CRUDE-OIL BLEND SCHEDULING OPTIMIZATION
OF AN INDUSTRIAL-SIZED REFINERY: A
DISCRETE-TIME BENCHMARK

JD Kelly, BC Menezes, F Engineer and IE Grossmann
Industrial Algorithms
Toronto, ON MIP-4C3
Carnegie Mellon University
Pittsburgh, PA 15213
SK-Innovation
Seoul, Jongno-gu

Abstract
We propose a discrete-time formulation for optimization of scheduling in crude-oil refineries considering both the logistics details practiced in industry and the process feed diet and quality calculations. The quantity-logic-quality phenomena (QLQP) involving a non-convex mixed-integer nonlinear (MINLP) problem is decomposed considering first the logistics model containing quantity and logic variables and constraints in a mixed-integer linear (MILP) formulation and, secondly, the quality problem with quantity and quality variables and constraints in a nonlinear programming (NLP) model by fixing the logic results from the logistics problem. Then, stream yields of crude distillation units (CDU), for the feed tank composition found in the quality calculation, are updated iteratively in the following logistics problem until their convergence is achieved. Both local and global MILP results of the logistics model are solved in the NLP programs of the quality and an ad-hoc criteria selects to continue those among a score of the MILP+NLP pairs of solutions. A pre-scheduling reduction to cluster similar quality crude-oils decreases the discrete search space in the possible superstructure of the industrial-sized example that demonstrates our tailor-made decomposition scheme of around 3% gap between the MILP and NLP solutions.

Keywords
Scheduling optimization, Crude-oil blend-shops, Phenomenological decomposition heuristic

Introduction
An enterprise-wide optimization (EWO) problem involving scheduling operations in crude-oil refineries integrates quantity, logic and quality variables and constraints, starting in the unloading of crude-oil and ending with the delivery of fuels as schematically shown in Figure 1. However, as this is a highly complex problem to be solved, decompositions in space and in terms of solution strategies are proposed to handle such complex problem where the benefits of doing so can be in the multi-millions of dollars (Kelly and Mann, 2003a; 2003b).

Previous literature in crude-oil scheduling optimization considered models covering crude-oil unloading to the products of distillation units (Lee et al., 1996; Jia et al., 2003; Mouret et al., 2008, Castro and Grossmann, 2014). They commonly use continuous-time formulations, except for Lee et al. (1996) who solved small instances of a discrete-time approach in an MILP model using a relaxation of the bilinear blending constraints that is the major drawback as there is no guarantee of quality conservation between different outlet streams from the same crude-oil quality tank.

Advance in MILP solvers have reduced the CPU time by two orders of magnitude in comparison with the 1990’s as a consequence of progress in processing speeds and more efficient optimization algorithms. Despite this, the modeling and solution aspects of the NP-hard problems
related to the crude-oil scheduling in both continuous- and discrete-time formulations have moved away the efforts from the latter by its combinatorial complexity, and since then, a series of works have covered mostly the more compact continuous-time approaches.

![Crude-Oil Management](image1)

**Figure 1. Crude-oil refining scheduling: from crude-oils to fuels.**

Today, powerful computer memory and new modeling and algorithm structures, as those proposed in this work, enable us to address a discrete-time formulation for crude-oil blend scheduling that uses a clustering procedure to reduce the size of the scheduling problems integrated with the so-called phenomenological decomposition heuristic (PDH) that partitions MINLP models into two simpler submodels namely logistics (quantity and logic) and quality (quantity and quality) problems in an iterative strategy until convergence of both solutions (Menezes et al., 2015).

An example in crude-oil scheduling demonstrated that for industrial-sized problems the full space MINLP solution becomes intractable, although it is solved using an MILP-NLP decomposition for a gap lower than 4% between both solutions (Mouret et al., 2009). Castro and Grossmann (2014) test several problems from the mentioned literature using the Resource-Task Network (RTN) superstructure for MINLP models that are solved to near global optimality by adopting the two-step MILP-NLP algorithm, where the mixed-integer linear relaxation is derived from multiparametric disaggregation, greatly reducing the optimality gap, although the increase in the number of variables by the method limits its application to small-sized instances.

In the crude-oil refining industry, most common concerns are about the difficulties to coordinate the execution of continuous-time scheduling by the operators and how to define a priori which points to select in the future for the continuous-time calculation. Therefore, discrete-time approaches using small time-steps to solve optimization of scheduling operations in industrial-sized problems are the desired representation among 635 crude-oil refining plants in operation around the world (Oil & Gas Research Center, 2015). Concern about the possible overflow in tanks by the discretization of the time can be easily circumvented in the field by the numerous alarms and measurements that operators and schedulers have in hand in real-time typically found in the Distributed Control Systems (DCS’s).

**Problem Algorithm and Pre-Scheduling Reduction**

Scheduling models integrating quantity, logic, and quality variables and constraints give rise to non-convex MINLP models. Limitations in the solution of large-scale problems mainly occurs in the relaxed NLP steps. Hence, the proposed model uses the PDH algorithm as seen in Fig. 2 that resembles Benders decomposition where the logic variables from the logistics problem, by neglecting the nonlinear blending constraints, are fixed such that a simpler program may be solved in the quality problem. Then, CDU stream yields for the crude-oil feed diet are found in the NLP model and updated in the following MILP solution in a new PDH iteration until their convergence within a PDH gap tolerance and similar logic results in consecutive iterations.

To decrease the discrete search space, there are two layers of clustering to segregate crude-oils with similar quality (Kelly et al., 2017). This is especially needed in industrial-sized problems. Besides, the crude-oil composition in initial inventories for the current selection of feed tanks generates initial CDU yields for the logistics problem using ±5% tolerance between the main distillates as in this stage the yields are the same despite the quality of the crude-oils.

![Crude-Oil Blend Scheduling Optimization](image2)

**Figure 2. Proposed Reductions and PDH Algorithm.**

**Crude-Oil Blend Scheduling Optimization**

The problem consists of determining the crude-oil blend scheduling involving crude-oil supply, storage and feed tank operations and the CDU production. Figure 3 shows a pair of marine vessels or feedstock tanks (CR1 and CR2) supplying a crude oil refinery with different quality raw materials to produce the main CDU distillates: fuel gas (FG), liquid petroleum gas (LPG), light naphtha (LN), heavy naphtha (HN), kerosene (K), light diesel (LD), heavy diesel (HD) and atmospheric residuum (AR).

A swing-cut unit (SW) is modeled to give a certain degree of fractionation for LN and HN applying not only a quantity variation, but including quality recalculation since the lighter and heavier SW splits have different qualities with respect to their amounts and blended crude-oil distillation curve (Menezes et al., 2013; Kelly et al., 2014). Storage tanks (S1 to S4) are connected to the feed tanks (F1...
to F3) using a crude-oil blender (COB) as indicated for improved performance of CDU operations by minimizing crude-oil composition disturbances in further real-time optimization and model predictive control.

**Figure 3. Crude-Oil Blend Scheduling.**

The network in Figure 3 is constructed using the unit-operation-port-state superstructure (UOPSS) formulation given by the objects and their connectivity involving arrows (→), in-ports (⊙), and out-ports (⊗). Unit-operations and arrows have binary $y$ and continuous $x$ variables, and the ports hold the states for the relationships among the objects, adding more continuous variables if necessary by the semantic and meaningfully configured programs.

In problem (P), the objective function (1) maximizes the gross margin from fuels revenues subtracting the performance of the CDU throughputs, giving by the deviation from the quantity in the previous time-period against the current time-period, minimizing the 1-norm or linear deviation of the flow in consecutive time-periods. This performance term smooths the CDU throughputs considering the lower $x_{\text{LOD}}$ and upper $x_{\text{UPD}}$ deviation of their adjacent amounts ($m \in \text{MCU}$). If $x_{m,t} + 1 \leq x_{m,t} \Rightarrow x_{\text{LOD},t} = x_{m,t} - x_{m,t+1}$ and $x_{\text{UPD},t} = 0$. If $x_{m,t} \geq x_{m,t} + 1 \Rightarrow x_{\text{UPD},t} = x_{m,t+1} - x_{m,t}$ and $x_{\text{LOD},t} = 0$. These deviation variables are set as the same as the bounds of the CDU throughputs, i.e., $0 \leq x_{\text{LOD},t} \leq x_{\text{UPD},t}$ and $0 \leq x_{\text{UPD},t} \leq x_{\text{UPD},t}$. The smoothing relationship for the CDU flow in Eq. (2) is satisfied if $x_{m,t+1} = x_{m,t}$: then $x_{m,t} = x_{\text{LOD},t} = 0$. Unit-operations $m$ for tanks, blenders and fuels belong, respectively, to the following sets: $M_{\text{TK}}, M_{\text{BL}}$ and $M_{\text{FU}}$. The UOPSS formulation given by the objects and their connectivity as in Figure 3 are specified in Eqs. (3) to (14). In the summations involving ports in Eqs. (6) to (9), (12) and (13), $j'$ and $i''$ represent upstream and downstream ports connected, respectively, to the in-port $i$ and out-port $j$ of unit-operations $m$. The set $U_m$ represents the unit-operations $m$ within the same physical unit. For $x \in \mathbb{R}^+$ and $y = \{0,1\}$:

$\text{(P) } \max Z = \sum_{t} \left( \sum_{m \in M_{\text{FU}}} \text{price}_{m,t} x_{m,t} - \sum_{m \in M_{\text{CU}}} \text{weight}(x_{\text{LOD},m,t} + x_{\text{UPD},m,t}) \right) \tag{1}$
Equations (12) is the quantity balance to control the inventory or holdup for unit-operations of tanks \((m \in M_{TK})\). The equality constraint calculates the current holdup amount \(x_{m,t}\) considering the material left in the past time-period (heals) plus and minus the summation of, respectively, the upstream and downstream connections to the tanks. Equation (13) is a material balance in fractionation columns \(M_{CDU}\) and blenders \(M_{BL}\) to ensure that there is no accumulation of material in these types of units.

It should be mentioned that quantity balances for in- and out-port-states is guaranteed because the UOPSS formulation does not perform explicit material balances for port-states, but only for unit-operations, so the flow of connected port-states are bounded by their lower and upper bounds. Stream flows involving ports are only for unit-operation-port-state to unit-operation-port-state (arrow streams). For quality balances, only in-port-states have explicit material balances since these are uncontrolled mixers. For out-port-states, which are uncontrolled splitters, the UOPSS does not create explicit splitter equations for the qualities because these are redundant with the value of the intensive property of upstream-connected unit-operations.

**Logistics Problem: MILP Crude-Oil Blend Scheduling**

The logistics problem includes Eqs. (1) to (14) and Eqs. (15) to (29) that involves: a) unit-operations in temporal transitions of sequence-dependent cycles, multi-use of objects and up-time or minimal time of their using, and zero downtime of an equipment, and b) tanks in fill-draw delay and fill-to-full and draw-to-empty operations.

The operation of the semi-continuous blender COB in Figure 3 is controlled by the temporal transition constraints (13) to (15) from Kelly and Zyngier (2007). The setup or binary variable \(y_{m,t}\) manages the dependent start-up, switch-over-to-itself and shut-down variables \(\overline{z}_{su,m,t}\), \(\overline{z}_{sw,m,t}\), and \(\overline{z}_{dw,m,t}\), respectively that are relaxed in the interval \([0,1]\) instead of considering them as logic variables. Equation (17) is necessary to guarantee the integrality of the relaxed variables.

\[ y_{m,t} - y_{m,t-1} - \overline{z}_{su,m,t} + \overline{z}_{sw,m,t} = 0 \quad \forall \ m \in M_{BL}, t \]  
\[ y_{m,t} + y_{m,t-1} - \overline{z}_{su,m,t} - \overline{z}_{dw,m,t} = 0 \quad \forall \ m \in M_{BL}, t \]  
\[ \overline{z}_{su,m,t} + \overline{z}_{sw,m,t} \leq 1 \quad \forall \ m \in M_{BL}, t \]  

In the multi-use procedure in Eq. (18), the lower and upper parameters \(USE_{U,t}^L\) and \(USE_{U,t}^U\) coordinate the use of the out-ports \((j \in I_{USE})\) by their connected downstream in-ports \(i\). This occurs in blenders to avoid the split of a mixture to different feed tanks at the same time. Equation (19) imposes \(USE_{L,t}^L\) and \(USE_{U,t}^U\) in the in-ports \((i \in I_{USE})\) by their connected upstream out-ports \(j\). It controls the maximum number of transfers within the same time-period to CDU. Equations (20) to (22) model the run-length or up-time considering \(UPT_U\) as the upper bound of using and \(t_{end}\) as the end of the time horizon with \(\Delta t\) as time-step, where Eq. (22) is the unit-operation up-time temporal aggregation cut constraint and number of period is \(n_t\); more details on these constraints can be found in Kelly and Zyngier (2007) and Zyngier and Kelly (2009). Equation (23) is the zero downtime constraint for the CDU to select at least one mode of operation \(m\) to be continuously operating.

\[ \frac{1}{USE_{U,t}^L} \sum_{j \in I_{USE}, t} y_{j,i''},t \leq \frac{1}{USE_{U,t}^U} \sum_{j \in I_{USE}, t} y_{j,i''},t \quad \forall \ j \in I_{USE}, t \]  
\[ \frac{1}{USE_{L,t}^U} \sum_{j \in I_{USE}, t} y_{j,i'},t \leq \frac{1}{USE_{L,t}^L} \sum_{j \in I_{USE}, t} y_{j,i'},t \quad \forall \ t \in I_{USE}, t \]  
\[ \overline{z}_{su,m,t} + \overline{z}_{sw,m,t-1} = y_{m,t+1} \quad \forall \ m \in M_{BL}, t > 1 \]  
\[ UPT_U \Delta t \sum_{t=t_{st}}^{t_{end}-UPT_U} y_{m,t} \leq U_{BL} t < t_{end} - UPT_U \]  
\[ \Delta t \sum_{t} \overline{z}_{su,m,t} \leq n_p \quad \forall \ m \in M_{BL} \]  
\[ \sum_{m \in M_{CDU}} y_{m,t} \geq 1 \quad \forall \ t \]  

For the operation of tanks in Eqs. (24) to (28), the formulation is found in Zyngier and Kelly (2009). The fill-draw delay constraints (24) and (25) controls, respectively, the minimum and maximum time between the last filling and the following drawing operations for the upstream \(j\) and downstream \(i''\) connections of a tank. The minimum and maximum fill-draw delays are \(\Delta d_{min}\) and \(\Delta d_{max}\), respectively, and the in- and out-ports \(i\) and \(j\) are connected to the tank \((I \in M_{TK})\).

\[ y_{j,i''},t + y_{j,i'''},t \leq 1 \quad \forall \ (j', i, i'') \text{ such that } I_j \in M_{TK}, \]  
\[ tt = 0. \Delta d_{min}, t = 1..t + tt < t_{end} \quad (24) \]  
\[ y_{j,i'},t - y_{j,i''},t - \sum_{t'=t_{st}}^{t_{end}} y_{j,i''},t' \leq 0 \quad \forall \ (j', i, i'') \]  
\[ \text{such that } I_j \in M_{TK}, t = 1..t - \Delta d_{max}, t + tt < t_{end} \quad (25) \]

The remaining tank operations are the fill-to-full and draw-to-empty constraints. They add the logic variable \(y_{d,m,t}\), representing the filling and drawing operations, which are equal to zero if the pool is filling and one if it is drawing, avoiding the use of two logic variables. The coefficients \(x_{\overline{h}_{m,t}}^{\overline{h}_{m,t}}\) and \(x_{\overline{h}_{m,t}}^{\overline{h}_{m,t}}\) are, respectively, the fill-to-full and draw-to-empty to force the tank to be filled and drawn to their values.

\[ y_{j,i'},t + y_{d,m,t} \leq 1 \quad \forall \ (j', i, m) \text{ for } m \in M_{TK}, t \quad (26) \]  
\[ y_{j,i''},t \leq y_{d,m,t} \quad \forall \ (m, j, i'') \text{ for } m \in M_{TK}, t \quad (27) \]  
\[ x_{m,t} - \overline{x}_{m,t}^{\overline{h}_{m,t}} (y_{d,m,t} - y_{d,m,t-1}) + \overline{x}_{m,t}^{\overline{h}_{m,t}} (\overline{x}_{m,t}^{\overline{h}_{m,t}} - \overline{x}_{m,t}^{\overline{h}_{m,t}}) \geq 0 \quad \forall \ m \in M_{TK}, t \quad (28) \]  
\[ x_{m,t} + \overline{x}_{m,t}^{\overline{h}_{m,t}} (y_{d,m,t} - y_{d,m,t-1}) - \overline{x}_{m,t}^{\overline{h}_{m,t}} (\overline{x}_{m,t}^{\overline{h}_{m,t}} + \overline{x}_{m,t}^{\overline{h}_{m,t}}) \leq 0 \quad \forall \ m \in M_{TK}, t \quad (29) \]
Quality Problem: NLP Crude-Oil Blend Scheduling

The quality problem includes Eqs. (1) to (9) (by fixing the binary results from the logistics), Eq. (12) for quantity balances in tanks, Eq. (13), the continuous part of Eq. (14) and the nonlinear Eqs. (30) to (35). The blending or pooling constraints involves volume-based quality balances in crude-oil components and specific gravity (density) and weight-based for sulfur concentration. They apply for: a) flow in unit-operations (except for fractionators) and in-ports; and b) holdup in tanks. The other nonlinear constraints are the transformations from crude-oil components to compounds (distillate amounts and properties) in fractionators as in Eqs. (33) to (35). See Kelly and Zyngier (2016) for the case of nonlinear equations.

Considering $p$ as the component-property (crude-oil component, specific gravity or sulfur concentration) and $v$ and $w$, respectively, volume- and weight-based properties, Eqs. (30) calculates the volume-based balance in unit-operations (in the case, only for blenders) and Eq. (31) in in-ports. When the in-ports are connected to tanks, their quality balances are redundant to Eq. (32) valid only for tanks, so that Eq. (31) is only true for in-ports not connected to tanks. It should be mentioned that there is no need of a quality variable for out-ports of a unit-operation. Instead we can use the quality of their unit-operation $m'$ itself.

\[
v_{m,p,t} \sum_{j} x_{j,i,t} = \sum_{j} v_{m',p,t} x_{j',i,t} \forall m, p, t \quad (30)
\]

\[
v_{j,p,t} \sum_{i} x_{j,i,t} = \sum_{j} v_{m',p,t} x_{j',i,t} \forall j, p, t \quad (31)
\]

\[
v_{m,p,t} x_{h,m,t} = v_{m,p,t-1} x_{h,m,t-1} + \sum_{j} v_{m',p,t} x_{j',i,t} - v_{m,p,t} \sum_{j} x_{j,i',t} \forall (i,j) \in M_{TK}, t \quad (32)
\]

The weight-based balances are skipped as they are similar to Eqs. (30) to (32) only replacing $v$ by the product $vw$. Finally, Eq. (33) converts CDU throughputs in amounts or yields of distillates ($j \in DIS$). Equations (34) and (35) calculate, respectively, the volume- and weight-based properties for $J_{DIS}$ considering the assay or renderings of each crude-oil $c$ with respect to their defined cuts. The renderings for yields and properties are $r_{j,c,cut}$ and $r_{j,c,cut}^p$ and SG means specific gravity in Eq. (35). Reformulating Eqs. (34) and (35), Eq. (33) can be substituted to cancel the term $\sum_{j'} X_{j',c,cut}$ and reduces both the nonlinearities and the number of non-zeros in these equations.

\[
\sum_{i} x_{j,i',t} = \sum_{j'} x_{j',c,cut} \sum_{c} \sum_{cut} v_{c,cut}^{p=c} \epsilon_{j,c,cut}^{yie} \forall j \in DIS, t \quad (33)
\]

\[
v_{j,p,t} \sum_{i} x_{j,i',t} = \sum_{j'} x_{j',c,cut} \sum_{c} \sum_{cut} v_{c,cut}^{p=c} \epsilon_{j,c,cut}^{yie} \epsilon_{j,c,cut}^{p} \forall j \in DIS, t \quad (34)
\]

Illustrative Example

The examples use the structural-based unit-operation-port-state superstructure (UOPSS) found in the semantic-oriented platform IMPL (Industrial Modeling and Programming Language) using Intel Core i7 machine at 2.7 Hz with 16GB of RAM. Figure 4 shows the unit-operation Gantt chart for the entire problem found in Figure 3. The past/present time-horizon has a duration of 24-hours and the future time-horizon is 336-hours discretized into 2-hour time-period durations (168 time-periods). The logistics problem has 8,333 continuous and 3,508 binary variables and 3,957 equality and 15,810 inequality constraints and it is solved in 176.0 seconds using 8 threads in CPLEX 12.6. The quality problem has 19,400 continuous variables and 14,862 equality and 696 inequality constraints and lasts 16.8 seconds in the IMPL’ SLP engine linked to CPLEX 12.6. The MILP-NLP gap between the two solutions is within 0.09% with only one PDH iteration.

The crude-oil blend header COB has an up-time or run-length of between 6 to 18-hours and this is clearly followed by the semi-continuous nature of the blender (i.e., see the 20 blends or batches of crude-oil mixes). The storage tanks have a lower fill-draw-delay of 6-hours which means that if there is a fill into the tank then there must be at least a 6-hour delay or hold before the draw out of the tank can occur. This is typical of receiving tanks in an oil-refinery when they receive crude-oil cargos from marine vessels that contain a significant amount of ballast water that needs to be decanted before it charges the desalter, preflash, furnace and CDU.

![Figure 4. Illustrative example Gantt chart.](image-url)
to or above 190.0 Kbbbl. This is necessary for managing tank inventories so that they have a well-defined fill-hold-draw profile. These restrictions are also necessary to minimize the crude-oil swings or runs to the CDU given that during the draw out of a single feed tank to a CDU, there are no crude-oil disturbances; i.e., between one to three days depending on the CDU charge rate and the inventory capacity of the charge tank.

Figure 5 plots the composition of crude-oils entering the feed tank F1 on its in-port. As can be easily seen, the compositions change when the blender charge crude-oil from the storage tanks.

![Figure 5. F1 in-port component plots.](image)

Industrial-Sized Example

The proposed model is applied in an industrial-sized refinery including 5 crude-oil distillation units (CDU) in 9 modes of operation and around 35 tanks among storage and feed tanks. The past/present time-horizon has a duration of 48-hours and the future time-horizon is 168-hours discretized into 2-hour time-period durations (84 time-periods). The logistics problem has 30,925 continuous and 29,490 binary variables and 6,613 equality and 79,079 inequality constraints (degrees-of-freedom = 53,802) and it is solved in 128.8 seconds using 8 threads in CPLEX 12.6. The quality problem has 102,539 continuous variables and 58,019 equality and 768 inequality constraints (degrees-of-freedom = 44,520) and lasts 10.3 minutes in the IMPL SLP engine linked to CPLEX 12.6. The MILP-NLP gap between the two solutions is within 3.5% after two PDH iterations.

Conclusion

In summary, we have highlighted a benchmark application of crude-oil blend scheduling optimization using a discrete, nonlinear and dynamic optimization with a uniform time-grid (see Menezes et. al. 2015 for more details). The fine points of the MINLP formulation are highlighted where a phenomenological decomposition of logistics and quality is applied to solve industrial-sized problems to feasibility. An illustrative example is modeled and solved to demonstrate the theory in practice.

References


