FEEDSTOCK STORAGE ASSIGNMENT IN PROCESS INDUSTRY QUALITY PROBLEMS

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

Abstract

We propose a mixed-integer linear (MILP) model for the design of assignments of various raw materials with different qualities when moving them from external supply sources to shared storages. This is especially important in process industries with limited storage and quality blend programs optimizing a plant feed diet for ongoing operations involving process units, inventory control and product demands, as found in crude-oil, ore/metal and food processing industries. This novel storage assignment problem minimizes the quality deviation when a larger number of feedstocks from marine vessels or ships are clustered into a smaller number of containers or storages in the plant, known as the Pigeonhole Principle, allocating the raw material to a definite place in an orderly system. Although the model only uses raw material quality data and neglects logistics details such as raw material supply amounts, timing and volume available in the storage, the simplification can be partially circumvented by splitting the raw material into two or more species with same qualities in order to fit into the storages. Examples dealing with 5 to 45 different crude-oil feedstocks clustered into 4 storage tanks demonstrate the proposed model, which yields the optimum storage assignment within minutes for industrial-scale problems.

Keywords

Storage assignment, Raw material clustering, Sharing of storage, Segregation rules, Pigeonhole principle.

Introduction

Process industry raw material differing in quality (i.e., percentage of its components as well as its properties) can yield different quantity of products with distinct properties as found in crude-oil, ore/metal, and food processing. The raw materials in these industries are supplied from reservoirs, mines and farms, respectively, varies depending on geological or agricultural conditions. The experimental analysis of their components and properties characterized in numbers (assay) are used in the modeling and optimization/simulation steps of the process operations, handling normally molar-, volume-, or mass-based values of the distinct raw materials. Crude-oil is represented by distillation curves of hydrocarbon molecules from methane to asphalt fractions; crushed stones in concentration of minerals with iron, nickel, copper, gold, aluminum, etc.; and fruits in terms of sugar content variants as fructose, glucose, sucrose, etc.

As known from industrial practice, the majority of the crude-oil refineries partition their raw materials for the scheduling operations based on its bulk specific gravity (light-to-heavy) and sulfur (sweet-to-sour) properties, which may not be the most appropriate criteria as explained in detail in Kelly and Forbes (1998) by considering the other compositional aspects of the crude-oil characteristics such as naphtha and diesel contents. As an improvement to the common sense rules for segregating raw material differing in quality, we present a new formulation of how to design the assignment of raw materials (feedstocks) to storage

* corresponding author: brennocm@andrew.cmu.edu

using the values of the raw material compound-properties or assay in order to minimize of the overall quality variance during the storage assignment. Ultimately, better segregation or clustering rules mean better scheduling and control of the raw material blendshops for both continuous (using on-line blenders) and batch mixtures, which can translate into significant savings per year for crude-oils as discussed by Kelly and Mann (2003a; 2003b).

When the raw material quality (based on their component-properties) determines process operations and product amounts and properties, the groups of similar quality raw material gathered in shared storages can be mixed again to obtain an optimal plant feed quality in the final blend. This potentially improves the production flexibility within a larger optimization search space if the assignment of raw material from feedstocks to storage minimizes their quality variance, although there is no guarantee of global optimality when solving the nonlinear (NLP) blending problem for the mixed storage raw material as this is a well-known non-convex problem.

We propose a mixed-integer linear (MILP) optimization model to categorize, place or separate raw material by grouping together those with similar qualities into the same cluster/storage. This membership or cardinality problem for clustering of raw material uses exact search in an MILP formulation. We also show the results of the heuristic searches as found in k-means clustering (KM) and fuzzy c-means clustering (FCM) algorithms (Bezdek et al., 1984) for the sake of comparison. It should be mentioned that clustering on latent variables or scores found in Principal Component Analysis (PCA) and Partial Least Squares (PLS) may also be applied (Ashe et al., 1997).

The clustering problem as addressed in this work considers only raw material quality data and neglects logistics details as amount of raw material to be transferred and holdup or heel in the storage. However, this can be bypassed by knowing in advance the ship or terminal lot-sizes and the available storage in the field, to separate the raw material in many others with the same quality as needed. By handling the raw material in an intelligent manner by pre-defining the storage assignment for further down-stream blend scheduling programs, the reduced-scope scheduling permits to explore discrete-time formulations with short time-steps in problems with hundreds of time-periods in highly complex process plants. The proposed model is applied to crude-oils in Menezes et. al. (2017) with an industrial-sized example including 5 crude-oil distillation units (CDU) in 9 modes of operations and around 35 tanks among storage and feed tanks.

Problem Statement

In Figure 1, a given raw material *RM* is assigned to the storage group *ST-CL*, represented with 4 modes of operations (*ST-CL1*, *ST-CL2*, *ST-CL3*, *ST-CL4*), of which only one mode can be setup or active for any given raw material. In the *ST-CL* groups inside the dotted rectangle, *ST* means storage to assign the raw material and *CL* is its mode to define the clustering groups or *RM* destinations. It is used instead of creating 4 physical places to assign the *RM*. The *CL* unit outside the *ST* group, is the hypothetical cluster where the compound-properties of the assigned *RM* to *ST-CL* is saved to be compared in the assignment-clustering procedure as seen in the following sections.

The circle structures with and without cross-hairs represent the compound-properties in the out (⊗) and in-port-states (○), respectively. For example, these compound-properties specific to crude-oils are typically whole-crude-oil specific-gravity, sulfur, etc. and individual compounds such as light-ends, naphtha, kerosene, jet-fuel, diesel, gas-oils, and their properties such as distillation temperatures, volatilities, viscosities, and contaminant concentrations.

In Figure 1, the compound-properties in the in-port sets I_{ST} ={i_1, i_2, i_3, i_4} for each mode *m* of *ST* (*ST-CL1*, *ST-CL2*, *ST-CL3*, *ST-CL4*) are connected to out-ports *J*={j_1, j_2, j_3, j_4} belonging to *RM* (J_{RM}) and to the respective cluster unit *CL* (J_{CL}). The number of compound-properties is typically taken as the same number of clusters for controllability reasons where the number of clusters corresponds to the number of raw material storages available. In this work, the raw material is crude-oil and the compound-properties are naphtha-yield (NY), diesel-yield (DY), diesel-sulfur (DS) and residue-yield (RY), although in Table 1 we also provide the crude-oil bulk specific-gravity (WCSG) and sulfur (WCSUL). These assays are taken from the ExxonMobil database (ExxonMobil , 2015).

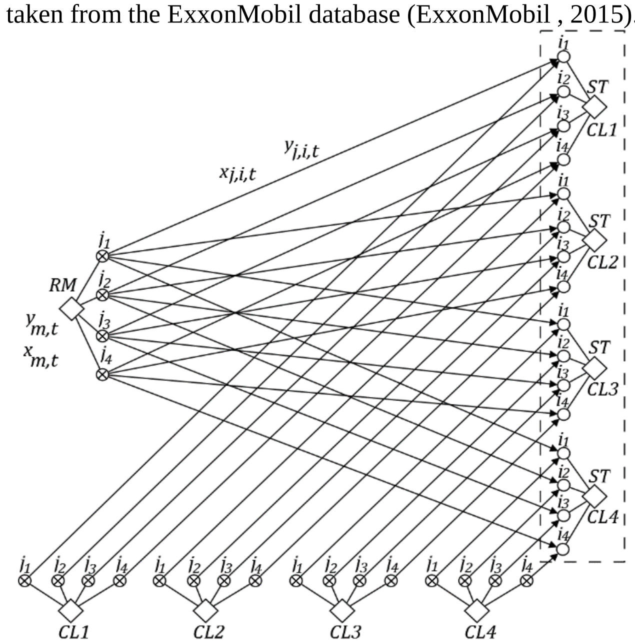

Figure 1. Storage assignment flowsheet in UOPSS.

We use the structural-based unit-operation-port-state superstructure (UOPSS) (Kelly, 2005; Zyngier and Kelly, 2012), where perimeter-units (◇) and arrows (⟶) shapes have binary variables *y* (setups) and continuous *x* (flows) and the ports hold the states for the relationships among the shapes, adding more continuous variables by the semantic and meaningfully configuration of the programs.

Each raw material *RM* arriving in the storage *ST* is considered as a time-step *t* of a discrete-time formulation, although it represents, for instance, the raw material sequence of transferring previously defined in the terminals by taking into account the supply arrivals or any preparation time of the raw material in transit.

Table 1. Crude-oil assay data.

Crude	NY	DY	DS	RY	WCSG	WCSUL
1	19.57	18.02	0.19	9.56	0.8265	0.3137
2	14.08	16.01	0.53	18.26	0.8686	0.9580
3	32.88	14.34	0.15	2.97	0.7826	0.1493
4	13.36	20.84	0.08	12.95	0.8468	0.1475
5	10.05	20.12	0.24	19.00	0.8895	0.5780
6	14.43	24.84	0.20	7.96	0.8745	0.2350
7	7.72	10.31	1.97	35.39	0.9358	3.7708
8	13.60	18.95	0.15	15.46	0.8642	0.2444
9	27.64	13.61	0.14	5.32	0.7909	0.1438
10	5.35	20.58	0.24	25.13	0.9153	0.5020
11	2.26	13.46	0.04	39.51	0.9153	0.0880
12	17.02	24.36	0.16	5.94	0.8478	0.1699
13	17.02	24.36	0.16	5.94	0.8478	0.1699
14	27.02	15.37	0.09	1.53	0.7616	0.0591
15	20.11	18.74	0.16	9.30	0.8358	0.2398
16	4.26	21.19	0.32	24.62	0.9330	0.8059
17	10.87	19.36	0.16	18.64	0.8767	0.3330
18	15.28	17.06	0.21	16.57	0.8509	0.4578
19	15.79	16.49	0.79	16.18	0.8628	1.2415
20	12.29	15.96	0.31	21.85	0.8855	0.6410
21	17.31	18.01	0.09	8.61	0.8203	0.1891
22	13.32	16.88	0.20	19.92	0.8718	0.3733
23	4.34	10.17	1.83	33.26	0.9324	3.8025
24	35.04	14.47	0.03	1.81	0.7848	0.0277
25	27.82	13.94	0.09	6.69	0.8013	0.1295
26	14.18	16.43	0.21	20.29	0.8800	0.4030
27	44.42	10.82	0.02	0.08	0.7366	0.0039
28	18.75	18.32	0.16	10.13	0.8358	0.2493
29	7.79	19.60	0.19	21.51	0.9002	0.3665
30	18.76	22.98	0.10	5.64	0.8423	0.1110
31	16.17	18.02	0.14	14.23	0.8519	0.2454
32	45.91	3.91	0.02	0.20	0.7347	0.0025
33	41.45	5.28	0.17	0.04	0.7335	0.0400
34	22.16	21.60	0.24	4.55	0.8404	0.2725
35	18.74	17.66	0.15	10.24	0.8284	0.2511
36	19.23	23.13	0.04	5.18	0.8123	0.0436
37	32.47	2.20	0.03	0.11	0.6916	0.0020
38	15.18	16.39	0.50	19.07	0.8660	0.9898
39	15.69	17.17	0.23	16.43	0.8539	0.5250
40	15.98	18.75	0.16	13.32	0.8446	0.3237
41	19.65	19.99	0.10	8.58	0.8458	0.1723
42	14.98	16.13	1.26	16.91	0.8555	1.8393
43	16.65	20.73	0.17	11.12	0.8729	0.2280
44	12.62	16.28	1.13	18.24	0.8827	1.7500
45	21.89	21.54	0.10	2.74	0.8246	0.0947

The proposed model does not calculate the nonlinear mixing in the shared storage. Instead, it uses the clusters to maintain the values of the compound-properties in their out-ports J_{CL} during all storage assignment design. When each raw material in *RM* is sent to the storage modes *ST-CL*, sequentially, within each time-period *t*, the values of the J_{CL} out-ports are compared to the values in the I_{ST} in-ports, and then the storage assignment or clustering of the *RM* to the *ST*,*CL* (storage-cluster) mode is determined. Therefore, although the *RM* transfer occurs at every time-period, a multi-period (multi *RM*) case should be considered as compound-properties of all raw materials assigned to the storage are compared with the cluster compound-properties as explained in the following.

Quality Variance Minimization in Shared Storages

In problem (*P*), which includes the clustering objective and target constraints plus the UOPSS flowsheet formulation, the objective function (1) minimizes the 1-norm variance of the raw material quality $(x_{i,t}^{LOD} + x_{i,t}^{UPD})$ in the in-ports *i* of the storage-operations *ST-CL* (I_{ST}) when several raw materials *RM* with different qualities are transferred to *ST-CL* in each time-period *t*. In the summation involving in-ports *i* and out-ports *j*, in Eqs. (2), (4) and (5) and the others in the next section, the ports are those connected as in Figure 1 and we omit their subsets in the indices for the sake of simplicity. For $x \in \mathbb{R}^+$ and $y = \{0,1\}$:

$$(P)\ \ Min\ Z = \sum_{i \in I_{ST}} \sum_{t} weight_{i,t} \left(x_{i,t}^{LOD} + x_{i,t}^{UPD}\right) \tag{1}$$

s.t.

$$\sum_{j} x_{j,i,t} - \bar{x}_{i,t} + x_{i,t}^{LOD} - x_{i,t}^{UPD} = 0\ \ \forall\ i \in I_{ST}, t \tag{2}$$

$$\bar{x}_{j,i,t}^{L}\ y_{j,i,t} \le x_{j,i,t} \le \bar{x}_{j,i,t}^{U}\ y_{j,i,t}\ \ \forall\ (j,i), t \tag{3}$$

$$\frac{1}{\bar{x}_{m,t}^{U}} \sum_{j} x_{j,i,t} \le y_{m,t} \le \frac{1}{\bar{x}_{m,t}^{L}} \sum_{j} x_{j,i,t}\ \ \ \forall\ (i,m), t \tag{4}$$

$$\frac{1}{\bar{x}_{m,t}^{U}} \sum_{i} x_{j,i,t} \le y_{m,t} \le \frac{1}{\bar{x}_{m,t}^{L}} \sum_{i} x_{j,i,t}\ \ \ \forall\ (m,j), t \tag{5}$$

$$y_{m',t} + y_{m,t} \ge 2 y_{j,i,t}\ \ \forall\ (m', j, i, m), t \tag{6}$$

$$x_{j,i,t}, x_{i,t}^{LOD}, x_{i,t}^{UPD} \in \mathbb{R}^+;\ y_{j,i,t}, y_{m,t} = \{0,1\} \tag{7}$$

The quality variance variables $x_{i,t}^{LOD}$ and $x_{i,t}^{UPD}$ are the lower and upper deviation of the raw material quality target $\bar{x}_{i,t}$ in $i \in I_{ST}$. If the quality value in the *ST* in-ports (the sum of the arrows arriving in) is $\sum_j x_{j,i,t} \le \bar{x}_{i,t} \Rightarrow x_{i,t}^{LOD} = \bar{x}_{i,t} - \sum_j x_{j,i,t}$ and $x_{i,t}^{UPD} = 0$. If $\sum_j x_{j,i,t} \ge \bar{x}_{i,t} \Rightarrow x_{i,t}^{UPD} = \sum_j x_{j,i,t} - \bar{x}_{i,t}$ and $x_{i,t}^{LOD} = 0$. The non-negative target variable bounds are set as the same as in the in-port bounds $0 \le x_{i,t}^{LOD} \le \bar{x}_{i,t}^{U}$ and $0 \le x_{i,t}^{UPD} \le \bar{x}_{i,t}^{U}$. The storage clustering relationship with respect to target variables in Eq. (2) forces the match to the target $\bar{x}_{i,t}$ in the in-ports $I_{ST} = \{i_1, i_2, i_3, i_4\}$ that is satisfied if $\sum_j x_{j,i,t} = \bar{x}_{i,t}$, and then $x_{i,t}^{LOD} = x_{i,t}^{UPD} = 0$. This is the key aspect in the clustering algorithm, where the values of the quality in the in-ports of the storages *ST* are minimized to a target $\bar{x}_{i,t}$ considering: a) the value of the transferred raw material in *t* given by $x_{j,i,t}$ for *j* in *RM*, and b) the value of the quality in the clusters given by $x_{j,i,t}$ for *j* in *CL*. It should be noticed that there are no variables for the total amounts in the in-ports. Instead we have a relational procedure involving the summation of the continuous variables $x_{j,i,t}$ there (the values of the compound-properties of the raw material *RM* and cluster *CL* transferred to *ST*,*CL* at time *t*).

Equations (3) to (6) are the main UOPSS constraints to control: a) the setups of perimeter unit-operations $y_{m,t}$ and arrows $y_{j,i,t}$, and b) the flows in the ports. Equation (3) is the semi-continuous constraint for the arrows between the ports. When the binary $y_{j,i,t}$ related to the arrow stream is setup, its value varies between bounds ($\bar{x}^L_{j,i,t}$ and $\bar{x}^U_{j,i,t}$). Equations (4) and (5) are the in-port *i* and out-port *j* balances. If the binary related to the unit-operation *m* in the perimeter is true (i.e., $y_{m,t}$ = 1), then the sum of the arrows arriving in the in-ports or leaving from the out-ports are between their perimeter unit-operation bounds ($\bar{x}^L_{m,t}$ and $\bar{x}^U_{m,t}$). Equation (6) is the structural transition interconnecting two unit-operations *m* and *m'* from different perimeters It says that if the setup variable of these unit-operations are turned-on (i.e., $y_{m,t}$ = 1 in *u* and $y_{m',t}$ = 1 in *u'*), the setup of the arrow stream between these unit-operations is turned-on by implication ($y_{j,i,t}$ = 1). It forms a group of 4 objects (*m', j, i, m*) with a logic valid cut that reduces the tree search in branch-and-bound methods.

Procedural Clustering of Raw Material

The full clustering problem (*CLP*) replaces the problem (*P*) using the weight in the objective function (8) similar to the one found in the KM and FCM clustering literature (Bezdek et. al., 1984) by defining the highest (UB_i) and lowest (LB_i) values among the compound-properties to be clustered. In Table 1, we select 4 compound-properties (raw material qualities, parameters or assay) given by pr_j or pr_i (NY, DY, DS, RY) that are represented in Figure 1 by the *J* and *I* sets of the ports.

The target values $\bar{x}_{i,t}$ in the in-ports I_{ST} are set to zero reducing Eq. (2) to Eq. (9) that is satisfied if the values of the qualities in the *RM* and *CL* out-ports are the opposite, since they vary with respect to the constraints (10) and (11) (both opened from Eq. (3) in *P*). When the assignment decision occurs, the binary variables $y_{j,i,t}$ of the *RM* and *CL* connected to *ST* should be active to the respective *ST* unit-operation (*ST*,*CL*). If there is no sharing of storage among the raw material to be transferred, Eq. (9) is satisfied and the lower and upper quality deviation variables in the objective function (8) are zero. If more than one of these raw material is clustered in the same storage, the continuous variables $x_{j,i,t}$ of the *RM* and *CL* connected to *ST* are different and a deviation of the quality in the I_{ST} in-ports is computed.

The set $JI_{RM\text{-}ST}$ contains the arrows connecting the raw material *RM* out-ports to the storage *ST* in-ports. Similarly, $JI_{CL\text{-}ST}$ connects cluster out-ports to the same *ST* in-ports in the storage-cluster modes M_{ST}. Using the defined $UB_i = max_i(pr_i)$ and $LB_i = min_i(pr_i)$ and the remaining constraints as the multi-use logic, zero downtime and the flowing equaling, we have the following for the full clustering problem *CLP* in Eqs. (8) to (21).

$$(CLP)\ \ Min\, Z = \sum_{i\in ST}\sum_{t}\frac{1}{UB_i - LB_i}\left(x^{LOD}_{i,t} + x^{UPD}_{i,t}\right) \tag{8}$$

$s.t.$

$$\sum_{j} x_{j,i,t} + x^{LOD}_{i,t} - x^{UPD}_{i,t} = 0\ \ \forall\ i \in I_{ST}, t \tag{9}$$

$$0 \le x_{j,i,t} \le UB_i\ y_{j,i,t}\ \ \forall\ \ (j,i) \in JI_{RM-ST}, t \tag{10}$$

$$-UB_i\ y_{j,i,t} \le x_{j,i,t} \le 0\ \ \forall\ \ (j,i) \in JI_{CL-ST}, t \tag{11}$$

$$\frac{1}{UB_i}\sum_{j} x_{j,i,t} \le y_{m,t} \le -\frac{1}{UB_i}\sum_{j} x_{j,i,t}\ \ \forall\ m \in M_{ST}, i \in I_{ST}, t \tag{12}$$

$$\frac{1}{pr_{j,t}}\sum_{i} x_{j,i,t} \le \bar{y}_{m,t} \le \frac{1}{pr_{j,t}}\sum_{i} x_{j,i,t}\ \ \forall\ m \in M_{RM}, j \in J_{RM}, t \tag{13}$$

$$\bar{y}_{m',t} + y_{m,t} \ge 2y_{j,i,t}\ \forall\ m' \in (M_{RM} \vee M_{CL}), m \in M_{ST},\ (j,i) \in (JI_{RM-ST} \vee JI_{CL-ST}), t \tag{14}$$

$$\frac{1}{D^L_{j,t}}\sum_{i} y_{j,i,t} \le \bar{y}_{m,t} \le \frac{1}{D^U_{j,t}}\sum_{i} y_{j,i,t}\ \ \forall\ m \in M_{RM}, j \in J_{RM}, t \tag{15}$$

$$\frac{1}{D^L_{i,t}}\sum_{j} y_{j,i,t} \le y_{m,t} \le \frac{1}{D^U_{i,t}}\sum_{j} y_{j,i,t}\ \ \forall\ m \in M_{ST}, i \in I_{ST}, t \tag{16}$$

$$D^L_{u,t} \le \sum_{m\in M_{ST}} y_{m,t} \le D^U_{u,t}\ \ \forall\ u \in U_{ST}, t \tag{17}$$

$$\sum_{m\in M_{ST}} y_{m,t} \ge 1\ \ \ u \in U_{ST} \forall\ t \tag{18}$$

$$x_{j,i,t} + \bar{x}^U_{j,t}\ y_{j,i,t} - \bar{x}^U_{j,t} \le xe_{j,t} \le x_{j,i,t} + \bar{x}^L_{j,t}\ y_{j,i,t} - \bar{x}^L_{j,t}\ \ \forall\ j \in J_{CL}, t \tag{19}$$

$$xe_{j,t} = xe_{j,t+1}\ \forall\ j \in J_{CL}, t < t_{end} \tag{20}$$

$$x_{j,i,t} \in \mathbb{R};\ x^{LOD}_{i,t}, x^{UPD}_{i,t} \in \mathbb{R}^+;\ y_{j,i,t}, y_{m,t} = \{0,1\} \tag{21}$$

Equations (4) and (5) are modified to Eqs. (12) and (13), respectively. For the I_{ST} in-ports in Eq. (12), UB_i of each quality in *i* is used as upper bounds of the stream values arriving in *ST*, related to the mode of operation (CL_1, CL_2, CL_3, CL_4) to be defined at each raw material transferring by the binary $y_{m,t}$ for $m \in M_{ST}$. Equation (13) is only constructed for the *RM* out-ports J_{RM} and its lower and upper bounds are set to the values of the compound-properties $pr_{j,t}$ (crude-oil assay data), and the binary $y_{m,t}$ of the sole unit-operation in *RM* is set to the unitary value every each time-period *t* ($\bar{y}_{m,t}$ = 1) given by the sequence of transferring already defined.

Equation (14) is the structural transition valid cut connecting unit-operations *m'* and *m* of different units to the sets *JI*. The binaries $y_{m',t}$ of the *RM* and *CL* are fixed to the unitary value as they are considered active during the whole time horizon ($\bar{y}_{m',t}$ = 1 for M_{RM} and M_{CL}). The *RM* unit-operation setup of each time-period turns on the quality values in the *RM* out-ports J_{RM} for each raw material transported within each discrete-time step since the lower and upper bounds of the J_{RM} are set to the compound-property values of each crude-oil time-step as in Eq. (13). The *CL* unit-operation setup is considered turned-on at all *t* as J_{CL} out-ports must be turned-on to maintain the components-properties active.

Constraints (15) to (21) complete the clustering program by including procedural relationships in the ports that connects the unit-operations and the arrows considering their inherent continuous and binary variables. In the multi-use procedural constraints (15) and (16), the lower and upper parameters ($D^L_{j,t}$ and $D^U_{j,t}$) coordinate the use of the J_{RM} ports in Eq. (15) and the $D^L_{i,t}$ and $D^U_{i,t}$ control the I_{ST} ports use in Eq. (16). Equation (17) is the multi-use constraint applied to the unit *ST* that controls its use by the *RM* links. Equation (18) is the so-called zero downtime constraint applied in the storage unit in U_{ST} to select at least one mode of operation *m* in M_{ST}, the storage-cluster mode.

Equations (19) and (20) are the flow-equaling constraints for the cluster out-ports J_{CL}. The structural equaling in Eq. (19) determines the equaling variable $xe_{j,t}$ when occurs both the transferring to the storage ($y_{j,i,t}$ = 1 for $JI_{RM\text{-}ST}$) and the clustering gets active ($y_{j,i,t}$ = 1 for $JI_{CL\text{-}ST}$), saving at the time *t*, the value of the J_{CL} ports. The temporal equaling variable in Eq. (20) maintains its value the same in all time periods. These equations are the main reason for the cluster objects *CL*. In Figure 1, when only a single raw material is clustered to the storage unit-operation *m*, the equaling variable $xe_{j,t}$ of each J_{CL} in all time-periods is the value of the $x_{j,i,t}$ in J_{RM}. When more than one raw material is transferred to the same storage-cluster pair, the $xe_{j,t}$ of each J_{CL} is the value of those clustered raw material ($x_{j,i,t}$ in J_{RM}) that reduces the quality variation of the overall transferring.

The final considerations are for Eq. (16) and Eq. (17), the multi-use constraints in the storage in-ports I_{ST} and perimeter-unit U_{ST}. By Figure 1, they can be reduced using only the lower bound inequality from Eq. (16) and the upper bound from Eq. (17), as $D^L_{i,t}$=$D^U_{i,t}$= 2 (the J_{RM} and J_{CL} use of I_{ST} at time *t*) and $D^L_{u,t}$=$D^U_{u,t}$=1 (only one raw material *RM* goes to the storage at time *t*).

Illustrative Examples

We model and solve the illustrative and industrial-sized examples using the structural-based unit-operation-port-state superstructure (UOPSS) found in the semantic-oriented platform IMPL (Industrial Modeling and Programming Language) with IBM's CPLEX 12.6 MILP solver on Intel Core i7 machine at 2.7 Hz with 16GB RAM.

The illustrative examples use 5 and 10 crude-oils from the beginning of Table 1. Their results show in Figures 2 and 3 the values of the $y_{m,t}$ in the storage *ST* considering the modes *CL1* to *CL4* for each time-period, which represents the crude-oils transferred to the storages. As the problem has 4 storage-clusters, when there are 4 different crude-oils, the objective function solution is trivially zero. In this case, there is no quality variance as each crude-oil is transferred to different storage-clusters. For a number of crude-oil higher than the number of storage-clusters, there is a need for sharing of storage so that a variation of the target variables $x^{LO}_{i,t}$ and $x^{UP}_{i,t}$ in the in-ports of the shared storage is computed in the objective function.

For the 5 first crude-oils from Table 1, Figure 2 shows the segregation, sharing of storage or clustering of the crude-oils 4 and 5 (time-periods 4 and 5 in the same storage-cluster *ST-CL2*). When the raw material transfer occurs during the storage assignment, only one set of 3 binary variables $y_{j,i,t}$ in $JI_{RM\text{-}ST}$, $y_{j,i,t}$ in $JI_{CL\text{-}ST}$ and $y_{m,t}$ in M_{ST} are true or turned-on. For the crude-oils 1, 2 and 3, during the transfers of each one within their own time-period *t*, the continuous variable values in the arrows ($x_{j,i,t}$ in $JI_{RM\text{-}ST}$ and $x_{j,i,t}$ in $JI_{CL\text{-}ST}$) have the same absolute value, but with opposite sign as controlled by Eqs. (10) and (11). So the quality deviation $x^{LO}_{i,t}$ and $x^{UP}_{i,t}$ in the storage in-ports I_{ST} of the crude-oils 1, 2 and 3 is zero since these crude-oils are not sharing storage.

To facilitate the analyses of the clustering procedures, we track the equaling variable $xe_{j,t}$ in the clusters out-ports J_{CL}. For the crude-oils 4 and 5 in cluster CL_2, the variable $xe_{j,t}$ can have values of the J_{RM} out-ports of one of them, so there is no match of the target $\bar{x}_{i,t}$ = 0 and Eq. (8) cannot be equal zero. Then $x^{LO}_{i,t}$and $x^{UP}_{i,t}$ can be different than zero. In the minimization problem, the compound-property 4 (*RY*) in the J_{CL2} out-port is the same of the crude-oil 4, and the other J_{CL2} out-port properties (*i*=1: *NY*; *i*=2: *DY*; *i*=3: *DS*) are those as in crude-oil 5. The solution is *Z* = 0.99 in 0.89s and the non-zero variable targets are $x^{LO}_{i=3,t=4}$ = 0.1632 and $x^{UP}_{i=1,t=4}$ = 3.1632; $x^{UP}_{i=2,t=4}$ = 0.7200 and $x^{LO}_{i=4,t=5}$ = 6.050. For these target deviation variables, we have, respectively, 0.14; 0.11; 0.36 and 0.38, when they are divided by their weights, giving their summation the value of *Z* = 0.99.

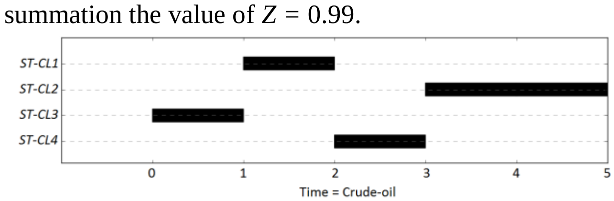

Figure 2. Gantt chart for the 5 first crude-oils.

For the 10 first crude-oils from Table 1, we can see in Figure 3 the clustering of 1, 2, 4, 6 and 8 to the storage-cluster *ST-CL1*. Crude-oils 5 and 10 in *ST-CL2*, 3 and 9 are in *ST-CL4* and the crude-oil 7 is isolated in *ST-CL3*. We can notice by inspection that these segregations are quite easy to understand as the crude-oils clustered in the *ST-CL* groups have similar compound-properties. Also, crude-oil 7 is set aside from others as it is an ultra-heavy oil as found by the lowest *NY* plus *DY* values and the highest *RY*.

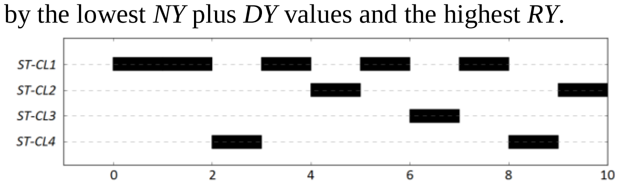

Figure 3. Gantt chart for the 10 first crude-oils.

Industrial-Scale Cases

Stated below are the cluster memberships (cardinality) and means for each compound-properties considering the 45 different crude-oils from Table 1. The highest quality deviation in the objective function, for all 45 crude-oils together, is 13.46. Figure 4 shows the computational statistics of the problems. In all cases the global optimum is found within 4 minutes and the time to close the MILP relaxation gap is pointed out for the cases with more than 25 crude-oils.

Cluster 1 (16 crude-oils): 1, 4, 6, 8, 12, 13, 15, 21, 28, 30, 34, 36, 40, 41, 43, 45
Cluster 2 (9 crude-oils): 3, 9, 14, 24, 25, 27, 32, 33, 37
Cluster 3 (18 crude-oils): 2, 5, 10, 11, 16, 17, 18, 19, 20, 22, 26, 29, 31, 35, 38, 39, 42, 44
Cluster 4 (2 crude-oils): 7, 23

Cluster 1 Means: 18.75, 20.84, 0.1594, 8.58
Cluster 2 Means: 32.88, 13.61, 0.0900, 1.53
Cluster 3 Means: 14.08, 17.06, 0.2300, 18.64
Cluster 4 Means: 7.72, 10.31, 1.83, 35.39

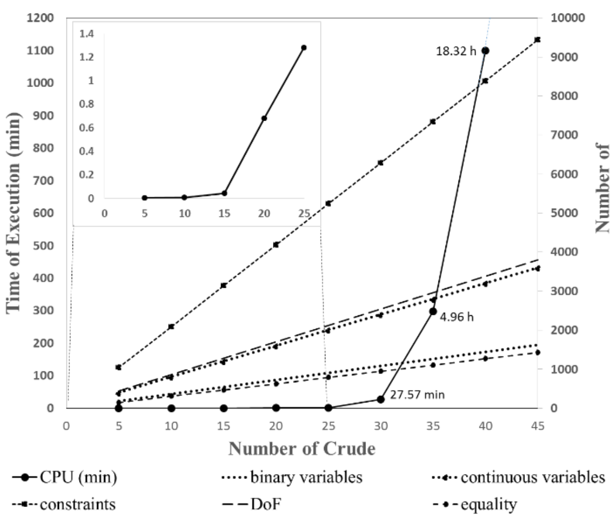

Figure 4. Computational statistics of the problems.

Comparing the proposed formulation with the k-means clustering (KM) and fuzzy c-means clustering (FCM) algorithms (Bezdek et. al., 1984), the objective function value is 14.25 using their formulation. The reason for the discrepancy is due to the fact that the KM and FCM are heuristic searches whereas as the MILP is an exact search. Of interest, the cluster mean values selected by the MILP search correspond directly or exactly to one of the crude-oil compound-properties. However, for the KM and FCM, the means do not correspond directly with the crude-oil compound-properties and hence the explanation for the observed difference. It should be noted that the KM and FCM algorithms are also non-convex, combinatorial and non-polynomial in time which require a higher-level randomized search heuristic to find the lowest possible objective function by reseeding the algorithms.

Conclusion

We have highlighted the application to create groupings, partitions, segregations or clusters of how to assign/allocate individual crude-oils or other feedstocks to a limited number of storage as well as providing a methodology of how to specify the cluster variables i.e., compound-properties. This is necessary to significantly improve the crude-oil blend property control as well as the blend scheduling optimization. This clustering is also similar to the notion of grouping into families, etc. found in various sequence-dependent changeover heuristics on shared resources such as using product-wheels and blocking. Finally, we believe that this is first time such a structured approach, using MILP, has been applied to the design of relatively simple and straightforward segregation rules with the ultimate goal of achieving better crude-oil management inside the oil-refinery which is not addressed by the ubiquitous crude-oil feedstock selection monthly planning model.

References

Ashe, T.R., Kelly, J.D., Lau, T.C., Pho, R.T., (1997). A Method for Predicting Chemical or Physical Properties of Complex Mixtures. US Patent #5,602,755.

Bezdek, J. C., Ehrlich, R., Full, W. (1984). FCM: the Fuzzy C-Means Clustering Algorithm. *Comput. Geosci.*, 10, 191.

ExxonMobil. (2015). http://www.exxonmobil.com/crudeoil, accessed April, 2016.

Kelly, J. D. (2005). The Unit-Operation-Stock Superstructure (UOSS) and the Quantity-Logic-Quality Paradigm (QLQP) for Production Scheduling in The Process Industries. *In Multidisciplinary International Scheduling Conference Proceedings*: New York, United States, 327.

Kelly, J. D., Forbes, J. F. (1998). Structured Approach to Storage Allocation for Improved Controllability, *AICHE*, 44,

Kelly, J. D., Mann, J. M. (2003a). Crude-Oil Blend Scheduling Optimization: An Application with Multi-Million Dollar Benefits. *Hydro. Proces*., June, 47.

Kelly, J. D., Mann, J. M. (2003b). Crude-Oil Blend Scheduling Optimization: An Application with Multi-Million Dollar Benefits. *Hydro. Proces*., July, 72.

Kelly, J. D., Menezes, B. C., Grossmann, I. E., Engineer, F. (2017). Crude-oil Blend Scheduling Optimization of an Industrial-Sized Refinery: a Discrete-Time Benchmark. *In Foundations of Computer Aided Process Operations*: Tucson, United States.

Zyngier, D., Kelly, J. D. (2012). UOPSS: A New Paradigm for Modeling Production Planning and Scheduling Systems. *In European Symposium in Computer Aided Process Engineering*: London, United Kingdom.