTUS_{d,c,t} \cdot xk_{k,c} \le FU_{d,t} \cdot xk^{U}_{k,d} \quad \forall d, k, t \tag{45}$$

$$PF^{L}_{d,j} \le \sum_{c} FCTUS_{d,c,t} \cdot Yield_{c,j,d} \le PF^{U}_{d,j} \quad \forall d, j, t \tag{46}$$

where $FU_{d,t}$ denotes the total amount of oil received by CDU d within period t, FU^{L}_{d} and FU^{U}_{d} are input rate limits of CDU d, $FCTUS_{d,c,t}$ is the amount of crude c in $FU_{d,t}$, $xc^{L}_{c,d}$ and $xc^{U}_{c,d}$ are composition limits, $xk_{k,c}$ is the property k of crude c, $xk^{L}_{k,d}$ and $xk^{U}_{k,d}$ are the property limits, $Yield_{c,j,d}$ is the yield of product j for crude c in CDU d, $PF^{L}_{d,j}$ and $PF^{U}_{d,j}$ are production rate limits of product j from CDU d.

The specific amount of crude oil processed by each CDU during the scheduling horizon needs to be given in advance in Reddy's model [6], which is impractical. A more common requirement in practice is that the crude supply for CDU d should be no less than its demand DM_d.

$$\sum_{t} FU_{d,t} \ge DM_d \quad \forall d \tag{47}$$

3.4. Total production cost

As shown in Eq. (49), the total production cost consists of the crude cost, vessel demurrage cost and changeover cost. $Cost_c$ is defined as the material cost of crude c. The demurrage cost of vessel v is represented by DC_v and can be calculated by Eq. (48), where ETD_v is the allowable dwell time of vessel v, SWC_v is the demurrage cost for vessel v per period, and $PV = \{(p, v) \mid$ parcel p is the last parcel of vessel v $\}$. COC is the cost per changeover.

$$DC_v \geq \left(TL_p - ETA_p - ETD_v\right) \cdot SWC_v \quad \forall (p,v) \in PV \tag{48}$$

$$\text{Total production cost} = \sum_{d,c,t} FCTU_{d,c,t} \cdot Cost_c + \sum_v DC_v + COC \cdot \sum_{d,t} CO_{d,t} \tag{49}$$

The deterministic crude oil scheduling optimization model (DCSO) in this paper is presented in the following Eq. (50).

$$\begin{aligned} \min \ & \text{Total production cost defined by Eq.(49)} \\ s.t. \ & \text{Eqs. (1) − (48)} \end{aligned} \tag{50}$$

4. Chance-constrained Crude Oil Scheduling Optimization Model

The demands of CDUs are uncertain in real world, so the risk of violation of the demand constraint due to the fluctuations should be considered. In chance-constrained optimization, constraint Eq. (47) is replaced by the following constraint Eq. (51).

$$P_d\left\{\sum_t FU_{d,t} \geq \xi_d\right\} \geq 1 - \varepsilon_d \quad \forall d \tag{51}$$

where $\xi_d \in R$ is defined as the uncertain demand of CDU d, P_d is the probability distribution of uncertain demand ξ_d, and $1-\varepsilon_d$ is the minimum required probability that the demand of CDU d is satisfied. But it is hard to obtain the probability distributions of demands in advance. A traditional method is to use an uncertain parameter $\widehat{\xi}_d \in R$ to approximate the ξ_d. The probability distribution of $\widehat{\xi}_d$ is defined as $\widehat{P}_d$, which is a discrete empirical probability and generated by the historical demand data:

$$\widehat{P}_d = \frac{1}{N_d} \sum_{m=1}^{N_d} \delta_{\widehat{\xi}_{d,m}} \quad \forall d \tag{52}$$

where $\widehat{\xi}_{d,m}$ $(m = 1,2, \cdots, N_d)$ is the m-th historical data of the demand of CDU d, and $\delta_{\widehat{\xi}_{d,m}}$ represents the Dirac distribution concentrating unit mass at $\widehat{\xi}_{d,m}$. The data are arranged in a non-increasing order such that:

$$\widehat{\xi}_{d,1} \geq \widehat{\xi}_{d,2} \geq \cdots \geq \widehat{\xi}_{d,N_d} \quad \forall d \tag{53}$$

Then, constraint Eq. (51) can be rewritten as follows:

$$\widehat{P}_d\left\{\sum_t FU_{d,t} \geq \widehat{\xi}_d\right\} \geq 1 - \varepsilon_d \quad \forall d \tag{54}$$

In order to transfer the constraint into a linear form, the $(1-\varepsilon)$ value-at-risk (VaR) is introduced [42].

$$VaR_{1-\varepsilon}(z) = \min_v\{v \mid P_Z\{z \leq v\} \geq 1 - \varepsilon\} \tag{55}$$

where z is an uncertain parameter, its probability distribution is P_Z, and v is an auxiliary variable. Constraint Eq. (43) is equivalent to:

$$\sum_t FU_{d,t} \geq DMV_d \quad \forall d \tag{56}$$

$$DMV_d = VaR_{1-\varepsilon_d}(\widehat{\xi}_d) = \widehat{\xi}_{d,q_d+1} \quad \forall d \tag{57}$$

where DMV_d is the minimum amount of oil that needs to be provided to CDU d during the scheduling horizon, $q_d = \lfloor N_d \cdot \varepsilon_d \rfloor$, and $\lfloor . \cdot \rfloor$ is the floor function. Therefore, the chance-constrained optimization model of crude oil scheduling based on empirical distribution (CCCSED) is as follows:

$$\begin{aligned} \min \ & \text{Total production cost defined by Eq.(38)} \\ s.t. \ & \text{Eqs. (1) − (46)} \\ & \text{Eqs. (48) and (56)} \end{aligned} \tag{58}$$

5. Data-driven WDRCC Model

Although the above chance-constrained approach is easy to apply, the empirical probability distribution may deviate from the true probability distribution, for which the chance constraint Eq. (51) may not be met actually. Therefore, a novel data-driven WDRCC optimization model is proposed to improve the robustness of decisions in crude scheduling.

The definition of the Wasserstein distance between probability distributions P_A and P_B is given in Eq. (59).

$$dw(P_A, P_B) = \min_{\Pi}\left\{\int_{R \times R} \|\xi_A - \xi_B\|_\beta \Pi(d\xi_A, d\xi_B)\right\} \tag{59}$$

where $dw(\cdot,\cdot)$ denotes the Wasserstein distance; $\xi_A \in R$ and $\xi_B \in R$ are uncertain parameters, whose probability distributions are P_A and P_B, respectively; the notation $\times$ denotes the Cartesian product; $\|\cdot\|_\beta$ denotes the L_β-norm; $\beta = 1$ in this paper; Π is a joint distribution of ξ_A and ξ_B with marginals P_A and P_B. The Wasserstein distance means the minimal cost to move a mass distribution described by P_A to the one described by P_B, if the decision variable Π is regarded as a transportation plan [24,43].

A Wasserstein ambiguity set A_d is established for CDU d to include possible probability distributions of ξ_d [33].

$$A_d = \{P_d \mid dw(P_d, \widehat{P}_d) \leq \theta_d\} \quad \forall d \tag{60}$$

The ambiguity set A_d is a Wasserstein ball with a center at the empirical probability distribution $\widehat{P}_d$. Its radius is denoted by $\theta_d > 0$. Corresponding distributionally robust chance constraint is as follows:

$$\min_{P_d \in A_d} P_d\left\{\sum_t FU_{d,t} \geq \xi_d\right\} \geq 1 - \varepsilon_d \quad \forall d \tag{61}$$

It requires that the chance constraint Eq. (29) can be met under the worst situation in the ambiguity set A_d. Eq. (61) can be transferred as

$$\max_{P_d \in A_d} P_d\left\{\sum_t FU_{d,t} < \xi_d\right\} \leq \varepsilon_d \quad \forall d \tag{62}$$

Since ξ_d is a one-dimensional uncertain parameter, constraint Eq. (62) is equivalent to

$$\sum_t FU_{d,t} \geq DMW_d \quad \forall d \tag{63}$$

$$DMW_d = \min_{\nu_d}\left\{\nu_d \middle| \max_{P_d \in A_d} P_d\{\nu_d < \xi_d\} \leq \varepsilon_d\right\} \quad \forall d \tag{64}$$

where DMW_d denotes the minimum amount of oil that needs to be supplied into CDU d during the scheduling horizon in the WDRCC model, and ν_d is an auxiliary variable.

$\max_{P_d \in A_d} P_d\{\nu_d < \xi_d\} \leq \varepsilon_d$ can be expressed in the form of the worst expectation of the indicator function.

$$\max_{P_d \in A_d} E_{P_d}\{I(\nu_d < \xi_d)\} \leq \varepsilon_d \quad \forall d \tag{65}$$

$$I(\nu_d < \xi_d) = \begin{cases} 1 & \nu_d < \xi_d \\ 0 & \text{otherwise} \end{cases} \quad \forall d \tag{66}$$

where $I(\cdot)$ is called the indicator function, and $E_{P_d}\{\cdot\}$ calculates the expectation with the probability distribution P_d. The $(1-\varepsilon)$ conditional value-at-risk (*CVaR*) of the uncertain parameter z is defined as follows [42].

$$CVaR_{1-\varepsilon}(z) = \min_\tau\left\{\tau + \frac{1}{\varepsilon}E_{P_Z}\{[z-\tau]_+\}\right\} \tag{67}$$

$$[z-\tau]_+ = \begin{cases} z-\tau, z-\tau \geq 0 \\ 0, \text{otherwise} \end{cases} \tag{68}$$

where τ is also an auxiliary variable. The constraint Eqs. (65) and (66) can be reformed as a *CVaR* representation by dual theory as follows [29].

$$\begin{aligned} &\frac{\theta_d}{\varepsilon_d} + \min_\tau\left\{\tau + \frac{1}{N_d \cdot \varepsilon_d}\sum_m\left[-\max(\nu_d - \widehat{\xi}_{d,m}, 0) - \tau\right]_+\right\} \\ &= \frac{\theta_d}{\varepsilon_d} + CVaR_{1-\varepsilon_d}[-\max(\nu_d - \widehat{\xi}_d, 0)] \leq 0 \quad \forall d \end{aligned} \tag{69}$$

An equivalent formula of *CVaR* is used to transfer the inequality. $CVaR_{1-\varepsilon}(z)$ can be further calculated by the weighted average of $VaR_{1-\varepsilon}(z)$ and $CVaR^+_{1-\varepsilon}(z)$.

$$CVaR_{1-\varepsilon}(z) = \lambda_{1-\varepsilon}(z) \cdot VaR_{1-\varepsilon}(z) + (1-\lambda_{1-\varepsilon}(z)) \cdot CVaR^+_{1-\varepsilon}(z) \tag{70}$$

$$CVaR^+_{1-\varepsilon}(z) = E_{P_Z}\{z | z > VaR_{1-\varepsilon}(z)\} \tag{71}$$

$$\lambda_{1-\varepsilon}(z) = \frac{P_Z\{z \leq VaR_{1-\varepsilon}(z)\} - (1-\varepsilon)}{\varepsilon} \tag{72}$$

where $CVaR^+_{1-\varepsilon}(z)$ is called the upper *CVaR* [44], which is the conditional expectation of z subject to $z > VaR_{1-\varepsilon}(z)$; and the $\lambda_{1-\varepsilon}(z)$ is the weight value. Thus, Eqs. (73) and (74) is gained.

$$\begin{aligned} CVaR_{1-\varepsilon_d}[-\max(\nu_d - \widehat{\xi}_d, 0)] = -\sum_{m=1}^{q_d+1} \alpha_{d,m} \cdot \max(\nu_d \\ - \widehat{\xi}_{d,m}, 0) \quad \forall d \end{aligned} \tag{73}$$

$$\alpha_{d,m} = \begin{cases} \dfrac{1}{N_d \cdot \varepsilon_d} & m = 1, 2, \cdots, q_d \\ 1 - \dfrac{q_d}{N_d \cdot \varepsilon_d} & m = q_d + 1 \end{cases} \quad \forall d \tag{74}$$

Constraint Eq. (64) can be expressed as follows:

$$DMW_d = \min_{\nu_d}\left\{\nu_d \middle| \frac{\theta_d}{\varepsilon_d} \leq \sum_{m=1}^{q_d+1} a_{d,m} \cdot \max(\nu_d - \widehat{\xi}_{d,m}, 0)\right\} \quad \forall d \tag{75}$$

The value of DMW_d can be gained beforehand by the bisection search, because the right side of the inequality increases monotonically with ν_d.

Finally, the data-driven WDRCC optimization model of crude scheduling (WDRCCCS) is shown in Eq. (76).

$$\begin{aligned} \min\ & \text{Total production cost defined by Eq.(49)} \\ \text{s.t.}\ & \text{Eqs. (1) – (46)} \\ & \text{Eq. (48)} \\ & \text{Eqs. (63) and (75)} \end{aligned} \tag{76}$$

It is worth noting that the out-of-sample performance of the WDRCC solution is significantly influenced by the value of the radius θ_d. A two-fold cross-validation method is adopted to select a proper radius θ_d of the Wasserstein ambiguity set for CDU d in this study. The procedure is as follows:

(1) The collected data $\widehat{\xi}_{d,m}$ ($m = 1,2,\cdots,N_d$) will be divided into a training set of size NA and a validation set of size $NV=N_d-NA$ by random sampling.
(2) NL candidate radii are given as $\theta_{d,l}$ ($l = 1,2,\cdots,NL$). $DMW_{d,l}$ is obtained for each candidate radius $\theta_{d,l}$ with the training dataset.
(3) The probability of $DMW_{d,l} < \xi_d$ will be calculated under the validation set.
(4) The process from the first step to the third step is repeated NR times. The ρ_d-th percentile of these probabilities will be gained for every radius.
(5) The smallest candidate radius $\theta_{d,l}$, whose ρ_d-th percentile of probabilities is less than the target level ε_d, is selected as the best radius for CDU d.

After the best radius is selected, all collected data are employed to recalculate a DMW_d for the subsequent optimization. The detailed process will be demonstrated through the case studies in the next section.

6. Application to the Crude Oil Scheduling

Cases are used to demonstrate the performance of the proposed approach in this paper. In Case 1, the proposed data-driven WDRCC optimization method is compared with the chance-constrained optimization method and the deterministic optimization method. The influence of the Wasserstein radii on the data-driven WDRCC model is also studied. Case 2 is used to show the applicability of the proposed method to a larger crude scheduling problem.

6.1. Case 1

One SBM station, one jetty, seven tanks, and two CDUs are involved in Case 1 with a 7-periods scheduling horizon. A time

Table 1
The properties and costs of crudes and their information in CDU1 and CDU2

	Attribute name	C1	C2	C3	C4	C5
General information	Density/g·cm^{-3}	0.827	0.880	0.873	0.916	0.911
	Sulfur content/% (mass)	1.00	2.35	1.43	3.43	0.12
	Acid[①]/mg·g^{-1}	0.05	0.20	0.60	1.23	4.38
	Cost/kUSD·kt^{-1}	565	553	550	543	540
Information in CDU1	Gasoline yield/% (mass)	27.60	17.32	16.30	17.42	3.70
	Diesel yield/% (mass)	24.08	19.60	19.42	17.08	10.71
	Residual oil yield/% (mass)	13.20	27.38	26.70	35.27	53.15
	Recipe upper limit/% (mass)	80	80	100	80	80
	Recipe lower limit/% (mass)	0	0	0	0	0
Information in CDU2	Gasoline yield/% (mass)	27.90	17.35	17.25	17.15	3.65
	Diesel yield/% (mass)	24.10	19.50	22.12	18.75	10.72
	Residual oil yield/% (mass)	13.18	27.20	28.08	35.60	53.15
	Recipe upper limit/% (mass)	80	80	100	80	80
	Recipe lower limit/% (mass)	0	0	0	0	0

① Measured by mg KOH/g crude.

Table 2
Data of vessels in Case 1

Vessel	Parcel	[Crude, parcel size/kt]	ETA_p	ETD_v	SWC_v/kUSD
V1	P1	[C5, 60.00]	0	1	1000
V2	P2	[C1, 60.00]	2	1	1000
VLCC1	P3	[C3, 80.00]	3	2	1000
	P4	[C4, 80.00]	3	2	1000
Maximum unloading rate of one parcel per period/kt					120
Minimum unloading rate of one parcel per period/kt					5

period corresponds to a day. The information of crudes is shown in Table 1. There are two single-parcel vessels (V1 and V2) and one VLCC (VLCC1). VLCC1 carries two crude parcels (P3 and P4). The data of arrival times, parcels, and unloading rates of vessels are given in Table 2.

Table 3 presents the initial inventories, compositions, and capacity bounds of tanks. Key parameters of CDUs are displayed in Table 4. At most four tanks can feed one CDU in one period. The cost of one changeover is 100 kUSD.

500 demand data samples of CDU1 and CDU2 are collected from the historical database of a real-world refinery. The histograms of these data are shown in Fig. 2. The target risk level ε_d is set to 0.05 for CCCSED and WDRCCCS. The cross-validation method is used to choose the right Wasserstain radii θ_{CDU1} and θ_{CDU2}, where NA is 250 and ρ_d is 90. Candidate radii $\{0.2 \times 10^{-3}, 0.4 \times 10^{-3}, \cdots, 2.0 \times 10^{-3}\}$ are considered for every CDU. $NR = 10$ independent runs are carried

Table 3
Initial data and capacity limits of tanks in Case 1

Tank	Initial inventory/kt	Initial composition: [Crude, mass/kt]	Lower limit of capacity/kt	Lower limit of capacity/kt
Tank1	50.00	[C1, 50.00]	100	10
Tank2	20.00	[C4, 20.00]	100	10
Tank3	80.00	[C5, 80.00]	100	10
Tank4	90.00	[C3, 90.00]	100	10
Tank5	40.00	[C2, 40.00]	100	10
Tank6	80.00	[C1, 80.00]	100	10
Tank7	30.00	[C2, 30.00]	100	10

Table 4
Key parameters of CDU1 and CDU2

CDU	Type of parameters	Attribute name	Upper limit	Lower limit
CDU1	Limits for mixed input	Transfer rate per period/kt	21.50	17.00
		Density/g·cm^{-3}	0.89	0.84
		Sulfur content/% (mass)	2.50	0
		Acid/mg·g^{-1}	1.50	0
	Limits for product output	Gasoline output per period/kt	5.28	1.44
		Diesel output per period/kt	5.28	2.4
		Residual oil output per period/kt	6.48	0
CDU2	Limits for mixed input	Transfer rate per period/kt	18.00	15.00
		Density/g·cm^{-3}	0.89	0.84
		Sulfur content/% (mass)	2.50	0
		Acid/mg·g^{-1}	2.00	0
	Limits for product output	Gasoline output per period/kt	2.88	0.96
		Diesel output per period/kt	4.08	1.92
		Residual oil output per period/kt	6.00	0

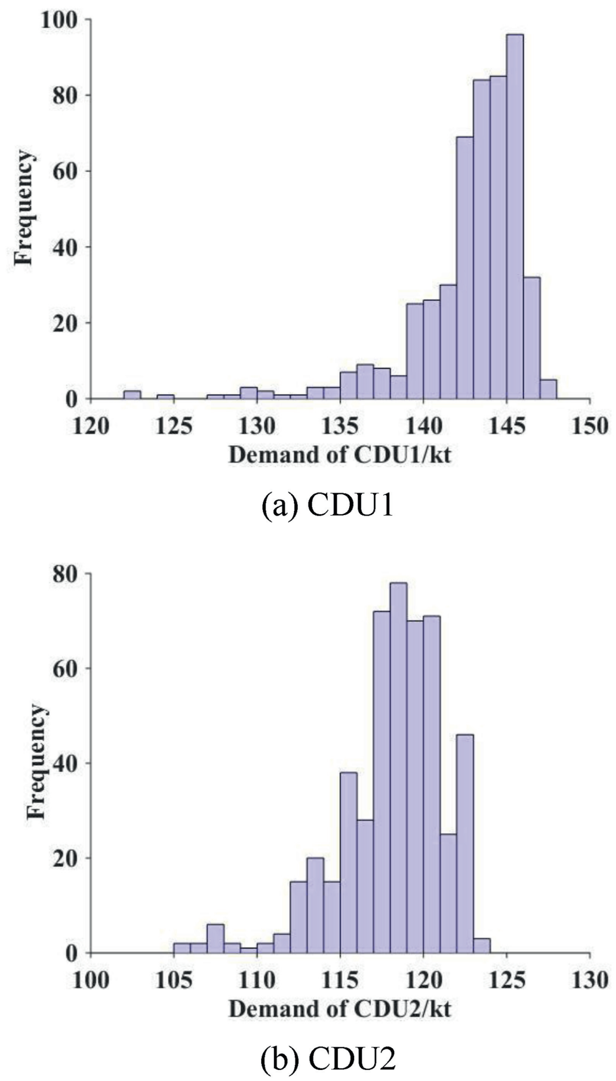

Fig. 2. Histograms of demand data in Case 1.

out for each radius to calculate probabilities of $DMW_{d,l} < \xi_d$ and ρ_d-th percentiles. Table 5 and Table 6 exhibit the results. As shown in Table 5, the 90th percentiles of probabilities of radius candidates 1.4×10^{-3}, 1.6×10^{-3}, 1.8×10^{-3} and 2.0×10^{-3} are below the 0.05. 1.4×10^{-3} is the smallest radius that satisfies the requirement and is selected as the optimal Wasserstein radius of CDU1. In the same way, 1.0×10^{-3} is chosen as the best θ_{CDU2} according to the results in Table 6.

Then, all 500 data samples are adopted to get DMW_d for WDRCCCS under selected radii. In the deterministic model, the demand DM_d is the average value of demand data of CDU d. The DMV_d is also calculated for the chance-constrained model by Eq. (57). Their values are shown in Table 7, where $DM_d < DMV_d < DMW_d$. All models are solved by GAMS 24.1.2/Glo-MIQO 2.2 on a computer of Intel(R) Core(TM) i7-8550U CPU @ 1.8 GHz, 2.00 GHz, and 8 GB RAM. The problem sizes and solutions are presented in Table 8. The data of model sizes are counted by GAMS 24.1.2 platform.

The number of variables and constraints of the three models are the same. It indicates no increase in computational complexity with the introduction of Wasserstein ambiguity sets, which is a merit of the proposed method. This is because DMV_d and DMW_d can be viewed as parameters in the established optimization models since they can be obtained respectively by Eq. (57) and Eq. (75) before the optimization. The total production cost is 143592.93 kUSD in the deterministic model. Uncertain demands are considered in CCSED, and simple empirical probability distributions are used. The objective value in CCSED is 147782.65 kUSD. The decisions in WDRCCCS are made based on the worst probability distributions in the ambiguity sets. Therefore, the objective value in WDRCCCS is the most conservative, which is 148136.64 kUSD. Furthermore, there is no demurrage cost in the results of all models, and changeover costs are all 400 kUSD. Since these models aim to minimize the total production cost, the crude oil consumption of each CDU during the scheduling horizon tends to be equal to the minimum feed requirement (*i.e.*, DM_d, DMV_d or DMW_d) in the respective model. Consequently, more crude oil is processed in the solution of WDRCCCS, which results in a higher crude oil cost than those in the other models.

The operational schedules are displayed in Figs. 3–5. As shown in these figures, VLCC parcels are unloaded one by one. For example, the unloading of VLCC parcel P4 is performed after the unloading of VLCC parcel P3 in Figs. 3–5. At most two tanks can receive oil from a parcel in one period. There are a maximum of two oil unloading operations from VLCC parcels and a maximum of 2NJ oil unloading operations from jetty parcels at the same time. A tank does not receive oil and send oil at the same time. A brine settling operation exists after a receiving oil operation. One tank can supply up to two sets of CDUs, and one CDU can receive oil from up to four tanks during one period. These phenomena demonstrate that these obtained crude oil scheduling schemes is qualified and prove the validity of the formulas in Section 3. Besides, each feed flow of a

Table 5
Selection of the radius for CDU1 *via* cross-validation

$\theta \times 10^3$	90th percentile	Violation probability for each run									
		1	2	3	4	5	6	7	8	9	10
0.2	0.060	0.036	0.052	0.024	0.024	0.060	0.048	0.056	0.048	0.060	0.036
0.4	0.060	0.036	0.052	0.024	0.024	0.060	0.044	0.052	0.040	0.060	0.032
0.6	0.060	0.036	0.052	0.024	0.024	0.060	0.044	0.048	0.036	0.060	0.028
0.8	0.056	0.032	0.052	0.024	0.024	0.056	0.040	0.048	0.036	0.060	0.028
1.0	0.056	0.032	0.052	0.024	0.024	0.056	0.040	0.044	0.032	0.060	0.028
1.2	0.052	0.032	0.048	0.024	0.024	0.052	0.036	0.044	0.032	0.056	0.028
1.4	0.048	0.032	0.048	0.024	0.024	0.048	0.036	0.044	0.032	0.056	0.024
1.6	0.048	0.032	0.048	0.024	0.024	0.048	0.036	0.044	0.032	0.056	0.024
1.8	0.048	0.028	0.048	0.020	0.024	0.044	0.032	0.044	0.032	0.052	0.024
2.0	0.044	0.028	0.044	0.020	0.024	0.044	0.032	0.040	0.032	0.052	0.024

Table 6
Selection of the radius for CDU2 *via* cross-validation

$\theta \times 10^3$	90th percentile	Violation probability for each run									
		1	2	3	4	5	6	7	8	9	10
0.2	0.064	0.036	0.032	0.064	0.056	0.060	0.016	0.044	0.072	0.016	0.020
0.4	0.064	0.032	0.032	0.064	0.044	0.048	0.016	0.044	0.068	0.016	0.020
0.6	0.056	0.028	0.020	0.056	0.044	0.040	0.016	0.044	0.064	0.016	0.020
0.8	0.052	0.024	0.016	0.052	0.044	0.040	0.016	0.040	0.060	0.016	0.020
1.0	0.044	0.024	0.016	0.044	0.040	0.032	0.012	0.036	0.056	0.016	0.020
1.2	0.044	0.024	0.016	0.044	0.040	0.032	0.012	0.036	0.056	0.012	0.016
1.4	0.044	0.024	0.016	0.044	0.040	0.024	0.012	0.032	0.052	0.012	0.016
1.6	0.044	0.024	0.016	0.044	0.040	0.020	0.012	0.024	0.052	0.012	0.016
1.8	0.040	0.024	0.016	0.040	0.040	0.020	0.008	0.024	0.052	0.012	0.016
2.0	0.040	0.024	0.016	0.040	0.032	0.020	0.008	0.024	0.052	0.008	0.016

Table 7
Values of DM_d, DMV_d and DMW_d

	CDU1	CDU2
DM_d/kt	142.77	118.12
DMV_d/kt	146.21	122.36
DMW_d/kt	146.42	122.50

Table 8
Problem sizes and solutions for DCSO, CCCSED and WDRCCCS in Case 1

	DCSO	CCSED	WDRCCCS
Binary variables	294	294	294
Continuous variables	1468	1468	1468
Constraints	2883	2883	2883
Total production cost/kUSD	143592.93	147782.65	148136.64
Cost of oil/kUSD	143192.93	147382.65	147736.64
Demurrage cost of ships/kUSD	0	0	0
Changeover cost/kUSD	400	400	400
Total input of CDU1/kt	142.77	146.21	146.42
Total input of CDU2/kt	118.12	122.36	122.50

CDU is stable regardless of the number of its supply tanks if there is no change in the combination of tanks, which shows the efficacy of the new production stability constraints.

The out-of-sample performance of solutions is examined by 10 sets of test data. Each data set contains 500 data samples. The violation probabilities of the inequality $\sum_t FU_{d,t} \geq \xi_d$ under these schedules are obtained and presented in Table 9. Similar to the cross-validation process, the 90th percentiles of probabilities are also provided. Because the distributions of demands are not symmetric, as shown in Fig. 2, violation levels under the production plan made by the deterministic model are even larger than 0.5, although the average values of modeling data are used. The violation probabilities are basically around the target risk level 0.05 in the chance-constrained model with the usage of the $(1\text{-}\varepsilon_d)$ value-at-risk of empirical distributions. However, there are four times when the violation probability exceeds the target risk level in CDU1 and six times in CDU2 because of the difference between the empirical probability distributions and the true probability distributions. The 90th percentiles in the two CDUs are 0.062 and 0.058, which are also larger than 0.05. It is difficult to keep the violation probability below the target risk level by CCSED. For the WDRCC model, the violation degrees are all less than 0.05, with the 90th percentiles 0.048 and 0.044. These tests imply that the production plan outputted by the WDRCC approach enjoys better robustness.

Experiments with different Wasserstein radii are carried out to reveal the influence of radii on the performance of the WDRCC method. Let $\theta_{\text{CDU1}} = \theta_{\text{CDU2}} = \theta^*$ for the brevity of the description. Fig. 6 shows the total production costs and 90th percentiles of violation probabilities of WDRCCCS under $\theta^* \in \{0.5 \times 10^{-3}, 1.0 \times 10^{-3}, 1.5 \times 10^{-3}, 2.0 \times 10^{-3}, 2.5 \times 10^{-3}, 3.0 \times 10^{-3}, 3.5 \times 10^{-3},$

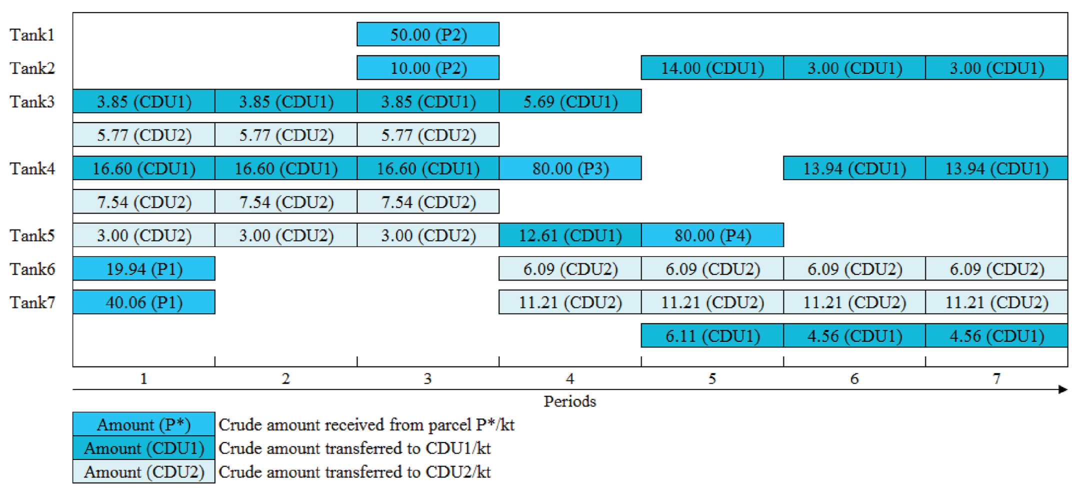

Fig. 3. Operational schedule of DCSO in Case 1.

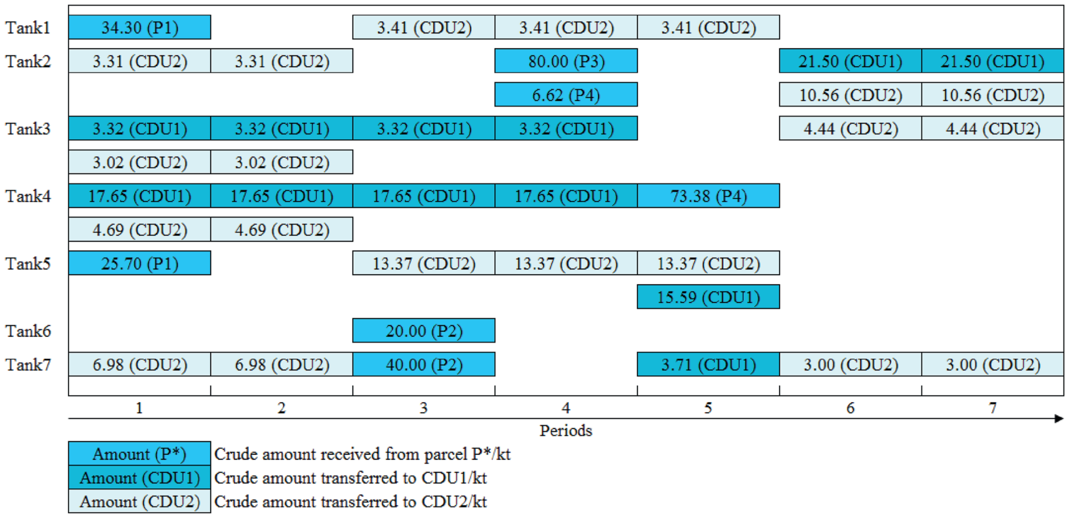

Fig. 4. Operational schedule of CCSED in Case 1.

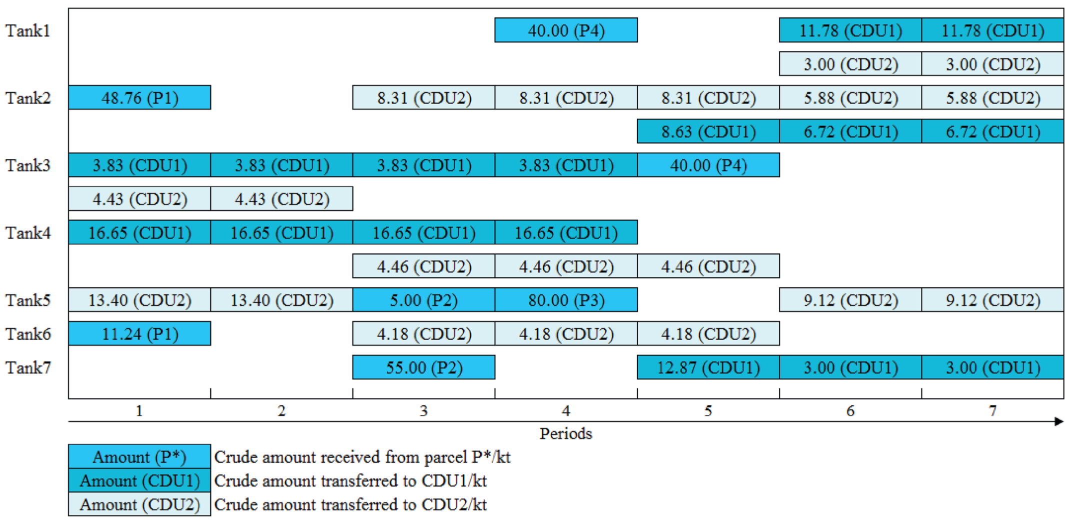

Fig. 5. Operational schedule of WDRCCCS in Case 1.

Table 9
Out-of-sample performance of schedules in Case 1

CDU	Model	90th percentile	Violation probability in each run									
			1	2	3	4	5	6	7	8	9	10
CDU1	DCSO	0.656	0.654	0.676	0.640	0.624	0.654	0.654	0.634	0.656	0.652	0.644
	CCSED	0.062	0.052	0.044	0.066	0.050	0.048	0.042	0.042	0.060	0.042	0.062
	WDRCCCS	0.048	0.032	0.028	0.048	0.036	0.034	0.03	0.03	0.046	0.028	0.048
CDU2	DCSO	0.570	0.542	0.564	0.548	0.562	0.570	0.572	0.542	0.562	0.532	0.542
	CCSED	0.058	0.056	0.058	0.044	0.052	0.064	0.046	0.046	0.058	0.050	0.054
	WDRCCCS	0.044	0.040	0.044	0.030	0.032	0.040	0.030	0.030	0.046	0.030	0.034

4.0×10^{-3}}. The 90th percentile decreases from 0.054 to 0.04 in CDU1 and from 0.055 to 0.022 in CDU2 with the increasing radius θ^*. The total production cost grows from 148044.05 kUSD to 148373.43 kUSD. This is because larger radii result in larger Wasserstein ambiguity sets, which brings a more conservative and robust result.

6.2. Case 2

The WDRCC approach is also applied to a larger case. Case 2 contains three VLCCs, whose information is described in Table 10. Table 11 displays the initial crude weights in eight tanks. An additional CDU (CDU3) is considered. The yields of products and the

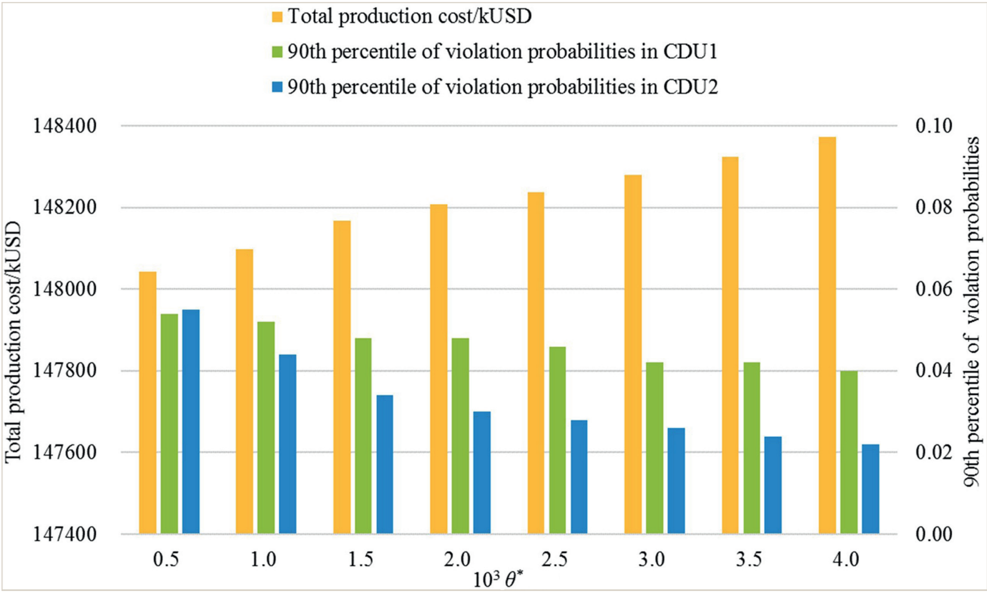

Fig. 6. Total production costs and 90th percentiles of violation probabilities of WDRCCCS under different θ^*.

Table 10
Data of vessels in Case 2

Vessel	Parcel	[Crude, parcel size/kt]	ETA_p	ETD_v	SWC_v/kUSD
VLCC1	P1	[C1, 80.00]	0	2	1000
	P2	[C2, 60.00]	0	2	1000
VLCC2	P3	[C5, 60.00]	2	2	1000
	P4	[C3, 60.00]	2	2	1000
VLCC3	P5	[C4, 80.00]	5	2	1000
	P6	[C2, 60.00]	5	2	1000
Maximum unloading rate of one parcel per period/kt					120
Minimum unloading rate of one parcel per period/kt					5

Table 12
Product yields and recipe limits of CDU3

Attribute name	C1	C2	C3	C4	C5
Gasoline yield/% (mass)	26.05	16.23	16.25	16.05	3.80
Diesel yield/% (mass)	24.10	19.55	19.40	17.02	10.60
Residual oil yield/% (mass)	12.08	26.30	26.50	34.14	52.60
Recipe upper limit/% (mass)	80	80	100	80	80
Recipe lower limit/% (mass)	0	0	0	0	0

recipe limits for crudes in CDU3 are shown in Table 12. Its key parameters are given in Table 13. A histogram of the demand data is shown in Fig. 7.

The θ_{CDU3} is chosen as 1.6×10^{-3} through the cross-validation method. The corresponding DMW_{CDU3} is equal to 168.27 kt. The solution of Case 2 is successfully found and shown in Table 14, although there are more variables and constraints in the model. Fig. 8 shows the detailed production plan. It can be seem that the production plan is compliant with operation rules of crude oil scheduling. The objective value is 241712.08 kUSD, which is higher than that in Case 1. This is mainly because more oil is processed with the introduction of CDU3. Another 10 sets of test data are sampled from the database to test the out-of-sample performance of the scheme. The violation probabilities of the demand constraints of three CDUs are not greater than 0.05, as shown in Table 15. The maximum values of the violation level are 0.048 in CDU1, 0.046 in CDU2 and 0.048 in CDU3.

The numerical experiments shows that the WDRCC model has better out-of-sample performance though it has the same model

Table 11
Initial data and capacity limits of tanks in Case 2

Tank	Initial inventory/kt	Initial composition: [Crude, mass/kt]	Upper limit of capacity/kt	Lower limit of capacity/kt
Tank1	10.00	[C4, 10.00]	100	10
Tank2	80.00	[C1, 80.00]	100	10
Tank3	90.00	[C3, 90.00]	100	10
Tank4	20.00	[C5, 20.00]	100	10
Tank5	90.00	[C2, 90.00]	100	10
Tank6	20.00	[C1, 20.00]	100	10
Tank7	80.00	[C3, 80.00]	100	10
Tank8	90.00	[C5, 90.00]	100	10

Table 13
Key parameters of CDU3

Type of limits	Attribute name	Upper limit	Lower limit
Limits for mixed input	Transfer rate per period/kt	24.50	20.00
	Density/g·cm^{-3}	0.89	0.84
	Sulfur content/% (mass)	2.50	0
	Acid/mg·g^{-1}	2.00	0
Limits for product output	Gasoline output per period/kt	5.28	1.20
	Diesel output per period/kt	5.28	2.40
	Residual oil output per period/kt	6.48	0

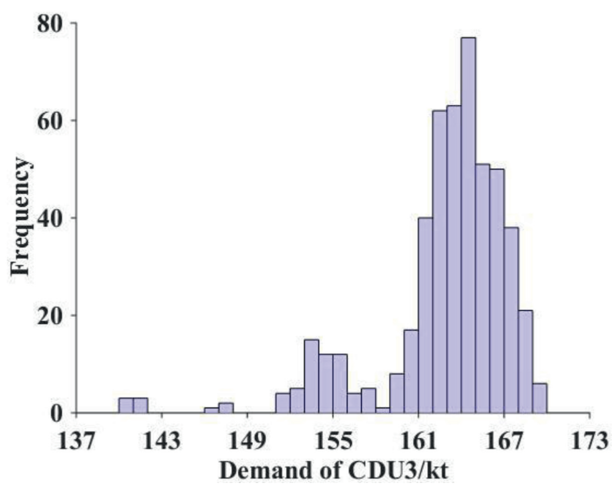

Fig. 7. Histograms of demand data of CDU3.

size with the deterministic and the chance-constrained models. The effectiveness of the cross-validation approach is also demonstrated by Case 1 and Case 2. The violation degrees are all in the

Table 14
Problem sizes and solutions for WDRCCCS in Case 2

	WDRCCCS
Binary variables	504
Continuous variables	2176
Constraints	4564
Total production cost/kUSD	241712.08
Cost of Oil/kUSD	241112.08
Demurrage cost of ships/kUSD	0
Changeover cost/kUSD	600
Total input of CDU1/kt	146.42
Total input of CDU2/kt	122.50
Total input of CDU3/kt	168.27

required range under these selected radii. Moreover, it is worth noting that an existing deterministic optimization model can be easily updated as the form of this data-driven WDRCC model by replacing DM_d with DMW_d despite of the complex derivation of the mathematical equations in this paper. Above all, the proposed method can give a reliable crude oil scheduling scheme without increasing the model complexity compared to traditional methods. This method is also easy to implement and has a rigorous statistical theoretical basis. These advantages make it promising for engineering applications.

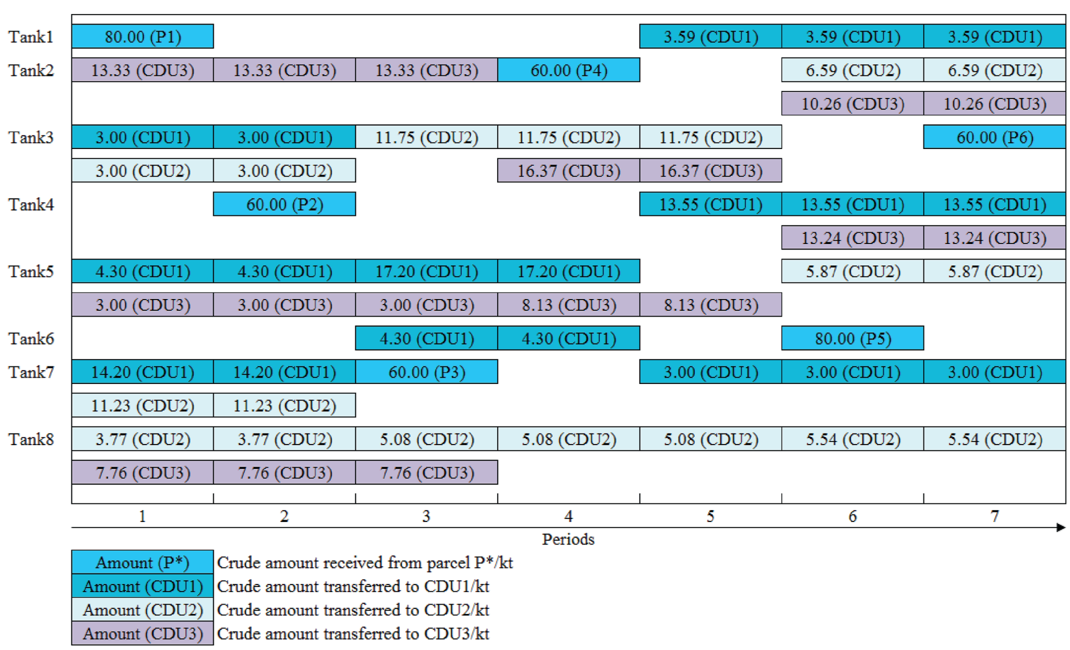

Fig. 8. Operational schedule of WDRCCCS in Case 2.

Table 15
The out-of-sample performance of the schedule of WDRCCCS in Case 2

CDU	90th percentile	Violation probability in each run									
		1	2	3	4	5	6	7	8	9	10
CDU1	0.046	0.030	0.030	0.036	0.046	0.034	0.040	0.034	0.042	0.048	0.038
CDU2	0.044	0.036	0.044	0.032	0.036	0.034	0.030	0.042	0.038	0.046	0.034
CDU3	0.046	0.022	0.038	0.024	0.042	0.048	0.032	0.046	0.032	0.030	0.034

7. Conclusions

This paper presents a data-driven WDRCC optimization model for crude oil scheduling under demand uncertainty. It provides a novel idea to improve the robustness of the industrial decision making through effective utilization of production data. First, the traditional deterministic crude oil scheduling optimization model is improved. Wasserstein ambiguity sets are then established from the demand data of CDUs, the center of which are empirical probability distributions. A cross-validation method is utilized to select suitable Wasserstein radii. The WDRCC optimization model is built based on the ambiguity sets. It is reformulated as a solvable form using dual theory and the equivalent equations of CVaR. The proposed optimization model is compared with the deterministic optimization model and the chance-constrained optimization model in Case 1. The objective value is 143592.93 kUSD in DCSO, 147782.65 kUSD in CCSED, and 148136.64 kUSD in WDRCCCS. However, out-of-sample tests show that the solution of the deterministic optimization model has high infeasible infeasibility probabilities because of the ignorance of uncertainties. The chance-constrained optimization model fails to keep violation probabilities of demand constraints in the required ranges within the desired range due to the discrepancy between the empirical probability distributions and the true probability distributions. Although the objective value of WDRCCCS is the highest, the proposed approach can effectively ensure that the violation probabilities are below the target risk value. The robustness of WDRCCCS can be strengthened by increasing the Wasserstein radii, which is accompanied by an enhancement in conservatism. Case 2 demonstrates the applicability of this method to a larger-scale problem. The proposed data-driven WDRCC approach also has great potential to be applied to other scheduling optimization problems under demand uncertainty, though the focuses of this paper is on the crude oil scheduling optimization problem.

Declaration of 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

The authors acknowledge the supports from National Natural Science Foundation of China (61988101, 62073142, 22178103), National Natural Science Fund for Distinguished Young Scholars (61925305) and International (Regional) Cooperation and Exchange Project (61720106008).

Nomenclature

COC	cost for once changeover, kUSD
$CO_{d,t}$	1 if there is a change of the feed of CDU d in period t
$Cost_c$	material cost of crude c, $\mathrm{kUSD}\cdot\mathrm{t}^{-1}$
DC_v	demurrage cost of ship v
DM_d	demand of CDU d during the scheduling horizon
DMV_d	minimum requirement for the total amount of feed for CDU d during the scheduling horizon in the CCCSED model
DMW_d	minimum requirement for the total amount of feed for CDU d during the scheduling horizon in the WDRCCCS model
ETA_p	estimated time of arrival of parcel p
ETD_v	allowable dwell time of vessel v
$FCTUS_{d,c,t}$	amount of crude c sent into CDU d during period t
$FCTU_{i,d,c,t}$	amount of crude c pumped into CDU d from tank i during period t
$FPT_{p,i,t}$	amount of oil unloaded into tank i from parcel p during period t
$FPT^{L/U}{}_{p,i}$	bounds on the unloading rate in the connection between parcel p and i
$FTU_{i,d,t}$	amount of oil transferred into CDU d from tank i during period t
$FTU^{L/U}{}_{i,d}$	bounds on the transmission rate in the connection between tank i and CDU d
$FU_{d,t}$	feed amount of CDU d during period t
$FU^{L/U}{}_d$	bounds on the crude oil input rate of CDU d
$f_{i,c,t}$	fraction of crude c in tank i at the end of period t
IC	set of pairs (tank i, crude c) that c can be stored in i
ID	set of pairs (tank i, CDU d) that i can feed d
JP	set of jetty parcels
M	a large number such as 10^9
NJ	number of jetties
NT_d	the upper limit of the number of tanks that can supply a CDU simultaneously
$PF^{L/U}{}_{d,j}$	bounds on the output rate of product j in CDU d
PC	set of pairs (parcel p, crude c) that p carries c
PI	set of pairs (parcel p, tank i) that p may unload oil into i
PS_p	total mass of oil in parcel p
PT	set of pairs (parcel p, period t) that crude oil in p can be unloaded in t
PV	set of pairs (parcel p, vessel v) that p is the last in the parcels carried by v
SP	set of VLCC parcels
SWC_v	demurrage cost of ship v per period, kUSD
TF_p	time at which parcel p first connects to unloading facilities
TL_p	time at which parcel p disconnects unloading facilities
$W_{i,t}$	crude amount in tank i at the end of period t
$W^{L/U}{}_i$	capacity bounds of tank i
$WCT_{i,c,t}$	amount of crude c in tank i at the end of period t
$X_{p,i,t}$	1 if parcel p charges tank i during period t
$XF_{p,t}$	1 if parcel p first connects to a discharge line in period t
$XL_{p,t}$	1 if parcel p disconnects the discharge facilities in period t
$XP_{p,t}$	1 if parcel p supply a tank during period t
$XT_{i,t}$	1 if tank i receives oil during period t
$xc^{L/U}{}_{c,d}$	bounds on the fraction of crude c in the feed of CDU d
$xk^{L/U}{}_{k,d}$	bounds on the property k of the feed of CDU d
$xk_{k,c}$	property k of crude c

$Y_{i,d,t}$	1 if CDU d is fed by tank i during period t
$Yield_{c,j,d}$	yield of product j of crude c in CDU d
$YY_{i,d,t}$	1 if the connection state of tank i and CDU d has changed in period t

Subscripts

c	crudes
d	crude distillation units
i	storage tanks
j	products
k	crude properties
p	crude parcels
t	periods
v	vessels

Superscripts

L	lower bounds
U	upper bounds

References

[1] L. Bai, Y.H. Jiang, D.X. Huang, X.G. Liu, A novel scheduling strategy for crude oil blending, *Chin. J. Chem. Eng.* 18 (5) (2010) 777–786.
[2] J. Li, W.K. Li, I.A. Karimi, R. Srinivasan, Improving the robustness and efficiency of crude scheduling algorithms, *AIChE J.* 53 (10) (2007) 2659–2680.
[3] H. Lee, J.M. Pinto, I.E. Grossmann, S. Park, Mixed-integer linear programming model for refinery short-term scheduling of crude oil unloading with inventory management, *Ind. Eng. Chem. Res.* 35 (5) (1996) 1630–1641.
[4] J.M. Pinto, M. Joly, L.F.L. Moro, Planning and scheduling models for refinery operations, *Comput. Chem. Eng.* 24 (9-10) (2000) 2259–2276.
[5] W.K. Li, C.W. Hui, B. Hua, Z.X. Tong, Scheduling crude oil unloading, storage, and processing, *Ind. Eng. Chem. Res.* 41 (26) (2002) 6723–6734.
[6] P.C.P. Reddy, I.A. Karimi, R. Srinivasan, Novel solution approach for optimizing crude oil operations, *AIChE J.* 50 (6) (2004) 1177–1197.
[7] R. Karuppiah, K.C. Furman, I.E. Grossmann, Global optimization for scheduling refinery crude oil operations, *Comput. Chem. Eng.* 32 (11) (2008) 2745–2766.
[8] P.C.P. Reddy, I.A. Karimi, R. Srinivasan, A new continuous-time formulation for scheduling crude oil operations, *Chem. Eng. Sci.* 59 (6) (2004) 1325–1341.
[9] B. Zimberg, E. Ferreira, E. Camponogara, A continuous-time formulation for scheduling crude oil operations in a terminal with a refinery pipeline, *Comput. Chem. Eng.* 178 (2023) 108354.
[10] Y.N. Yang, R.C. He, G. Yu, W. Du, M.L. Yang, W.L. Du, Efficient rolling horizon approach to a crude oil scheduling problem for marine-access refineries, *Comput. Chem. Eng.* 170 (2023) 108121.
[11] 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.
[12] Y. Chen, Z.H. Yuan, B.Z. Chen, Process optimization with consideration of uncertainties—An overview, *Chin. J. Chem. Eng.* 26 (8) (2018) 1700–1706.
[13] S. Gupta, N. Zhang, Flexible scheduling of crude oil inventory management, *Ind. Eng. Chem. Res.* 49 (3) (2010) 1325–1332.
[14] C.F. Li, X.R. He, B.Z. Chen, Q. Xu, C.W. Liu, A hybrid programming model for optimal production planning under demand uncertainty in refinery, *Chin. J. Chem. Eng.* 16 (2) (2008) 241–246.
[15] J. Li, R. Misener, C.A. Floudas, Scheduling of crude oil operations under demand uncertainty: A robust optimization framework coupled with global optimization, *AIChE J.* 58 (8) (2012) 2373–2396.
[16] C.P. Luo, G. Rong, A strategy for the integration of production planning and scheduling in refineries under uncertainty, *Chin, J. Chem. Eng.* 17 (1) (2009) 113–127.
[17] D. Panda, M. Ramteke, Preventive crude oil scheduling under demand uncertainty using structure adapted genetic algorithm, *Appl. Energy* 235 (2019) 68–82.
[18] X. Dai, L. Zhao, Z. Li, W.L. Du, W.M. Zhong, R.C. He, F. Qian, A data-driven approach for crude oil scheduling optimization under product yield uncertainty, *Chem. Eng. Sci.* 246 (2021) 116971.
[19] B. Liu, Y.F. Wang, X. Feng, Optimization of circulating cooling water systems based on chance constrained programming, *Chin. J. Chem. Eng.* 40 (2021) 167–178.
[20] J. Wang, Optimizing crude oil operations under Uncertainty13th IFAC symposium on information control problems in manufacturing, in: IFAC Proceedings Volumes, Elsevier, 2009.
[21] C.W. Cao, X.S. Gu, Z. Xin, Chance constrained programming models for refinery short-term crude oil scheduling problem, *Appl. Math. Model.* 33 (3) (2009) 1696–1707.
[22] C.W. Cao, X.S. Gu, Z. Xin, Stochastic chance constrained mixed-integer nonlinear programming models and the solution approaches for refinery short-term crude oil scheduling problem, *Appl. Math. Model.* 34 (11) (2010) 3231–3243.
[23] L. Zhao, C. Ning, F.Q. You, Operational optimization of industrial steam systems under uncertainty using data-Driven adaptive robust optimization, *AIChE J.* 65 (7) (2019) e16500.
[24] P.M. Esfahani, D. Kuhn, Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations, *Math. Program. Ser. A B* 171 (1–2) (2018) 115–166.
[25] J.M. Zhao, L. Zhao, W.L. He, Data-driven Wasserstein Distributionally Robust Optimization for Refinery Planning under Uncertainty, IECON 2021 – 47th Annual Conference of the IEEE Industrial Electronics Society, Canada. IEEE, Toronto, ON, 2021, pp. 1–6.
[26] Q.Y. Bian, H.H. Xin, Z. Wang, D.Q. Gan, K.P. Wong, Distributionally robust solution to the reserve scheduling problem with partial information of wind power, *IEEE Trans. Power Syst.* 30 (5) (2015) 2822–2823.
[27] Z. Wang, Q.Y. Bian, H.H. Xin, D.Q. Gan, A distributionally robust co-ordinated reserve scheduling model considering CVaR-based wind power reserve requirements, *IEEE Trans. Sustain. Energy* 7 (2) (2016) 625–636.
[28] C. Shang, F.Q. You, Distributionally robust optimization for planning and scheduling under uncertainty, *Comput. Chem. Eng.* 110 (2018) 53–68.
[29] W.J. Xie, On distributionally robust chance constrained programs with Wasserstein distance, *Math. Program. Ser. A B* 186 (1–2) (2021) 115–155.
[30] C. Ning, F.Q. You, Data-driven Wasserstein distributionally robust optimization for biomass with agricultural waste-to-energy network design under uncertainty, *Appl. Energy* 255 (2019) 113857.
[31] E. Guevara, F. Babonneau, T. Homem-de-Mello, S. Moret, A machine learning and distributionally robust optimization framework for strategic energy planning under uncertainty, *Appl. Energy* 271 (2020) 115005.
[32] B.T. Liu, Q. Zhang, Z.H. Yuan, Two-stage distributionally robust optimization for maritime inventory routing, *Comput. Chem. Eng.* 149 (2021) 107307.
[33] R. Ji, M.A. Lejeune, Data-driven distributionally robust chance-constrained optimization with Wasserstein metric, *J. Glob. Optim.* 79 (4) (2021) 779–811.
[34] B.K. Poolla, A. Hota, S. Bolognani, D. Callaway, A. Cherukuri, Wasserstein Distributionally Robust Look-Ahead Economic Dispatch, 2021 IEEE Power & Energy Society General Meeting (PESGM), IEEE, Washington, DC, USA, 2021, p. 1.
[35] Y.Q. Zhou, W.B. Yu, S.Y. Zhu, B. Yang, J.P. He, Distributionally robust chance-constrained energy management of an integrated retailer in the multi-energy market, *Appl. Energy* 286 (2021) 116516.
[36] W.J. Xie, S. Ahmed, Bicriteria approximation of chance-constrained covering problems, *Oper. Res.* 68 (2) (2020) 516–533.
[37] C. Ordoudis, V.A. Nguyen, D. Kuhn, P. Pinson, Energy and reserve dispatch with distributionally robust joint chance constraints, *Oper. Res. Lett.* 49 (3) (2021) 291–299.
[38] S. Zymler, D. Kuhn, B. Rustem, Distributionally robust joint chance constraints with second-order moment information, *Math. Program.* 137 (1) (2013) 167–198.
[39] A. Arrigo, C. Ordoudis, J. Kazempour, Z. De Grève, J.F. Toubeau, F. Vallée, Wasserstein distributionally robust chance-constrained optimization for energy and reserve dispatch: An exact and physically-bounded formulation, *Eur. J. Oper. Res.* 296 (1) (2022) 304–322.
[40] C. Duan, W.L. Fang, L. Jiang, L. Yao, J. Liu, Distributionally robust chance-constrained approximate AC-OPF with Wasserstein metric, *IEEE Trans. Power Syst.* 33 (5) (2018) 4924–4936.
[41] Y. Guo, K. Baker, E. Dall'Anese, Z.C. Hu, T.H. Summers, Data-based distributionally robust stochastic optimal power flow—part II: Case studies, *IEEE Trans. Power Syst.* 34 (2) (2019) 1493–1503.
[42] R.T. Rockafellar, S. Uryasev, Conditional value-at-risk for general loss distributions, *J. Bank. Finance* 26 (7) (2002) 1443–1471.
[43] V.M. Panaretos, Y. Zemel, Statistical aspects of Wasserstein distances, *Annu. Rev. Stat. Appl.* 6 (2019) 405–431.
[44] S. Sarykalin, G. Serraino, S. Uryasev, Value-at-risk *vs.* conditional value-at-risk in risk management and optimization. State-of-the-Art Decision-Making Tools in the Information-Intensive Age, *Informs* (2008) 270–294.