Chemical Engineering Science 66 (2011) 4307–4318

Contents lists available at ScienceDirect

Chemical Engineering Science

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

A multiscale decomposition method for the optimal planning and scheduling of multi-site continuous multiproduct plants

Sebastian Terrazas-Moreno, Ignacio E. Grossmann *

Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15232, USA

ARTICLE INFO

Article history:
Received 24 December 2010
Received in revised form
25 February 2011
Accepted 10 March 2011
Available online 21 March 2011

Keywords:
Scheduling
Planning
Multiproduct plant
Multiscale decomposition
Lagrangean decomposition
Bi-level decomposition

ABSTRACT

This paper addresses the solution of simultaneous scheduling and planning problems in a production–distribution network of continuous multiproduct plants that involves different temporal and spatial scales. Production planning results in medium and long-term decisions, whereas production scheduling determines the timing and sequence of operations in the short-term. The production–distribution network is made up of several production sites distributing to different markets. The planning and scheduling model has to include spatial scales that go from a single production unit within a site, to a geographically distributed network. We propose to use two decomposition methods to solve this type of problems. One method corresponds to the extension of the bi-level decomposition of Erdirik-Dogan and Grossmann (2008) to multi-site, multi-market networks. A second method is a novel hybrid decomposition method that combines bi-level and spatial Lagrangean decomposition methods. We present four case studies to study the performance of the full space planning and scheduling model, the bi-level decomposition, and the bi-level Lagrangean method in profit maximization problems. Numerical results indicate that in large-scale problems, decomposition methods outperform the full space solution and that as problem size increases the hybrid decomposition method becomes faster than the bi-level decomposition alone.

1. Introduction

The process systems engineering (PSE) community has been concerned with the integration of planning and scheduling in the process industry for the last two decades. One of the first examples is the work by Birewar and Grossmann (1990), where aggregate sequencing constraints were included as part of the production planning problem of a multiproduct flow shop batch plant. A later review by Grossmann et al. (2002) established the need for discrete optimization models for solving the integrated planning and scheduling problem. Since then, a considerable number of papers have appeared on the subject of integration of planning and scheduling. The works by Hooker (2005), Stefansson et al. (2006), Sung and Maravelias (2007), Erdirik-Dogan and Grossmann (2006, 2008), Verderame and Floudas (2008), Liu et al. (2008), and Li and Ierapetritou (2009, 2010), are but a few recent examples. The review paper by Maravelias and Sung (2009) offers a more extensive description of recent developments.

The main argument for integration of planning and scheduling is that it is required for ensuring the feasibility of the scheduling decisions at the planning level and for improving the enterprise-wide operational and economic objectives (Grossmann, 2005). The single most important challenge is the computational intractability of optimization models dealing with simultaneous planning and scheduling. This is a result of integrating a medium or long term planning problem dealing with months and years with short term scheduling dealing with days or hours. In the case of a geographically distributed network (as in the case of a multi-site process network distributing to several markets), the problem involves not only different temporal scales, but also a range of spatial scales that go from single units in one plant to a multi-site distribution network.

Maravelias and Sung (2009) classify works in planning and scheduling in terms of the characteristics of the planning model and in terms of the solution method. The modeling approach can involve a detailed scheduling formulation extended to cover the planning horizon, aggregate or relaxed scheduling constraints, off-line-surrogate formulations, or hybrid methods for rolling horizon. In terms of solution approach they can be hierarchical, iterative, or full space. This paper proposes an algorithm for solving a multi-site planning and scheduling problem for continuous multiproduct plants in which we use an existing aggregate model for the planning problem (Erdirik-Dogan and Grossmann, 2008) of a single site, but extends it to multi-site setting (Jackson and Grossmann, 2003). The solution strategy uses an iterative bi-level decomposition for dealing with the different temporal scales that result from the planning and scheduling integration, combined with Lagrangean decomposition (Guignard and Kim, 1987) to address the integration of spatial scales. The scheduling model embedded within the integrated planning and

* Corresponding author.
E-mail address: grossmann@cmu.edu (I.E. Grossmann).

0009-2509/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ces.2011.03.017

scheduling problem corresponds to a continuous time, slot-based model for a single-stage multiproduct continuous plant with parallel lines and sequence dependent changeovers (Erdirik-Dogan and Grossmann, 2008). Recent papers by Liu et al. (2010), Lima et al. (2011), and Kopanos et al. (2011) are relevant to this work. Liu et al. (2010) discuss planning formulations with TSP-type sequencing models, where they propose subtour elimination constraints different from the ones used by Erdirik-Dogan and Grossmann (2008). Lima et al. (2011) address the planning and scheduling of a single-unit continuous multiproduct plant that manufactures glass. They solve large-scale industrial case studies through bi-level decomposition and implementing several variations of rolling horizon algorithms. Kopanos et al. (2011) also address an industrial-scale problem, in this case a real-world bottling facility. Both, Lima et al. (2011) and Kopanos et al. (2011) introduce constraints that allow product transitions to start in one planning period and end in the next one. This type of constraints is included in this work, since we assume that sequence dependent changeovers can have significant duration.

Li and Ierapetrirou (2009, 2010) developed solution methods based on Augmented Lagrangean Relaxation (ALR) for planning and scheduling of batch processes where the objective is cost minimization. They exploit the bi-level structure of the problem and decompose it into scheduling subproblems for every time period and production site. Our work also uses a bi-level representation of the problem with an upper planning level and a lower scheduling level. However, we use a planning model with aggregated sequencing constraints in the upper level as opposed to a surrogate model with implicit objective function as they do. The solution method is also different since we decompose the upper-level planning problem into single-site and single-market problems to derive a relaxation bound for the integrated problem, while they decompose the lower level scheduling problem and enforce a penalty term through ALR in order to obtain a feasible solution of the original formulation.

In the next sections we define the integrated multi-site multi-market planning and scheduling problem for continuous multi-product plants followed by the corresponding mathematical model. We then present a new hybrid algorithm for obtaining optimal or near optimal solutions of the simultaneous planning and scheduling problem. In a later section, several case studies of different sizes are solved using the proposed method. The paper concludes with some remarks and discussion of the numerical results.

2. Problem statement

We model the market-facing end of the chemical supply chain as a process network where the objective is profit maximization from product revenues. As seen in Fig. 1, nodes in this network are either production sites or markets. Arcs in the network correspond to potential transportation links for shipping the products in the network. Each production node corresponds to a single stage continuous manufacturing facility that can produce several products, giving rise to a scheduling problem within each production node. The flow of products in the network is limited by the capacity of the production nodes, the shipping capacity of the arcs, and the demand in the market nodes. Production changeovers can have a significant duration that is sequence-dependent. Therefore, scheduling decisions affect the extent to which it is possible to exploit the capacity of each production node.

The description of the model of the chemical supply chain involves two different time scales. The long term time scale involves decisions such as the amount of product manufactured in the production nodes and the flows of those products to the markets. The short term time scale involves scheduling and sequencing decisions in each production site. The objective is to maximize

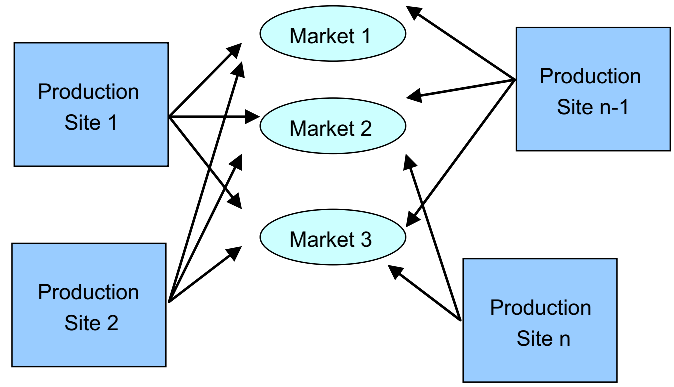

Fig. 1. Production and distribution network.

profit over a medium to long term planning horizon (weeks to months), while operating the network in the short-term (daily or hourly) so as to meet the product demands. In addition to the medium and long-term decisions for the planning problem and the short-term decisions for the scheduling at each production site, spatial scales due to the geographical distribution of the sites have to be integrated to achieve the optimum of the overall system. The integrated optimization problem is then as follows:

Given are:

- A set of products.
- A set of production sites.
- A set of single stage multiproduct continuous plants with parallel production lines.
- A set of markets for the products.
- A set of time periods for production planning.
- The forecasted demand and price for each product at the end of the time periods.
- The production cost and inventory holding cost for each product.
- The changeover cost and time for each pair of products in each site.
- The shipment cost for each product between each site and each market.
- Production, storage, and shipment maximum capacities.

The problem is to *determine*:

- Planning decisions for each time period
  - Amounts to be produced of each product at each site.
  - Amounts to be stored in inventory at each site.
  - Shipments between each production site and each market.
- The production schedule at each site
  - Assignment of products and sequence in each production line at each time period.
  - Start and end time for the production runs at each time period.
  - Amount of product manufactured at each production run.

with the *objective* of maximizing the profit, calculated as

Sales − Production cost − Manufacturing cost − Changeover cost − Shipment cost.

3. Mathematical model

The models in the following section are an extension of the MILP model by Erdirik-Dogan and Grossmann (2008) for continuous multiproduct plants with parallel lines modified to allow for transitions beginning in one time period and ending in the next (Lima et al., 2011; Kopanos et al., 2011).

Nomenclature

Index/set

$i,k/I$	indices/set of products
ℓ/L	index/set of time slots
s/S	index/set of manufacturing sites
t/T	index/set of time periods
m/M	index/set of markets
n/N	index/set of parallel lines
I_n	set of products that can be manufactured in line n
L_n	set of slots that are defined for line n

Parameters

$\beta_t^{i,m}$	sale price of product i in market m during time period t
$\omega_s^{i,n}$	operating cost of product i in line n of production site s
δ_s^i	inventory cost of product i in production site s
$\gamma_s^{i,m}$	shipment cost between production site s and market m
$TC_s^{i,k}$	cost incurred during a transition from product i to product k in production site s
$a_s^{i,n}$	production rate of product i in line n of production site s
$bt_s^{i,k,n}$	transition time from product i to product k in line n of production site s
$v_{s,i}^{UP}$	maximum storage capacity for product i in production site s
$f_{s,i}^{UP}$	maximum shipping amount between production site s and market m
$sh_{s,m,i}^{UP}$	maximum capacity for shipping product i between production site s and market m
$d_t^{i,m}$	demand for product i in market m during time period t
H_t	length of time period t
h_t	hours at the end of time period t
$mds_t^{i,m}$	minimum percentage of demand of product i that has to be satisfied in market m during time period t

Variables

$x_{t,s}^{i,n,\ell}$	production of product i in site s, slot ℓ, of line n in time period t
$xt_{t,s}^{i,n}$	production variable upper level problem
$p_{t,s}^i$	aggregated production or product i of all lines in site s during period t
$v_{t,s}^i$	inventory of product i in site s at the end of time period t
$vt_{t,s}^i$	inventory variable for upper level problem
$\theta_{t,s}^{i,n,\ell}$	production time i in slot ℓ during period t in site s and line n for lower level problem
$\theta t_{t,s}^{i,n,\ell}$	production time variable for upper level problem
$ts_{t,s}^{n,\ell}$	start time of slot ℓ
$\tau e_{t,s}^{i,k,n}$	part of the transition time between time periods assigned to the end of period t
$\tau s_{t,s}^{i,k,n}$	part of the transition time between time periods assigned to the start of period t
$\tau ep_{t,s}^{i,k,m}$	part of the transition time between time periods assigned to the end of period t in the upper level planning problem
$\tau sp_{t,s}^{i,k,m}$	part of the transition time between time periods assigned to the start of period t in the upper level planning problem
$te_{t,s}^{n,\ell}$	end time of slot ℓ
$sh_{t,s}^{i,m}$	shipment of product i from manufacturing site s to market m at time period t
$sht_{t,s}^{i,m}$	shipment variable for upper level problem
$z_{t,s}^{i,k,n,\ell}$	continuous variable bounded a 0 and 1 to denote a transition between products i and k
$trt_{t,s}^{i,k,n}$	binary variable denoting changeovers between products i and k across adjacent periods
$ny_{t,s}^{i,n}$	number of slots of product i in time t
$w_{t,s}^{i,n,\ell}$	binary variable to denote if product i is produced in slot ℓ of line n at site s in period t
$yp_{t,s}^{i,n}$	binary variable in upper level problem to denote if product i is produced in line n at site s in period t
$yop_{t,s}^{i,n}$	binary variable for limiting the search space of $w_{t,s}^{i,n,\ell}$ in the lower level problem
$xf_{t,s}^{i,n}$	binary variable to denote if product i is the first product manufactured in site s during period t
$xl_{t,s}^{i,n}$	binary variable to denote if product i is the last product manufactured in site s during period t
$zp_{t,s}^{i,k,n}$	binary variable to denote if product i precedes product k in unit m during time period t
$zzp_{t,s}^{i,k,n}$	binary variable that indicates if the link between products i and k is broken
$zzz_{t,s}^{i,k,n}$	binary variable denoting changeovers between products i and k across adjacent periods

3.1. Integrated planning and scheduling model

Eq. (1) represents the objective function of the integrated problem. The first term corresponds to the product sales minus the shipment costs; we assume that all products shipped to a market are sold. The second term are the production and inventory costs, and the third term represents the cost of transitions

$$\max \pi = \sum_{t \in T}\left[\sum_{i \in I}\sum_{m \in M}\sum_{s \in S}(\beta_s^{i,m}-\gamma_s^{i,m})sh_{t,s}^{i,m} - \sum_{i \in I}\sum_{s \in S}\left(\sum_{n \in N}\sum_{\ell \in L}\omega_s^{i,n}x_{t,s}^{i,n,\ell}+\delta_{t,s}^i v_{t,s}^i\right)\right] - \sum_{i \in I}\sum_{k \in I}\sum_{n \in N}\sum_{t \in T}\sum_{s \in S}TC_s^{i,k}\left(\sum_{\ell \in L}z_{t,s}^{i,k,\ell,n}+trt_{t,s}^{i,k,n}\right) \tag{1}$$

In Eq. (2) the sum of the production at each slot is aggregated into the variable p that stands for the total production of product i in site s during period t

$$\sum_{n \in N}\sum_{\ell \in L}x_{t,s}^{i,n,\ell} = p_{t,s}^i, \quad i \in I,\ s \in S,\ t \in T \tag{2}$$

The set of equations in (3) represents the mass balances for every product at the end of each time period in every site

$$v_{t-1,s}^i + p_{t,s}^i = \sum_{m \in M} sh_{t,s}^{i,m} + v_{t,s}^i, \quad i \in I,\ s \in S,\ t \in T \tag{3}$$

Eq. (4) constraints the shipments (sales) to be less than or equal to the demand at each market, and greater than or equal to a minimum demand satisfaction level

$$mds_t^{i,m} \le \sum_{s \in S} sh_{t,s}^{i,m} \le d_t^{i,m}, \quad i \in I,\ m \in M,\ t \in T \tag{4}$$

The summation in Eq. (5) limits the production in each slot to one product

$$\sum_{i \in I} w_{t,s}^{i,n,\ell} = 1, \quad n \in N,\ \ell \in L_n,\ t \in T,\ s \in S \tag{5}$$

According to constraint (6) the production time of i in slot ℓ can only be nonzero if the product has been assigned to it

$$\theta_{t,s}^{i,n,\ell} \le H_t w_{t,s}^{i,n,\ell}, \quad n \in N,\ i \in I_n,\ \ell \in L_n,\ t \in T,\ s \in S \tag{6}$$

Eq. (7) sets the production of i equal to the production rate times the production time

$$x_{t,s}^{i,n,\ell} = a_s^{i,n}\theta_{t,s}^{i,n,\ell}, \quad n \in N,\ i \in I_n,\ \ell \in L_n,\ t \in T,\ s \in S \tag{7}$$

In Eq. (8) the end time of every time slot except the last in each time period is computed as the starting time plus the production and the transition times. Eq. (9) determines the end of the last slot in each time period as its start time plus the production time

$$te_{t,s}^{n,\ell} = ts_{t,s}^{n,\ell} + \sum_{i \in I_n} \theta_{t,s}^{i,n,\ell} + \sum_{i \in I_n} \sum_{k \in I_n} bt_s^{i,k,n} z_{t,s}^{i,k,n,\ell}, \quad n \in N, \ \ell \in L_n, \ \ell < |L_n| \, t \in T, \ s \in S \tag{8}$$

$$te_{t,s}^{n,\ell} = ts_{t,s}^{n,\ell} + \sum_{i \in I_n} \theta_{t,s}^{i,n,\ell}, \quad n \in N, \ \ell \in L_n, \ \ell = |L_n| \, t \in T, \ s \in S \tag{9}$$

Constraint (10) determines the starting time of the first slot of each time period to be greater than or equal to the time when the previous period ended plus the part of the transition across time periods assigned to the current time period

$$ts_{t+1,s}^{n,1} \geq h_t + \sum_{i \in I_n} \sum_{k \in I_n} \tau s_{t+1,s}^{i,k,n}, \quad n \in N, \ \ell \in L_n, \ t \in T, \ t < |T|, \ s \in S \tag{10}$$

From Eq. (11), the ending time of a slot is equal to the starting time of the one immediately following it, except for the last slot of each time period

$$ts_{t,s}^{n,\ell+1} = te_{t,s}^{n,\ell}, \quad n \in N, \ i \in I_n, \ \ell \in L_n, \ \ell < |L_n|, \ t \in T, \ s \in S \tag{11}$$

Constraint (12) ensures that the ending time of the last slot plus the part of the transition time across time periods assigned at the end of the period is less than or equal to the ending point of the time period

$$te_{t,s}^{n,|L_n|} + \sum_{i \in I_n} \sum_{k \in I_n} \tau e_{t,s}^{i,k,n} \leq h_t, \quad n \in N, \ \ell \in L_n, \ t \in T, \ s \in S \tag{12}$$

Eqs. (13) and (14) assign the value of one to the relevant transition variable between products i and k produced in adjacent slots in the same time period

$$\sum_{k \in I_n} z_{t,s}^{i,k,n,\ell} = w_{t,s}^{i,n,\ell}, \quad n \in N, \ i \in I_n, \ \ell \in L_n, \ t \in T, \ s \in S \tag{13}$$

$$\sum_{i \in I_n} z_{t,s}^{i,k,n,\ell} = w_{t,s}^{k,n,\ell+1}, \quad n \in N, \ k \in I_n, \ \ell \in L_n, \ t \in T, \ s \in S \tag{14}$$

Eqs. (15) and (16) assign the value of one to the relevant transition variable between products i and k produced in adjacent slots in different time periods. Eq. (17) splits the transition time across time periods into a part belonging to t and another to $t+1$

$$\sum_{k \in I_n} trt_{t,s}^{i,k,n,\ell} = w_{t,s}^{i,n,|L_n|}, \quad n \in N, \ i \in I_n, \ t \in T, \ s \in S \tag{15}$$

$$\sum_{i \in I_n} trt_{t,s}^{i,k,n,\ell} = w_{t+1,s}^{k,n,1}, \quad n \in N, \ k \in I_n, \ t \in T, \ s \in S \tag{16}$$

$$bt_s^{i,k,n} trt_{t,s}^{i,k,n} = \tau e_{t,s}^{i,k,n} + \tau s_{t+1,s}^{i,k,n}, \quad n \in N, \ i \in I_n, \ k \in I_n, \ t \in T, \ t < |T|, \ s \in S \tag{17}$$

The same product can be assigned to several adjacent slots if the total number of slots is greater than the number of products assigned for production during a time period. The extra slots in this situation are labeled "flexible slots". Since the duration of time slots is variable, this can result in degenerate assignment of products. The set of constrains (18)–(22) prevents such degeneracy by enforcing that all flexible slots be assigned to the first product in the production sequence. Erdirik-Dogan and Grossmann (2006) provide a detailed explanation of these constraints

$$yop_{t,s}^{i,n} \geq w_{t,s}^{i,n,\ell}, \quad n \in N, \ i \in I_n, \ \ell \in L_n, \ t \in T, \ s \in S \tag{18}$$

$$yop_{t,s}^{i,n} \leq ny_{t,s}^{i,n}, \quad n \in N, \ i \in I_n, \ t \in T, \ s \in S \tag{19}$$

$$ny_{t,s}^{i,n} \leq |L_n| yop_{t,s}^{i,n}, \quad n \in N, \ i \in I_n, \ t \in T, \ s \in S \tag{20}$$

$$ny_{t,s}^{i,n} \geq |L_n| - \sum_{k \in I_n} yop_{t,s}^{i,n} - 1 - bigM(1 - w_{t,s}^{i,n,\ell}) \quad n \in N, \ i \in I_n, \ \ell \in L_n, \ t \in T, \ s \in S \tag{21}$$

$$ny_{t,s}^{i,n} \leq |L_n| - \sum_{k \in I_n} yop_{t,s}^{i,n} - 1 + bigM(1 - w_{t,s}^{i,n,\ell}) \quad n \in N, \ i \in I_n, \ \ell \in L_n, \ t \in T, \ s \in S \tag{22}$$

Constraints (23) and (24) enforce maximum capacity for storage and transportation

$$v_{t,s}^{i} \leq v_{s,i}^{UP}, \quad i \in I, \ s \in S, \ t \in T \tag{23}$$

$$sh_{t,s}^{i,m} \leq sh_{s,m,i}^{UP}, \quad i \in I, \ m \in M, \ s \in S, \ t \in T \tag{24}$$

Model (PS) given by Eqs. (1)–(24) defines the optimal multi-site multi-market integrated planning and problem.

4. Solution method

The formulation (PS) involves detailed scheduling of several manufacturing sites of continuous multiproduct plants over a planning horizon that can cover weeks or months. Short-term scheduling for a single site is a challenging problem that when integrated with medium or long term planning of a process network can become computationally intractable. In fact most production scheduling problems such as the one considered in this paper are NP-hard (Garey and Johnson, 1979). Even though recent developments in computational capabilities allow the solution of larger optimization problems, decomposition algorithms provide a practical approach for solving industrial scheduling problems. Therefore, our goal in this work is to develop an optimization approach to solve realistic problem instances of (PS) in a reasonable amount of time, meaning in the order of a few minutes to an hour. In this section, we describe an algorithm based on bi-level and Lagrangean decompositions to solve instances of (PS) considerably faster than in full space. We first describe the bi-level decomposition of (PS), followed by the hybrid bi-level Lagrangean decomposition.

4.1. Bi-level decomposition

We extend the bi-level decomposition strategy for planning and scheduling as proposed by Erdirik-Dogan and Grossmann (2008) to the multi-site multi-market case. The main idea is to only address the temporal part of the problem decomposing the planning and scheduling parts of the problem in such a way that the planning problem over all sites and markets contains some of the information required to obtain a feasible solution at the scheduling level and that an iterative solution of both problems converge to the optimal solution of (PS). This is achieved by: (i) using an aggregate sequencing model at the planning level to account for sequence-dependent transitions using traveling salesman problem (TSP) constraints; (ii) limiting the search space of the binary variables at the scheduling level using the solution of the planning problem; (iii) adding integer cuts to the planning problem to provide a tighter representation of the scheduling problem, which we should note is optimized simultaneously over all production sites and markets. The planning problem with aggregate sequencing considerations is called the upper-level planning problem (UP) and is given by the following model.

4.2. Upper-level planning problem

$$\max \pi p = \sum_{t \in T} \sum_{s \in S} \sum_{m \in M} \sum_{i \in I} \beta_{t,s}^{i,m} sht_{t,s}^{i,m} - \sum_{t \in T} \sum_{s \in S} \sum_{i \in I} \left(\sum_{n \in N} \omega_s^{i,n} xt_{t,s}^{i,n} + \delta_s^i vt_{t,s}^i \right) - \sum_{t \in T} \sum_{s \in S} \sum_{m \in M} \sum_{i \in I} \gamma_s^{i,m} sht_{t,s}^{i,m}$$

$$-\sum_{t\in T}\sum_{s\in S}\sum_{i\in I}\sum_{k\in I}\sum_{n\in N}TC_s^{i,k,n}(zp_{t,s}^{i,k,n}+zzz_{t,s}^{i,k,n}-zzp_{t,s}^{i,k,n}) \quad (25)$$

s.t.

$$\sum_{n\in N}xt_{t,s}^{i,n}=pt_{t,s}^{i},\quad i\in I,\ s\in S,\ t\in T \quad (26)$$

$$v_{t-1,s}^{i}+pt_{t,s}^{i}=\sum_{m\in M}sht_{t,s}^{i,m}+v_{t,s}^{i},\quad i\in I,\ s\in S,\ t\in T \quad (27)$$

$$mds_t^{i,m}\le\sum_{s\in S}sht_{t,s}^{i,m}\le d_t^{i,m},\quad i\in I,\ m\in M,\ t\in T \quad (28)$$

$$\theta_{t,s}^{i,n}\le H_t yp_{t,s}^{i,n},\quad n\in N,\ i\in I_n,\ t\in T,\ s\in S \quad (29)$$

$$xt_{t,s}^{i,n}=a_s^{i,n}\theta t_{t,s}^{i,n},\quad n\in N,\ i\in I_n,\ t\in T,\ s\in S \quad (30)$$

$$yp_{t,s}^{i,n}=\sum_{k\in I_n}zp_{t,s}^{i,k,n},\quad n\in N,\ i\in I_n,\ t\in T,\ s\in S \quad (31)$$

$$yp_{t,s}^{k,n}=\sum_{i\in I_n}zp_{t,s}^{i,k,n},\quad n\in N,\ k\in I_n,\ t\in T,\ s\in S \quad (32)$$

$$\sum_{i\in I_n}\sum_{k\in I_n}zzp_{t,s}^{i,k,n}=1,\quad n\in N,\ s\in S,\ t\in T \quad (33)$$

$$zzp_{t,s}^{i,k,n}\le zp_{t,s}^{i,k,n},\quad n\in N,\ i\in I_n,\ k\in I_n,\ t\in T,\ s\in S \quad (34)$$

$$yp_{t,s}^{i,n}\ge zp_{t,s}^{i,i,n},\quad n\in N,\ i\in I_n,\ t\in T,\ s\in S \quad (35)$$

$$zp_{t,s}^{i,i,n}+yp_{t,s}^{k,n}\le 1,\quad n\in N,\ i\in I_n,\ k\in I_n,\ i\neq k,\ t\in T,\ s\in S \quad (36)$$

$$zp_{t,s}^{i,i,n}\ge yp_{t,s}^{i,n}-\sum_{k\neq i,k\in I}yp_{t,s}^{k,n},\quad n\in N,\ i\in I_n,\ t\in T,\ s\in S \quad (37)$$

$$xf_{t,s}^{k,n}\ge\sum_{i\in I_n}zzp_{t,s}^{i,k,n},\quad n\in N,\ k\in I_n,\ t\in T,\ s\in S \quad (38)$$

$$xl_{t,s}^{i,n}\ge\sum_{k\in I_n}zzp_{t,s}^{i,k,n},\quad n\in N,\ i\in I_n,\ t\in T,\ s\in S \quad (39)$$

$$\sum_{i\in I_n}xf_{t,s}^{i,n}=1,\quad n\in N,\ i\in I_n,\ t\in T,\ s\in S \quad (40)$$

$$\sum_{i\in I_n}xl_{t,s}^{i,n}=1,\quad n\in N,\ i\in I_n,\ t\in T,\ s\in S \quad (41)$$

$$\sum_{k\in I_n}zzz_{t,s}^{i,k,n}=xl_{t,s}^{i,n},\quad n\in N,\ i\in I_n,\ t\in T,\ s\in S \quad (42)$$

$$\sum_{i\in I_n}zzz_{t,s}^{i,k,n}=xf_{t+1,s}^{k,n},\quad n\in N,\ K\in I_n,\ t\in T,\ s\in S \quad (43)$$

$$trnp_{t,s}^{n}=\sum_{i\in I_n}\sum_{k\in I_n}bt_s^{i,k,n}(zp_{t,s}^{i,k,n}-zzp_{t,s}^{i,k,n}),\quad n\in N,\ t\in T,\ s\in S \quad (44)$$

$$bt_s^{i,k,n}zzz_{t,s}^{i,k,n}=\tau ep_{t,s}^{i,k,n}+\tau sp_{t+1,s}^{i,k,n},\quad n\in N,\ i\in I_n,\ k\in I_n,\ t\in T,\ t<|T|,\ s\in S \quad (45)$$

$$\sum_{i\in I_n}\theta t_{t,s}^{i,n}+trnp_{t,s}^{n}+\sum_{i\in I_n}\sum_{k\in I_n}\tau ep_{t,s}^{i,k,n}\le H_t,\quad n\in N,\ t=1,\ s\in S \quad (46)$$

$$\sum_{i\in I_n}\sum_{k\in I_n}\tau sp_{t,s}^{i,k,n}+\sum_{i\in I_n}\theta t_{t,s}^{i,n}+trnp_{t,s}^{n}+\sum_{i\in I_n}\sum_{k\in I_n}\tau ep_{t,s}^{i,k,n}\le H_t,\quad n\in N,\ t\in T,\ 1<t<|T|,\ s\in S \quad (47)$$

$$\sum_{i\in I_n}\sum_{k\in I_n}\tau sp_{t,s}^{i,k,n}+\sum_{i\in I_n}\theta t_{t,s}^{i,n}+trnp_{t,s}^{n}\le H_t,\quad n\in N,\ t=|T|,\ s\in S \quad (48)$$

$$v_{t,s}^{i}\le v_{s,i}^{UB},\quad i\in I,\ s\in S,\ t\in T \quad (49)$$

$$sht_{t,s}^{i,m}\le sht_{s,m,i}^{UB},\quad i\in I,\ s\in S,\ m\in M,\ t\in T \quad (50)$$

$$\sum_{(i,n,t,s)\in Z_1^r}yp_{t,s}^{i,n}\le|Z_1^r|-1,\quad r\in R_{cut} \quad (51)$$

Eq. (51) corresponds to an integer cut, and it is used to exclude combinations of assignment variables that are proven to result in a worse solution to the problem (UP) than that found at iteration r. These integer cuts are explained in more detail by Erdirik-Dogan and Grossmann (2008). Here $Z_1^r=\{i,n,t,s : yp_{t,s}^{i,n,r}=1\}$, and R_{cut}={r: r is a bi-level iteration used to construct an integer cut}.

4.3. Lower-level scheduling problem

The lower level subproblem (PL) corresponds to Eqs. (1)–(24) plus constraint (52) given below, which restricts the products i that are to be produced at each line n of site s at time period t to those in the optimal solution of the upper level problem ($y\hat{p}_{t,s}^{i,n}$) is the optimal assignment in the planning problem (UP)). This last equation has the effect of limiting the search space of the binary assignment values and speeds up the solution of (PL)

$$yop_{t,s}^{i,n}\le y\hat{p}_{t,s}^{i,n}\quad i\in I,\ s\in S,\ m\in M,\ t\in T \quad (52)$$

4.4. Hybrid bi-level Lagrangean decomposition

This section describes the proposed algorithm for solving large-scale instances of problem (PS) by addressing both the temporal and spatial parts of the problem. It builds on the methodology of bi-level decomposition by using Lagrangean decomposition to solve the upper-level problem for *individual sites*. The idea of the hybrid algorithm is to add a new shipping variable $s\hat{h}t_{t,s}^{i,m}$ and constraint $sht_{t,s}^{i,m}=s\hat{h}t_{t,s}^{i,m}$ to the upper-level problem defined by Eqs. (25)–(51). The variable $sht_{t,s}^{i,m}$ is substituted for $s\hat{h}t_{t,s}^{i,m}$ in Eq. (28), yielding:

$$mds_t^{i,m}\le\sum_{s\in S}s\hat{h}t_{t,s}^{i,m}\le d_t^{i,m},\quad i\in I,\ m\in M,\ t\in T \quad (53)$$

The constraint $sht_{t,s}^{i,m}=s\hat{h}t_{t,s}^{i,m}$ is dualized using Lagrangean relaxation (Geoffrion, 1974), which has the effect of spatially decomposing the problem as shown in Fig. 2. In this way the objective function (25) becomes

$$\max \pi pl(\lambda)=\sum_{t\in T}\sum_{s\in S}\sum_{m\in M}\sum_{i\in I}\beta_{t,s}^{i,m}s\hat{h}t_{t,s}^{i,m}-\sum_{t\in T}\sum_{s\in S}\sum_{i\in I}\left(\sum_{n\in N}\omega_s^{i,n}xt_{t,s}^{i,n}+\delta_{t,s}^{i}vt_{t,s}^{i}\right)$$

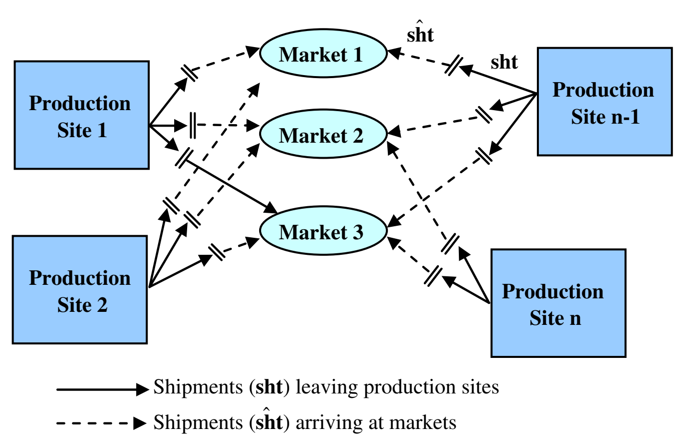

Fig. 2. Schematic representation of the relaxation of constraint $sht_{t,s}^{i,m}=s\hat{h}t_{t,s}^{i,m}$.

$$-\sum_{t\in T}\sum_{s\in S}\sum_{m\in M}\sum_{i\in I}\gamma_s^{i,m}\hat{sht}_{t,s}^{i,m}$$
$$-\sum_{t\in T}\sum_{s\in S}\sum_{i\in I}\sum_{k\in I}\sum_{n\in N}TC_s^{i,k,n}(zp_{t,s}^{i,k,n}+zzz_{t,s}^{i,k,n}-zzp_{t,s}^{i,k,n})$$
$$+\sum_{t\in T}\sum_{s\in S}\sum_{m\in M}\sum_{i\in I}\lambda_{t,s}^{i,m}(sht_{t,s}^{i,m}-\hat{sht}_{t,s}^{i,m}) \tag{54}$$

The parameter $\lambda_{t,s}^{i,m}$ is the Lagrange multiplier of the dualized constraint. We call the modified upper level problem as Upper-level Lagrangean Problem (UPL). It is defined by Eqs. (26), (27), (29)–(50), (53), (54), and it is decomposable into subproblems for individual sites ($\mathrm{UPL_s}$) $s\in S$ and individual markets ($\mathrm{UPL_m}$) $m\in M$ as shown in Fig. 2. The subproblem for an individual site s^* is as follows:

$$\max\ \pi pl_{s^*}(\lambda)=-\sum_{t\in T}\sum_{i\in I}\left(\sum_{n\in N}\omega_s^{i,n}xt_{t,s^*}^{i,n}+\delta_{t,s^*}^{i}vt_{t,s^*}^{i}\right)$$
$$-\sum_{t\in T}\sum_{i\in I}\sum_{k\in I}\sum_{n\in N}TC_{s^*}^{i,k,n}(zp_{t,s^*}^{i,k,n}+zzz_{t,s^*}^{i,k,n}-zzp_{t,s^*}^{i,k,n})$$
$$+\sum_{t\in T}\sum_{m\in M}\sum_{i\in I}\lambda_{t,s^*}^{i,m}sht_{t,s}^{i,m}$$

s.t.

Eqs. (26), (27), (29)–(50) specified for $s=s^*$ ($\mathrm{UPL_{s*}}$)

while the subproblem for market m^* corresponds to

$$\max\ \pi pl_m(\lambda)=\sum_{t\in T}\sum_{s\in S}\sum_{i\in I}\beta_s^{i,m^*}\hat{sht}_{t,s}^{i,m^*}-\sum_{t\in T}\sum_{s\in S}\sum_{i\in I}\lambda_{t,s}^{i,m^*}(\hat{sht}_{t,s}^{i,m^*})$$
$$-\sum_{t\in T}\sum_{s\in S}\sum_{i\in I}\gamma_s^{i,m^*}(\hat{sht}_{t,s}^{i,m^*})$$

s.t.

$mds_t^{i,m^*}\le\sum_{s\in S}\hat{sht}_{t,s}^{i,m^*}\le d_t^{i,m^*}\ \forall i\in I,\ t\in T$ and $\hat{sht}_{t,s}^{i,m^*}\le sht_{s,m,i}^{UB}\ \forall s\in S,\ i\in I,\ t\in T$ ($\mathrm{UPL_{m^*}}$)

It can be verified that the optimal objective value in (UPL) is equal to the sum of optimal values of ($\mathrm{UPL_s}$) $s\in S$ and ($\mathrm{UPL_m}$) $m\in M$. That is, $\max\ \pi pl(\lambda)=\sum_{s\in S}\max\ \pi pl_s(\lambda)+\sum_{m\in M}\max\ \pi pl_m(\lambda)$.

Fig. 3 contains the flow diagrams of the bi-level and hybrid algorithms (bi-level + Lagrangean decomposition).

5. Remarks

1. Computational results in Erdirik-Dogan and Grossmann (2008) show that the upper-level problem, (UP) in this paper, is the bottleneck in the bi-level decomposition algorithm. Speeding up the solution of (UP) could speed up the convergence of the bi-level method. For a given $\lambda_{t,s}^{i,m}$, upper-level Lagrangean problem (UPL) can be solved for each manufacturing site and each market at a time, decreasing the computation time of (UP).
2. The value of $\pi pl(\lambda)$ in (UPL) for any λ is guaranteed to yield an upper bound for πp in (UP) (Guiganrd and Kim, 1987). We also know that πp in (UP) is an upper bound to π in (PS). Since $\pi pl(\lambda)\ge\pi p\ge\pi$ for any λ, (UPL) is always an upper bound to (PS).
3. Terrazas-Moreno et al. (2011) show that the optimal Lagrange multiplier of the relaxed constraint $sht_{t,s}^{i,m}=\hat{sht}_{t,s}^{i,m}$ can be approximated by the transfer price of product i between site s and market m. A value close to this transfer price can be obtained by adding to problem (UPL) $sht_{t,s}^{i,m}=\hat{sht}_{t,s}^{i,m}$, solving a linear relaxation of the augmented problem, and reading the optimal dual variable of $sht_{t,s}^{i,m}=\hat{sht}_{t,s}^{i,m}$.
4. The solution to the optimization problem $\min_{\lambda}\pi pl(\lambda)$ yields the tightest upper bound to (UP). The value of λ initialized as described in the previous remark, can be improved using subgradient method (Held et al., 1974) to bring closer to the optimal multiplier λ^*,where $\pi pl(\lambda *)=\min_{\lambda}\pi pl(\lambda)$.
5. The solution πs of lower-level problem (PL) with assignment variables fixed to those obtained in the solution of the upper-level problem ($yop_{t,s}^{i,n}=\hat{yp}_{t,s}^{i,n}$) provides a heuristic lower bound that is valid on π in the original planning and scheduling problem (PS). Considering remark 2, $\pi s\le\pi\le\pi p\le\pi pl$. The hybrid algorithm described in this paper uses πs as a lower bound and πpl as an upper bound to (PS), and provides an iterative procedure that can minimize the magnitude of $\pi pl-\pi s$.

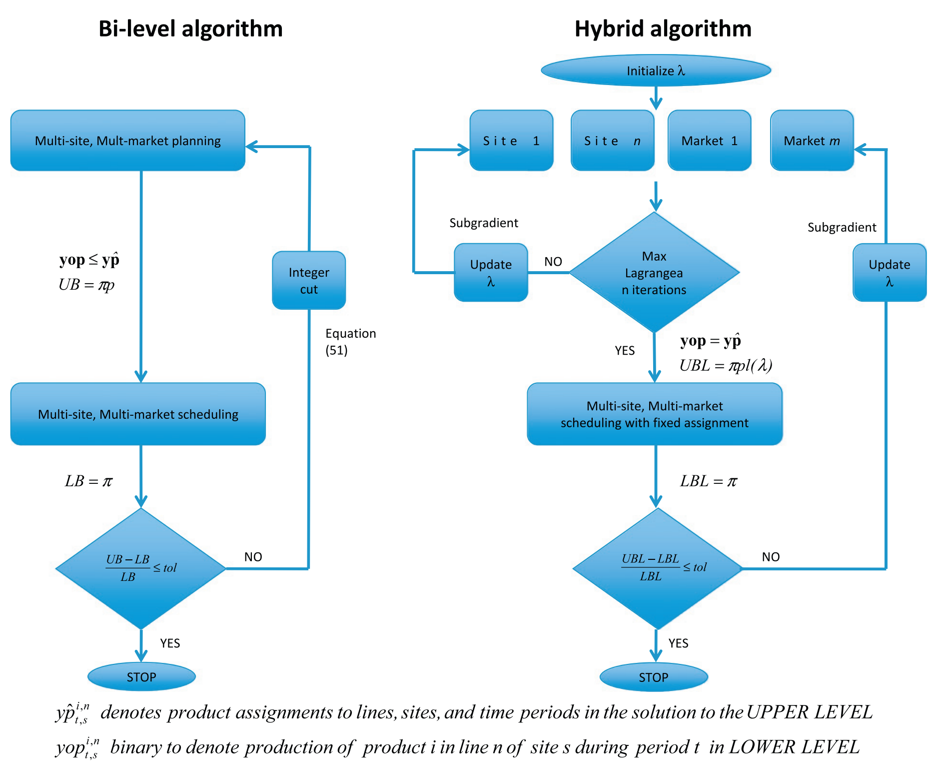

Fig. 3. Comparison of bi-level and hybrid algorithms.

5.1. Algorithmic steps

The detailed description of the hybrid algorithm that is contained in Fig. 3 is as follows:

Step 1: Initialize algorithm parameters:

B^{it}	*maximum number of bi-level iterations*
L^{it}	*maximum number of Lagrangean iterations at each bi-level iteration*
UBL^1	*initial upper bound*
LBL^1	*initial lower bound*
$r=1,\dots,B^{it}$	*bi-level iteration index*
$q=1,\dots,L^{it}$	*Lagrangean iteration index*
$mds_t^{i,m}$	*minimum required percentage of demand satisfaction*

Step 2: Solve the linear relaxation of the following problem:

$$\begin{aligned}\max \pi pl(\lambda) = &\sum_{t\in T}\sum_{s\in S}\sum_{m\in M}\sum_{i\in I}\beta_{t,s}^{i,m}\hat{sht}_{t,s}^{i,m}\\ &-\sum_{t\in T}\sum_{s\in S}\sum_{i\in I}\left(\sum_{n\in N}\omega_s^{i,n}xt_{t,s}^{i,n}+\delta_s^i vt_{t,s}^i\right)\\ &-\sum_{t\in T}\sum_{s\in S}\sum_{m\in M}\sum_{i\in I}\gamma_s^{i,m}\hat{sht}_{t,s}^{i,m}\\ &-\sum_{t\in T}\sum_{s\in S}\sum_{i\in I}\sum_{k\in I}\sum_{n\in N}TC_s^{i,k,n}(zp_{t,s}^{i,k,n}+zzz_{t,s}^{i,k,n}-zzp_{t,s}^{i,k,n})\\ &+\sum_{t\in T}\sum_{s\in S}\sum_{m\in M}\sum_{i\in I}\lambda_{t,s}^{i,m}(sht_{t,s}^{i,m}-\hat{sht}_{t,s}^{i,m})\ \text{(UPLa)}\end{aligned}$$

s.t.

Eqs. (26),(27),(29)–(50),(53), $sht_{t,s}^{i,m}=\hat{sht}_{t,s}^{i,m}\forall i\in I,\ m\in M,\quad t\in T,\ s\in S$

Step 3: Assign the value of the optimal dual multiplier of constraint $sht_{t,s}^{i,m}=\hat{sht}_{t,s}^{i,m}$ to $\lambda_{t,s}^{i,m}$ for iteration $q=1$, $r=1$.

Step 4

For $r=1,\dots,B^{it}$,
  For $q=1,\dots,L^{it}$,
    With $\lambda_{t,s}^{i,m,q,r}$ solve the site subproblems (UPL_s) $\forall s\in S$ and market subproblems (UPL_m) $\forall m\in M$ and compute the upper bound as

$$\pi pl(\lambda^{q,r})=\sum_{s\in S}\pi pl_s(\lambda^{q,r})+\sum_{m\in M}\pi pl_m(\lambda^{q,r}).$$

    If $\pi pl(\lambda^{q,r})\le UBL^r$,

$$UBL^{r+1}=\pi pl(\lambda^{q,r})$$
$$\hat{yp}_{t,s}^{i,n,r}=yp_{t,s}^{i,n}$$

    end if

$$\lambda_{t,s}^{i,m,q+1,r}=\lambda_{t,s}^{i,m,q,r}+\varepsilon\frac{\pi pl(\lambda^{q,r})-LBL^r}{\sum_{t\in T}\sum_{s\in S}\sum_{m\in M}\sum_{i\in I}(sht_{t,s}^{i,m}-\hat{sht}_{t,s}^{i,m})^2}\times(sht_{t,s}^{i,m}-\hat{sht}_{t,s}^{i,m})$$

  end For
  $q=1$

$$\lambda_{t,s}^{i,m,q+1,r}=\lambda_{t,s}^{i,m,L^{it},r}$$

Solve (PL) with fixed product assignment ($yop_{t,s}^{i,n}=\hat{yp}_{t,s}^{i,n,r}$)
If $\pi\ge LBL^r$
  $LBL^{r+1}=\pi$
end if

If $\dfrac{UBL^{r+1}-LBL^{r+1}}{|LBL^{r+1}|}\le tol$
  STOP
Else
  $r=r+1$
End if
End For

Remark: We keep the set R_{cut} empty at all times to avoid generating integer cuts in Eq. (51). The reason is as follows. In the bi-level decomposition Eq. (51) has the objective of avoiding the same integer solution and certain subset and superset of the solution obtained in the upper level problem at previous iterations. Since the coefficients in the objective function of the upper level problem (Eq. (25)) do not change, it is guaranteed that the integer cuts do not cut off better solutions (higher in maximization problems) than those found in previous iterations. When the upper level is decomposed using Lagrangean decomposition, the parameter $\lambda_{t,s}^{i,m}$ that stands for the Lagrange multiplier is updated at every iteration in the decomposable objective function (Eq. (54)). Since some of the coefficients in the objective function (namely, $\lambda_{t,s}^{i,m}$) change at every iteration, it is not possible to guarantee that a better optimal solution in new iterations will not correspond to the same, or a subset or superset, of a previous integer solution. This fact destroys the properties that allow the use of Eq. (51) in problem (UP).

6. Case studies

Example 1. A small manufacturing and distribution network consists of three production sites with single production lines serving two markets that demand products A–C. The objective is to obtain the optimal planning and detailed production schedule for 4 weeks.

Table 1 shows the maximum production rates in each site, while Tables 2a and 2b contain values of market demand for all products and the minimum levels of demands that must be satisfied. Tables 3 and 4 show the market price as well as the production, inventory, and distribution costs. Finally, Table 5 contains transition times and costs.

Table 1
Maximum production rates.

Product	Site 1 (kg/h)	Site 2 (kg/h)	Site 3 (kg/h)
A	1110	4860	2780
B	4170	3470	4860
C	4170	4270	1390

Table 2a
Demand for products at markets.

Product	Week	Market 1 (10^3 kg/week)	Market 2 (10^3 kg/week)
A	1	175.0	187.5
	2	275.0	150.0
	3	500.0	600.0
	4	750.0	750.0
B	1	200.0	100.0
	2	150.0	312.5
	3	600.0	600.0
	4	900.0	950.0
C	1	175.0	200.0
	2	100.0	50.0
	3	700.0	700.0
	4	950.0	990.0

Table 2b
Minimum percentage of demands that must be satisfied.

Product	Market 1 (%)	Market 2 (%)
A	50	50
B	50	50
C	50	50

Table 3
Market price and production, and inventory costs.

Product		Site 1 (10^{-3} \$/kg)	Site 2 (10^{-3} \$/kg)	Site 3 (10^{-3} \$/kg)	Market 1 (10^{-3} \$/kg)	Market 2 (10^{-3} \$/kg)
A	Price				20	23
	Production	2.5	3.5	3.5		
	Inventory	5.0	5.0	5.0		
B	Price				12	19
	Production	4.5	4.0	4.5		
	Inventory	5.0	5.0	5.0		
C	Price				16	15
	Production	3.5	2.5	4.5		
	Inventory	6.0	6.0	6.0		

Table 4
Distribution costs.

Product		Site 1 (10^{-3} \$/kg)	Site 2 (10^{-3} \$/kg)	Site 3 (10^{-3} \$/kg)
A	Market 1	4.0	3.0	3.0
	Market 2	5.0	6.5	4.0
B	Market 1	2.0	1.0	2.0
	Market 2	5.0	6.0	5.5
C	Market 1	5.0	4.0	6.5
	Market 2	2.5	2.0	3.0

Table 5
Transition time and cost.

	From/to	A (h)	B (h)	C (h)
Transition time	A	0	24	12
	B	24	0	12
	C	24	24	0
	From/to	A (\$)	B (\$)	C (\$)
Transition cost[a]	A	0	100	200
	B	150	0	300
	C	100	250	0

[a] Transition costs are independent of amount of product.

The market demand is low during the first two weeks and high in the last two. The spare capacity in the first couple of weeks can be used to build inventories and take advantage of the later increase in demand. By requiring that a minimum of 50% of demand be satisfied, we prevent the less profitable products from being completely ignored in favor of more profitable ones. This consideration is important in industrial settings where a minimum demand must be satisfied to remain competitive.

Tables 3–5 show that the production of every product is profitable. Production capacity and market demand are such that production sites can be run at, or close to, their maximum capacity, especially in the third and fourth weeks. Therefore, the optimal solution will involve the best possible utilization of production capacity to produce and sell as much as possible.

Example 1 is a small case study so it is possible to obtain an optimal solution without decomposition techniques. Its small size also makes it easy to determine the optimal solution. We use bi-level and bi-level-Lagrangean decompositions to solve Example 1, in addition to the full space model, with the objective of comparing each of the solutions. It is worth to clarify that the hybrid algorithm is not expected to be faster than the full space model in this example. It is almost always the case that the comparative performance of decomposition methods versus full space models improves as problem size grows.

The full space model consists of 144 discrete variables, 1117 continuous variables, and 1303 constraints. We solve it using CPLEX 12.2.0 in GAMS version 23.6 for Windows, using an Intel Core i7 CPU at 2.93 GHz, and 4.00 GB of RAM. We use the same software and hardware for all the calculations in this paper. The solution time is less than 1 CPU second.

The following analysis corresponds to the solution obtained using the full space model. Fig. 4 is a Gantt chart that corresponds to the optimal production schedule for every site. Product A is assigned more production time since it is the most profitable. This is expected since the objective function of the problem is profit. In this small example the number of products and production sites is equal, so there is no need for many product transitions. In fact, after week 2, each site is dedicated to a single product. Fig. 5 reveals that there is spare capacity in the production sites during weeks 1, 2 and 3, so the product inventory is built up in preparation for high demand at the end of the time horizon.

Table 6 shows the sales as a percentage of the demand in the markets (Table 2). The information in this table confirms that during weeks 1 and 2 all the demand is satisfied. During weeks 3 and 4 the demand exceeds production capacity, and even with inventories carried over from the first two weeks, it is not possible to satisfy all demands. The optimal profit calculated as revenues minus costs is shown in Table 7.

Next, the bi-level and bi-level-Lagragean decompositions are used to solve Example 1. The objective is to compare the solutions obtained with these two methods. The advantage in terms of

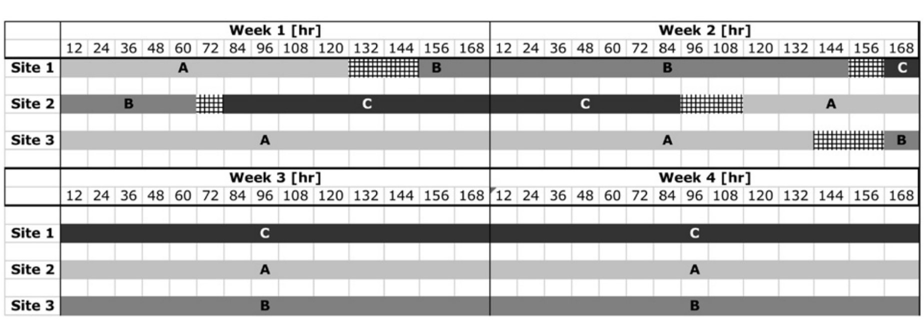

Fig. 4. Optimal production schedule.

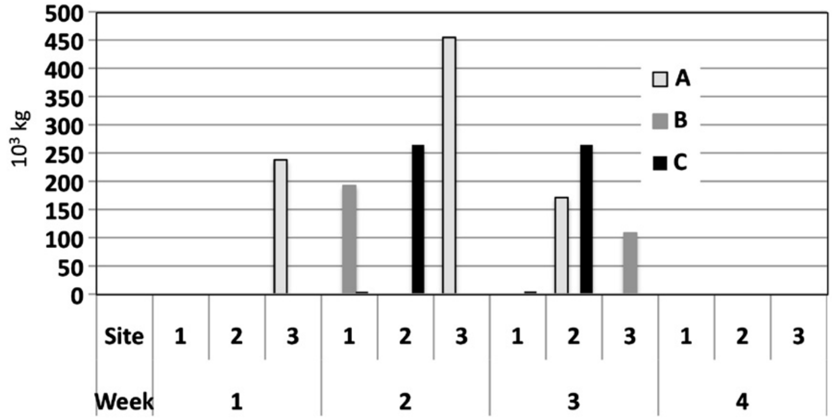

Fig. 5. End of week inventories.

Table 6
Sales as percentage of demand.

Product	Week	Market 1	Market 2
A	1	100%	100%
	2	100%	100%
	3	100%	100%
	4	82%	50%
BB	1	100%	100%
	2	100%	100%
	3	50%	100%
	4	50%	50%
C	1	100%	100%
	2	100%	100%
	3	50%	50%
	4	50%	50%

Table 7
Optimal profit for Example 1.

Revenues ($)	136,928
Operating costs ($)	28,417
Inventory costs ($)	9051
Transition costs ($)	900
Distribution costs ($)	28,849
Profit ($)	69,711

Table 8
Value of upper and lower bounds in $ for each iteration of the hybrid algorithm.

Bi-level iterations	Bound	Lagrangean iterations for solving the upper level problem				
		1	2	3	4	5
1	Upper	71,029	71,5 11	71,615	71,816	71,582
	Lower	Infeasible				
2	Upper	71,782	71,394	72,224	71,656	71,636
	Lower	67,136				
3	Upper	71,523	71,183	71,011	70,931	70,926
	Lower	63,087				
4	Upper	70,793	70,847	70,824	**70,726**	70,805
	Lower	**69,513**				

computational time, particularly of the hybrid method, will be demonstrated in Examples 2–4.

The bi-level decomposition is implemented as described in earlier sections of this paper. The upper level problem has 324 discrete variables, 499 continuous variables, and 838 constraints. As mentioned before, the upper level planning problem contains an aggregated sequencing model that is able to take into account sequence dependent product transitions. On the other hand, the lower level corresponds to the full space model where the search space of the binary variables is reduced according to the solution of the upper level problem. The lower level has 144 discrete variables, 1117 continuous variables, and 1339 constraints. The bi-level decomposition algorithm converges in one single iteration (upper and lower level have the same solution) and takes less than one CPU second in total. The detailed schedule and optimal profit is the same as the full space problem.

The bi-level Lagrangean decomposition algorithm takes 12 s to converge within a 2% tolerance between the upper and lower level solutions. It is worth mentioning that it is expected for this algorithm to perform worse than the solution using the full space model, and even the bi-level algorithm by itself, *for small examples like Example* 1. Major reason is that the algorithm requires several iterations to obtain a relaxed solution to the upper level problem of the bi-level algorithm. Relaxing the upper level pays off with less computational time for larger example. In this example, we use a maximum of 20 bi-level iterations (although the algorithm usually converges in fewer) and 5 Lagrangean iterations per bi-level iteration. Table 8 shows the value of the bounds for different iterations. After 4 major, bi-level iterations, the bound between the solution to the relaxed upper level problem and the solution to the lower level problem are within a 2% tolerance. The best upper and lower bounds in Table 8 are highlighted.

Remark: The duality gap between the optimal solution of the upper level problem and the solution to the Lagrangean relaxation of the upper level problem makes convergence to 0% tolerance between upper and lower level problems unlikely. The trade-off is that *in larger problems* the solution to the relaxed upper level problem can be more efficient than the direct solution.

Remark 3 in the Section 4 states that by initializing the Lagrange multipliers $\lambda_{t,s}^{i,m}$, with the optimal value of the dual variable that corresponds to $sht_{t,s}^{i,m} = \hat{sht}_{t,s}^{i,m}$ in the linear relaxation of (UPL) we obtain a value close to the transfer value between site s and market m of product i in time period t. Furthermore, it is also stated that this value is close to the optimum for the corresponding Lagrange multiplier. Terrazas-Moreno et al. (2011) explain how in a similar network where the maximum sales are limited by market demand, the equilibrium transfer prices are close to the production costs $\alpha_{t,s}^{i}$, whereas in the case where sales are limited by production capacity, the equilibrium transfer prices are close to the product price minus the shipment cost. We illustrate this using product A transferred between site 1 and market 1 in all time periods as an example. Table 9 contains, for the example just mentioned, the initial value for the Lagrange multipliers ($\lambda_{t,1}^{A,1}$), their best value found (corresponding to bi-level iteration 4 and Lagrangean iteration 4 in Table 8), as well as the magnitudes of the production costs ($\alpha_{1,t}^{A}$), and of the product price minus the average shipment cost ($\beta_t^{A,1}-((\gamma_{t,1}^{A,1}-\gamma_{t,2}^{A,1})/2)$). The first observation from Table 9 is that the initial and best values for Lagrange multipliers are practically the same, validating the choice of initialization using the dual variables of the linear relaxation. The second observation is that in the first week the Lagrange multiplier is of the order of magnitude of the production cost, and in the last two it is of the order of magnitude of the price minus shipment cost value. This is consistent with the beginning of the time horizon being low demand (sales are limited by demand) and the later periods being high demand (sales limited by production capacity). The differences between the multipliers and the transfer prices calculated as $\alpha_{1,t}^{A}$ or $\beta_t^{A,1}-((\gamma_{t,1}^{A,1}-\gamma_{t,2}^{A,1})/2)$ can be attributed to the effect of transition cost (which must be added to the production cost) and inventory costs. The effect of inventory cost is more important in period 2, which corresponds to a low demand week, but where the Lagrange multipliers are more than twice the production costs.

Fig. 6 shows the detailed scheduling obtained by using the hybrid algorithm; it corresponds to the best lower bound of \$69,513. Although the schedule is not exactly the same as that in Fig. 4, it follows the same pattern: the first couple of weeks there are a number of transitions until in the last two weeks each site is dedicated to the production of one product. Also, manufacturing of product A is favored since it is the most profitable. Comparing Figs. 5 and 7 confirms that the solution obtained with the hybrid algorithm follows the same pattern of inventory building during the first weeks. The main difference is that site 1 is idle during a fraction of the first week in the schedule obtained by the hybrid algorithm. It is worth clarifying that in this work we do not consider time required for start-up. Overall, the bi-level-Lagrangean algorithm obtained a close to optimal solution that could be useful if the true optimal solution is not available within reasonable computational time. In this example the full space model and the bi-level decomposition performed better. However, in the following examples we will see that finding near optimal solutions with the hybrid algorithm has computational advantages over the full space or bi-level method.

Table 9
Lagrange multiplier values, production cost and product price minus shipments costs for product A.

	$t=1$	$t=2$	$t=3$	$t=4$
Initial $\lambda_{t,1}^{A,1}$	0.003	0.008	0.013	0.019
Best$\lambda_{t,1}^{A,1}$	0.004	0.008	0.013	0.019
$\alpha_{t,1}^{A,1}$	0.003	0.003	0.003	0.003
$\beta_t^{A,1}-\frac{\gamma_{t,1}^{A,1}-\gamma_{t,2}^{A,1}}{2}$	0.017	0.017	0.017	0.017

Examples 2, 3 and 4. The objective of these examples is to compare the computational performances of the full space model, bi-level decomposition, and bi-level Lagrangean algorithm in one medium-sized and two large planning and scheduling problem instances as shown in Table 10. Table 11 contains the number of variables and constraints in the three examples. Data for solving these examples is available by request from the authors.

The size of the model of Example 2 is about one order of magnitude larger than Example 1, whereas Example 3 is about

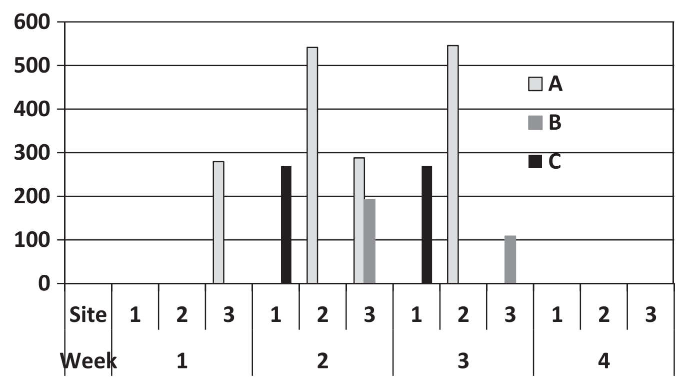

Fig. 7. End of week inventories obtained with hybrid bi-level-Lagrangean algorithm.

Table 10
Number of production sites, markets, products, and weeks in Examples 2–4.

Example	Production sites	Markets	Products	Weeks
2	3	2	3	4
3	6	6	16	24
4	6	6	25	12

Table 11
Number of variables and constraints in full space models of Examples 2–4.

	Discrete variables	Continuous variables	Constraints
Example 2	3264	64,857	28,694
Example 3	39,168	799,489	353,563
Example 4	46,800	1,369,801	419,389

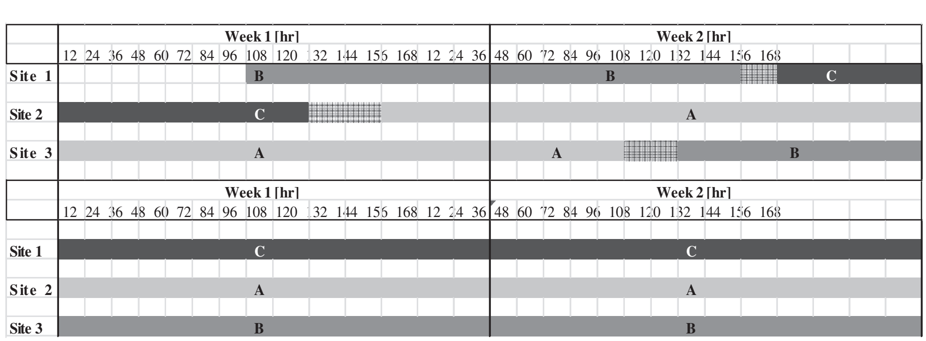

Fig. 6. Optimal production schedule obtained with hybrid bi-level-Lagrangean algorithm.

one order of magnitude larger than Example 2. Example 4 is about the same order of magnitude as the previous one; it includes more products but the planning horizon is shorter. The idea of these differences is to analyze the effect of model size on the comparative performance of the full space model, bi-level decomposition, and bi-level-Lagrangean decomposition.

Table 12 shows that the upper level subproblem of the bi-level method has more binary variables than the full space model. This is a result of the traveling salesman problem (TSP) sequencing constraints introduced into the model to capture the sequence-dependent changeovers. These sequencing constraints correspond to Eqs. (31)–(43). In other words, the price to pay for a tight upper level planning problem is an increase in the number of binary variables required to model sequence-dependent transitions. Precisely because of this fact, it is worth using spatial Lagrangean decomposition on the upper level (planning) problem, as is done in the hybrid bi-level-Lagrangean decomposition algorithm. Note that the size of the lower level scheduling problem is very similar in both decomposition schemes. Table 13 contains all the parameters that were used for the algorithm. The numerical results are summarized in Table 14.

In the three examples we observe that the bi-level decomposition yields a solution closer to the optimum than the hybrid algorithm. This is expected since the hybrid algorithm is basically a bi-level decomposition with a relaxed upper level problem. In Example 2 both decomposition algorithms require almost the same computational time. Even in cases such as this, the hybrid decomposition offers the practical advantage that planning can be solved separately at each site. If a good set of Lagrange multipliers (transfer prices) can be obtained, the resulting production plan can be close to optimal. This represents a significant advantage in industrial settings where the transfer of detailed information between sites is challenging.

Table 12
Size of subproblems in bi-level and bi-level-Lagrangean algorithm.

Algorithm	Subproblem	Discrete variables	Continuous variables	Constraints
Example 2				
Bi-level	Upper level	6720	8237	8765
	Lower level	3264	59,713	26,582
Hybrid	Site subproblem	2240	1413	2821
	Market subproblem	–	193	129
	Lower level	3264	59,713	26,534
Example 3				
Bi-level	Upper level	80,640	60,049	106,705
	Lower level	39,168	725,761	320,539
Hybrid	Site subproblem	13,440	10,009	17,001
	Market subproblem	–	2305	769
	Lower level	39,168	725,761	320,443
Example 4				
Bi-level	Upper level	95,400	163,073	115,561
	Lower level	46,800	1,279,799	379,939
Hybrid	Site subproblem	15,900	10,513	18,636
	Market subproblem	–	1801	601
	Lower level	46,800	1,279,799	379,789

Table 13
Values of the parameters used in the bi-level Lagrangean algorithm.

Parameter	Example 2	Example 3	Example 4
B^{it}—Max bi-level iterations	20	20	20
L^{it}—Lagrangean iterations in each bi-level iterations	3	2	2
UB^1—Initial upper bound	$+\infty$	$+\infty$	$+\infty$
LB^1—Initial lower bound	1000	5000	6000
mds^i_m—Min percentage of demand satisfaction	0	0	0
tol—Convergence tolerance (%)	2	2	2

Table 14
Numerical results for Examples 2–4.

	Algorithm	Best solution ($)	Solution time (CPU s)	Gap (%)	Iterations
Example 2	Full space[a]	2184	11	0	–
	Bi-level	2184	3	0	1
	Hybrid	2157	5	1.2	4
Example 3	Full space[a]	19,885	1056	0	–
	Bi-level	19,819	390	0.3	1
	Hybrid	19,556	79	1.7	3
Example 4	Full space[a]	9986	3492	1.3	–
	Bi-level	10,105	667	0	1
	Hybrid	10,007	243	1.1	6

[a] 2% tolerance in CPLEX.

In Examples 3 and 4 the hybrid algorithm requires less time than the bi-level decomposition, and between one and two orders of magnitude less than the full space model. These results are summarized in Fig. 8. Table 15 shows the percentage of computational time spent in solving each of the subproblems with the two decomposition algorithms. It is interesting to see how in the bi-level decomposition most of the time in the larger examples is spent solving the lower level problem. This is in contrast with the single-site results reported by Erdirik-Dogan and Grossmann (2008). It seems that the effect of adding multiple sites and markets to the problems particularly affects the complexity of the lower level problem. More importantly, from Table 15 we can conclude that fixing the assignments obtained from the upper level in the lower level, as we do for the hybrid decomposition ($yop = yp$), is a heuristic rule that greatly decreases the solution time required to obtain a feasible solution to the lower level problem.

7. Conclusions

This paper has proposed two algorithms for solving multi-site multi-market simultaneous planning and scheduling problems involving continuous multiproduct plants. The idea is to use a bi-level decomposition or a hybrid method where the upper level, involving the planning problem, is decomposed through spatial Lagrangean relaxation (Jackson and Grossmann, 2003). We initialize the Lagrange multiplier of the relaxed constraint using the linear relaxation. It has been observed that this value is close to the equilibrium transfer price between sites and markets, which in turn is close to the optimal Lagrange multiplier (Terrazas-Moreno et al., 2011). We show in a small example that the solution obtained with this hybrid decomposition method is similar to the one obtained with the full space planning and scheduling model. In other examples we show that the computational performance of the hybrid algorithm is similar to bi-level decomposition in medium-sized problems (Example 2) and faster in large-scale problems (Examples 3 and 4). In all four case studies the hybrid algorithm converges within a 2% of optimality tolerance. We believe that this method is efficient for problems with tens of thousands of discrete variables, and hundreds of thousands of continuous variables and constraints. Yet, it is simple enough to be a practical alternative for obtaining tight bounds for industrial-sized planning and scheduling problems.

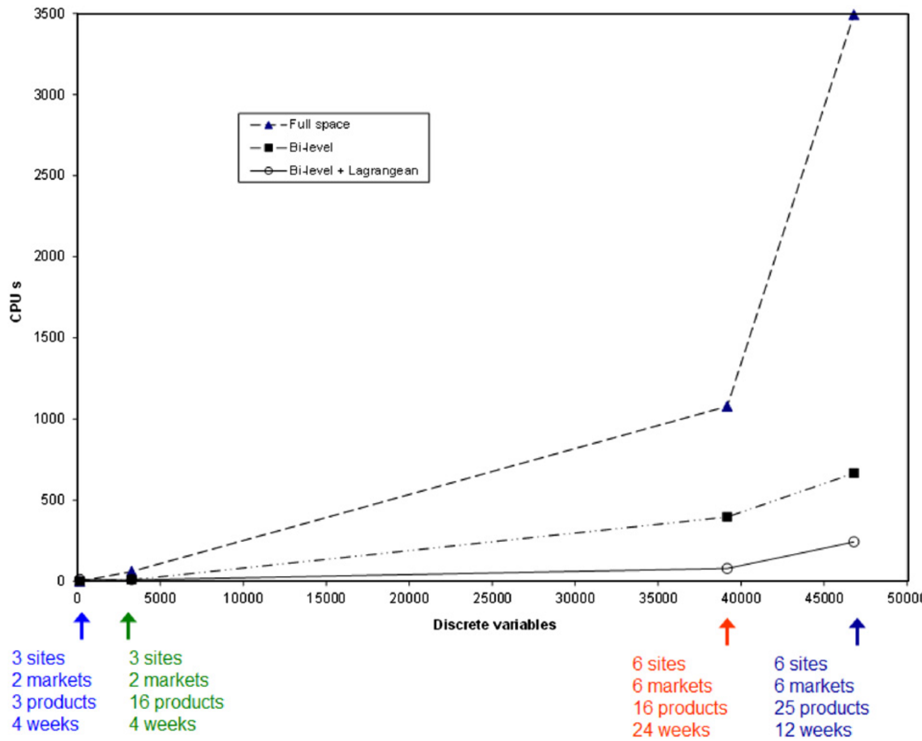

Fig. 8. Solution time for case studies using full space, bi-level, and hybrid algorithms.

Table 15
Percent of time spent in each subproblem for bi-level and hybrid algorithm.

	Upper level in bi-level (% CPU time)	Lower level in bi-level (% CPU time)	Upper level in hybrid (% CPU time)	Lower level in hybrid (% CPU time)
Example 2	52	48	73	27
Example 3	23	77	81	19
Example 4	33	67	83	17

Acknowledgment

The authors would like to thank Dow Chemical for financial support of this project.

References

Birewar, D.B., Grossmann, I.E., 1990. Simultaneous production planning and scheduling in multiproduct batch plants. Industrial and Engineering Chemistry Research 29, 570–580.
Erdirik-Dogan, M., Grossmann, I.E., 2006. A decomposition method for simultaneous planning and scheduling of single-stage continuous multiproduct plants. Industrial and Engineering Chemistry Research 45, 299–315.
Erdirik-Dogan, M., Grossmann, I.E., 2008. Simultaneous planning and scheduling of single-stage multi-product continuous plants with parallel lines. Computers and Chemical Engineering 32, 2664–2683.
Garey, M.R., Johnson, D.S., 1979. Computers and Intractability A Guide to the Theory of NP-Completeness. W.H. Freeman and Co, New York.
Geoffrion, A.M., 1974. Lagrangian relaxation for integer programming. Mathematical Programming Study 2, 82–114.
Grossmann, I.E., van den Heever, S.A., Harjunkoski, I., 2002. Discrete optimization methods and their role in the integration of planning and scheduling. AIChE Symposium Series 326, 150–168.
Grossmann, I.E., 2005. Enterprise-wide optimization: a new frontier in process systems engineering. AIChE Journal 51, 1846–1857.
Guignard, M., Kim, S., 1987. Lagrangean decomposition: a model yielding stronger lagrangean bounds. Mathematical Programming 39, 215–228.
Held, M., Wolfe, P., Crowder, H.P., 1974. Validation of subgradient optimization. Mathematical Programming 6, 62–88.
Hooker, J.N., 2005. A hybrid method for planning and scheduling. Constraints 10, 385–401.
Jackson, J.R., Grossmann, I.E., 2003. Temporal decomposition scheme for nonlinear multisite production planning and distribution models. Industrial and Engineering Chemistry Research 42, 3045–3055.
Kopanos, G.M., Puigjaner, L., Maravelias, C.T., 2011. Production planning and scheduling of parallel continuous processes with product families. Industrial and Engineering Chemistry Research 50, 1369–1378.
Li, Z., Ierapetritou, M., 2009. Integrated production planning and scheduling using a decomposition framework. Chemical Engineering Science 64, 3585–3597.
Li, Z., Ierapetritou, M., 2010. Production planning and scheduling integration through augmented Lagrangian optimization. Computers and Chemical Engineering 34, 996–1006.
Lima, R.M., Grossmann, I.E., Jiao, Y., 2011. Long-term scheduling of a single-unit multi-product continuous process to manufacture high performance glass. Computers and Chemical Engineering 35, 554–574.
Liu, S., Pinto, J.M., Papageorgiou, L.G., 2008. A TSP-based MILP model for the medium-term planning of single-stage continuous multiproduct plants. Industrial and Engineering Chemistry Research 47, 7733–7743.
Liu, S., Pinto, J.M., Papageorgiou, L.G., 2010. A TSP-based MILP model for the medium-term planning of single-stage continuous multiproduct plants with parallel units. Computational Management Science 7, 407–435.
Maravelias, C.T., Sung, C., 2009. Integration of planning and scheduling: overview, challenges, and opportunities. Computers and Chemical Engineering 33, 1919–1930.
Stefansson, H., Shah, N., Jensson, P., 2006. Multiscale planning and scheduling in the secondary pharmaceutical industry. AIChE Journal 52, 4133–4149.
Sung, C., Maravelias, C.T., 2007. An attainable region approach for production planning of multiproduct processes. AIChE Journal 53, 1298–1315.
Terrazas-Moreno, S., Trotter, P.A., Grossmann, I.E., 2011. Temporal and spatial Lagrangean decomposition in multi-site multi-period production planning problems with sequence-dependent changeovers. Computers and Chemical Engineering. doi:10.1016/j.compchemeng.2011.01.004.
Verderame, P.M., Floudas, C.A., 2008. Integrated operational planning and medium-term scheduling for large-scale industrial batch plant. Industrial and Engineering Chemistry Research 47, 4845–4860.