Optimal production scheduling with customer-driven demand substitution

Journal:	*International Journal of Production Research*
Manuscript ID	TPRS-2015-IJPR-1899
Manuscript Type:	Special Issue Paper
Date Submitted by the Author:	17-Nov-2015
Complete List of Authors:	Zeppetella, Luca; University of Modena and Reggio Emilia, Engineering Faculty, Department of Engineering Sciences and Methods Gebennini, Elisa; University of Modena and Reggio Emilia, Engineering Faculty, Department of Engineering Sciences and Methods Grassi, Andrea; Universita' di Modena e Reggio Emilia, Dip.to di Scienze e Metodi dell'Ingegneria; Rimini, Bianca; University of Modena and Reggio Emilia, Department of Engineering Sciences and Methods
Keywords:	PRODUCT MIX, SCHEDULING, LINEAR PROGRAMMING

Keywords (user):	DEMAND SUBSTITUTION, PRODUCT ASSORTMENT

To appear in the *International Journal of Production Research*
Vol. 00, No. 00, 00 Month 20XX, 1–26

Optimal production scheduling with customer-driven demand substitution

L. Zeppetella*, E. Gebennini, A. Grassi, B. Rimini

Dipartimento di Scienze e Metodi dell'Ingegneria, Università degli Studi di Modena e Reggio Emilia
Via Amendola 2, 42122 Reggio Emilia, Italy

(*Received 00 Month 20XX; accepted 00 Month 20XX*)

This paper deals with the production scheduling problem with customer-driven demand substitution.

We distinguish between the so-called *long-term* product assortment, that is the whole set of alternative product options that a system is able to produce in the long term, and the *short-term* product assortment, that is the subset of options that are currently available. In such a context, typical of fields where *high-variety* strategies are applied, the first-choice option of the customer could be unavailable at a certain time instant. In that case, we assume that the customer has the possibility to substitute with other options or, if he/she is not willing to substitute, either a lost sale occurs or a backorder is placed.

Thus, this paper proposes two mixed-integer linear programming models (for both the lost sale case and the backorder case) for optimizing the production schedule and, consequently, the *short-term* product assortment, by jointly considering (i) capacity and production constraints and costs on one hand, (ii) and demand substitution issues on the other hand.

An extensive experimental analysis has allowed us to evaluate the models' behaviour in a wide variety of operative scenarios and to draw some concluding remarks.

Keywords: Product mix; Scheduling; Linear programming; Demand substitution; Product assortment

1. Introduction

Nowadays, customer expectations and buying behaviour should be taken into consideration not only in the final stages, but along all the processes that are needed to provide the customer with value-added products and services. Such a statement is emphasized by the widespread application of the so-called *high-variety strategies* (see Kahn 1998, for more detail), which take the form of *customization* strategies, aiming at producing exactly the product option that the customer desires, and *variety-seeking* strategies, aiming at producing more variety in a product category in order to allow each consumer to enjoy a diversity of options over time.

In particular, the notion of customer-driven demand substitution, strictly related to customization and variety, should already be considered in the production management process. However, to the authors' knowledge, while these concepts are widely investigated in the literature on retailing (and marketing in general), it is much less common to include product substitution and product assortment issues in production planning and control.

Thus, this paper aims at contributing to the literature by addressing the production scheduling problem under substitutable demand. Specifically, the focus is on the optimal allocation of the available production resources over time by considering (i) capacity and production constraints and costs on one hand, and (ii) product assortment requirements on the other hand, when demand-substitution effects dominate.

In particular, a manufacturing system able to produce a variety of alternative options (or variants) of the same product category is studied under a dynamic perspective. In such a system we can distinguish between:

*Corresponding author. Email: luca.zeppetella@unimore.it

- *long-term* product assortment: the whole set of options that the system is able to produce in the long term (according to its production resources);
- *short-term* product assortment: the subset of options that the system offers to its customers at a time.

In many real-world systems the *short-term* product assortment does not always coincide with the *long-term* product assortment. This can be due to reasons of cost (e.g., inventory holding and setup costs) and/or capacity constraints, so that only a subset of the possible product options are available at the same time. On the other hand, as stated above, it is always important (i) to meet the customer's expectations by offering his/her favourite option (the first-choice) or a similar one and (ii) to rotate the products in the short term so that all the product options can be eventually offered to the customer. In this way, the customer is satisfied (or, at least, partially satisfied) in the short term (according to the *customization* strategies) and he/she has the opportunity of choosing among a wider variety of products in the long term (according to *variety-seeking* strategies).

Hence, the problem is to dynamically optimize the *short-term* product assortment under a given *long-term* product assortment by jointly considering capacity and production constraints, economic issues and customer-driven demand substitution.

An early study of this problem can be found in Gebennini et al. (2015) where the production schedule and, consequently, the *short-term* product assortment are optimized for an application in the brewery industry. In that case the lost sales, that may occur if a product is not available and the customer is not willing to substitute, have been taken into account by introducing "dummy" products.

In the present paper, two distinct operative situations in which the manufacturing system under study can operate are addressed separately, i.e.,

- the situation with lost sales, where any demand that is not satisfied in a period (by the first-choice option or substitutive options) is considered lost: the model proposed in this paper (see Section 3) improves on the model in Gebennini et al. (2015) by taking explicitly into consideration the amount of lost sales of each option in each period;
- the situation with backorders, where any demand that is not satisfied in a period (by the first-choice option or substitutive options) is backordered: a new optimization model is proposed in this paper (see Section 4).

1.1 *Literature review*

In this study the concept of "substitutable demand" plays a fundamental role.

The benefits of product substitution have long been recognized in several studies. For example, product substitution may offer opportunities for economies of scale or the possibility of inventory pooling to hedge against demand uncertainties and to help reduce safety stocks (Hsu et al. 2005).

The literature about substitutable demand distinguishes between one-way substitution and two-way substitution. One-way substitution assumes that products can be ordered based on an attribute, such as quality or speed of service, so that products with higher levels of the attribute can substitute for products with lower levels of the attribute. This kind of substitution is also called "hierarchical substitution" (see, e.g. Bitran and Dasu 1992; Bassok et al. 1999; Rao et al. 2004; Stavrulaki 2011). On the other hand, models that allow two-way substitutability enable customers to substitute among products within the same category. Two-way substitution is the perspective that we adopt in this paper. Note that in case of hierarchical substitution an adaptation cost should be taken into consideration (Tripathy et al. 1999; Jans and Degraeve 2008). Such a cost can be seen as a more general substitution cost in case of two-way substitution, as in the case under analysis where products are not characterized through quality.

Another interesting distinction that can be found in literature is between stockout-based substitution and assortment-based substitution (see, e.g., Kök and Fisher 2007; Yücel et al. 2009).

Stockout-based or dynamic substitution occurs when the customer replaces his/her favourite option with another one because it is stocked-out at the moment of the purchase decision; assortment-based or static substitution behaviour occurs when the favourite product is not in the assortment (i.e., in the portfolio of products). According with these definitions, the problem addressed in this paper should refer to the stockout-based substitution. However, in our case this distinction is not conclusive. This is because the *long-term* assortment is given and it is supposed to meet the customers' preferences on the whole. The objective is to optimize the *short-term* assortment, i.e., to choose the options of the *long-term* assortment that must be available (not stocked-out) at a certain instant of time. Thus, both inventory and assortment issues are jointly taken into consideration.

The number of substitution attempts is another feature that differentiates the models involving demand substitution issues. In general, since Kök and Fisher (2007) shows that a multi attempts model can be approximated with a single attempt substitution model with increased rates, most of the literature proposes single attempt models. On the contrary, in the present study multiple substitution attempts are taken into account in order to avoid any approximation error.

An interesting point is that while product substitution issues are widely discussed in the retail sector (see, e.g. Mantrala et al. 2009; Drezner et al. 1995), works regarding manufacturing aspects are mainly restricted to the inventory control problem (see, e.g. Bassok et al. 1999; Rao et al. 2004; Gallego et al. 2006). In fact, to the authors' knowledge, few studies include product substitution issues into production scheduling and control.

Thus, by taking inspiration from Yücel et al. (2009) (which dealt with the product assortment problem in the retail sector) and by significantly extending Gebennini et al. (2015), we developed two mixed-integer linear programming models that dynamically optimize the production schedule (and, consequently, the *short-term* product assortment) under substitutable demand, by considering both the lost sale case and the backorder case. The optimal solution is generated by minimizing the sum of setup costs, holding costs, substitution costs and lost-sale/backorder costs.

The remainder of this paper is organized as follows. Section 2 describes the problem and the main assumptions. Section 3 presents the mathematical model for the lost sale case, while in Section 4 a new model is proposed where backorders are allowed. Section 5 presents and discusses an extensive experimental analysis, and some conclusions are drawn in Section 6.

2. Problem statement

In this study we distinguish between *long-term* product assortment and *short-term* product assortment.

The *long-term* product assortment, that is the whole set of product options that can be produced and offered to the customers on a planning horizon of several periods, is supposed to be fixed and given.

On the other hand, the subset of options available (not stocked-out) at a certain instant of time, called the *short-term* product assortment, is supposed to vary over the planning horizon. In such a way, a proper rotation of product options is guaranteed, according to the *high-variety strategies* (this constraint is called "rotation requirement" in the sequel).

We assume here that, if an option is not in the *short-term* product assortment at a certain instant of time, it can be substituted by another one. Specifically, for each couple of options of the *long-term* product assortment a substitution rate is defined. On the other hand, if the customer is not willing to substitute when his/her first-choice option is not available, either a lost sale occurs or a backorder is placed. These two distinct operative situations are addressed separately throughout the paper.

Anyway, in both situations (lost sales or backorders) the objective is to schedule the production orders so that the *short-term* product assortment is optimized over the planning horizon by considering set-up costs, inventory holding costs, substitution costs and lost-sale/backorder costs.

The remaining assumptions and conventions of the two mixed-integer linear programming (MILP) models proposed in this paper can be summarized as follows:

- Time is discrete.
- The demand of all the products of the *long-term* product assortment is known and deterministic for each time unit.
- The substitution rates are known and fixed for each couple of product options. This assumption is common in stockout-based substitution models (see, e.g., Netessine and Rudi 2003). In this way we do not attempt to model the decision process of the individual customer, but we consider the aggregate customers' behaviour. Then, according to Smith and Agrawal (2000), we assume that customers choose independently of each other and the substitution rates are independent of the total number of customers in any time unit.
- Multi-level substitution is allowed. Thus, when a product options is not available its demand is redistributed in a number of distribution attempts (see Yücel et al. 2009). In particular, for each product option we have to consider: (i) the first-choice demand incoming to that option, (ii) the substitution demand incoming to that option, one for each level of substitution and (iii) the substituted demand outgoing from that option, one for each level of substitution.
- There is a given number of production resources (or stations) that work in a parallel manner and the following aspects are taken into consideration: (i) batch capacity (depending on the specific production resource), (ii) production lead time (depending on the specific option), (iii) single product option per batch.
- When a certain quantity of a product option becomes available either it is sold to the customer in that time unit or it is stored. The timing convention used to mathematically treat the products' flows is as follows: the demand is realized at the beginning of any time unit while changes in the inventory levels occur at the end of the time unit. The present formulation of the problem does not include storage capacity constraints.
- The following (linear) costs are considered:
  - set-up costs;
  - inventory holding costs;
  - substitution costs;
  - lost-sale costs (for the model proposed in Section 3) or backorder costs (for the model proposed in Section 4);

  where the last two types of costs are related to the lower degree of customer satisfaction.

  The objective is to minimize the sum of these costs over the planning horizon, both in the lost sale case (see Section 3) and the backorder case (see Section 4).

3. Model with lost sales

In this section a MILP model is proposed for the production scheduling problem under customer-driven demand substitution with lost sales. Hence, in any time unit t, if a product is not in the *short-term* product assortment or if it has run out of stock in that time unit, whether the customer substitutes the product with another one or a lost sale occurs.

3.1 *Notation*

The following notation is adopted:

Indexes

- $t = 1, \ldots, T$: time units along the planning horizon;

- $j = 1, \ldots, J$: product options of the *long-term* product assortment;
- $s = 1, \ldots, S$: production stations;
- $m = 1, \ldots, M$: levels (or attempts) of substitution (with $M \leq J$);

Variables

- x_{jst}: quantity of product option j that becomes available from station s at the beginning of time unit t;
- $x0_{jt}$: amount of satisfied demand of product option j (without substitution) in time unit t;
- xs_{mjkt}: amount of product option j used to satisfy the m-th level of substitution from option k in time unit t.
- xl_{mjt}: lost sales of product option j at the m-th level of substitution in time unit t.
- I_{jt}: inventory level of product option j at the end of time unit t;
- y_{jst}: 1, if a production order of product option j is scheduled and launched on station s in time unit t; 0, otherwise;
- z_{jst}: 1, if station s is occupied by product option j in time unit t (except for the first time unit when the production order is launched); 0, otherwise.

Input Data

- d_{jt}: demand of product option j in time unit t;
- I_j^0: inventory level of product option j at the beginning of the planning horizon;
- C_s: batch capacity of station s;
- LT_j: production lead time (in number of time units) of product option j;
- w_{jk}: substitution rate, i.e., proportion of customers whose preference is product option k that substitute option k with option j ($w_{jj} = 0$);
- w_k^l: proportion of customers whose preference is product option k and that refuse to substitute option k with any other option;
- T^r: maximum number of time units between two production order of a given product option (this parameter defines the product rotation requirement);
- c^{setup}: unit setup cost;
- c_j^{h}: unit holding cost for product option j;
- c_{mj}^{sub}: penalty cost of the m-th level of substitution from product option j;
- c_{mj}^{lost}: penalty cost of the lost sales of product option j at the m-th level of substitution.

3.2 *Model formulation*

The MILP model is formulated as follows:

$$\min \quad TC = TCP + TCH + TCS + TCL \tag{1}$$

subject to

$$TCP = \sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{s=1}^{S} y_{jst}\, c^{\text{setup}}, \tag{2}$$

$$TCH = \sum_{j=1}^{J} \frac{I_j^0 + \sum_{s=1}^{S} x_{js1} + I_{j1}}{2}\, c_j^{\text{h}} +$$

$$+ \sum_{t=2}^{T} \sum_{j=1}^{J} \frac{I_{j,t-1} + \sum_{s=1}^{S} x_{jst} + I_{jt}}{2} \, c_j^{\mathrm{h}} , \tag{3}$$

$$TCS = \sum_{t=1}^{T} \sum_{m=1}^{M} \sum_{j=1}^{J} \sum_{\substack{k=1 \\ k \neq j}}^{J} xs_{mjkt} \, c_{mj}^{\mathrm{sub}} , \tag{4}$$

$$TCL = \sum_{t=1}^{T} \sum_{m=1}^{M} \sum_{j=1}^{J} xl_{mjt} \, c_{mj}^{\mathrm{lost}} , \tag{5}$$

$$x0_{j1} + \sum_{m=1}^{M} \sum_{\substack{k=1 \\ k \neq j}}^{J} xs_{mjk1} + I_{j1} = I_j^0 + \sum_{s=1}^{S} x_{js1} , \quad \forall j , \tag{6}$$

$$x0_{jt} + \sum_{m=1}^{M} \sum_{\substack{k=1 \\ i \neq j}}^{J} xs_{mjkt} + I_{jt} = I_{j,t-1} + \sum_{s=1}^{S} x_{jst} , \quad \forall j, \forall t > 1 , \tag{7}$$

$$x0_{jt} + \sum_{m=1}^{M} \sum_{\substack{k=1 \\ i \neq j}}^{J} xs_{mkjt} + \sum_{m=1}^{M} xl_{mjt} = d_{jt} , \quad \forall j, \forall t , \tag{8}$$

$$xs_{1jkt} \leq (d_{kt} - x0_{kt}) \, w_{jk} , \quad \forall j, k, \text{ with } k \neq j, \forall t , \tag{9}$$

$$xl_{1kt} \leq (d_{kt} - x0_{kt}) \, w_k^l , \quad \forall k , \forall t , \tag{10}$$

$$xs_{2jkt} \leq (d_{kt} - x0_{kt} - \sum_{\substack{r=1 \\ r \neq j,k}}^{J} xs_{1rkt} - xl_{1kt}) \sum_{\substack{r=1 \\ r \neq j,k}}^{J} w_{rk} w_{jr} ,$$
$$\forall j, k, \text{ with } k \neq j, \forall t , \tag{11}$$

$$xl_{2kt} \leq (d_{kt} - x0_{kt} - \sum_{\substack{r=1 \\ r \neq j,k}}^{J} xs_{1rkt} - xl_{1kt}) \sum_{\substack{r=1 \\ r \neq k}}^{J} w_{rk} w_r^l ,$$
$$\forall j, k, \text{ with } k \neq j, \forall t , \tag{12}$$

$$\sum_{\tau=t-LT_j+1}^{t-1} z_{js\tau} = (LT_j - 1) \, y_{js(t-LT_j)} , \quad \forall j , \forall s , \forall t > LT_j , \tag{13}$$

$$y_{jst} \leq \left(1 - \sum_{j=1}^{J} z_{jst}\right) , \quad \forall j , \forall s , \forall t , \tag{14}$$

$$\sum_{j=1}^{J} y_{jst} \leq 1 , \quad \forall s , \forall t , \tag{15}$$

$$\sum_{j=1}^{J} z_{jst} \leq 1 , \quad \forall s , \forall t , \tag{16}$$

$$x_{jst} = C_s \, y_{js(t-LT_j)} , \quad \forall j , \forall s , \forall t > LT_j , \tag{17}$$

$$\sum_{\tau=t-T^r+1}^{t} \sum_{s=1}^{S} y_{js\tau} \geq 1 , \quad \forall j , \forall t \geq T^r , \tag{18}$$

$$x_{jst} \geq 0 , \quad \forall j , \forall s , \forall t , \tag{19}$$

$$I_{jt}\,, \quad \forall j\,, \forall t\,, \tag{20}$$
$$x0_{jt}\,, \quad \forall j\,, \forall t\,, \tag{21}$$
$$xs_{mjkt} \geq 0\,, \quad \forall j\,, \forall j,k \text{ (with } k \neq j) \;, \forall t\,, \tag{22}$$
$$y_{jst} \in [0,1] \quad \forall j\,, \forall s\,, \forall t\,, \tag{23}$$
$$z_{jst} \in [0,1] \quad \forall j\,, \forall s\,, \forall t\,. \tag{24}$$

where:

- constraint (2) defines the total setup cost;
- constraint (3) defines the total inventory holding cost, where expected inventory is calculated as the average of the initial and the final inventory levels for each time unit;
- constraint (4) defines the total substitution cost, where a unit cost is paid for the demand portion satisfied by the substitutive option;
- constraint (5) defines the total lost sale cost;
- constraints (6)-(7) guarantee the conservation of material flow for each product and time unit;
- constraint (8) states that the demand for a product option can be satisfied by the very product or with substitution, and the potential unsatisfied demand results into lost sales;
- constraints (9)-(12) represent the substitution inequalities and the lost-sales inequalities. For any level of substitution, the amount of option j that substitutes for product option k is less than or equal to a certain proportion of the unsatisfied demand of product option k. This proportion depends on the substitution rates. The remaining unsatisfied demand results into a lost sale. Substitution and lost-sales inequalities are written for each of the M levels of substitution. For the sake of clarity, as in Yücel et al. (2009), only the two substitution and lost-sales inequalities are provided;
- constraints (13)-(14) guarantee that when a production order is launched on a station, the station is occupied by that product option until the production is completed (i.e., for a number of time units equal to the product lead time) and no other production orders can be launched. Note that variable y_{jst} is 1 only when the production of product j begins on station s (i.e., the first time unit of the lead time). Variable z_{jst} is set to 1 for all the time units of the lead time except the first one (where $y_{jst} = 1$).
- constraints (15)-(16) guarantee that when a product option is in process on a station, that station is devoted to that option only;
- constraint (17) states that if a product option j becomes available at a certain time t, the corresponding production order must have been scheduled LT_j time units in advance. Then, the quantity of product option j that becomes available depends on the batch capacity related to that station;
- constraint (18) represents the rotation requirement. For each product option, at least one production order must be scheduled every T^r time units. This prevents that product options that are highly substitutable and/or with a low demand are never, or rarely, produced (with the objective of preserving the variety in the long term);
- constraints (19)-(24) define non-negative and binary variables.

The objective function is in the form of a cost minimization. The costs taken into account are the total setup cost (TCP), the total inventory holding cost (TCH), the total substitution cost (TCS) and the total lost-sale cost (TCL).

The solution of the model provides the decision maker with:

- the schedule of the production orders over the planning horizon: variable y_{jst} states, for each time unit, the production orders that should be launched on the available stations;

- the *short-term* product assortment $J_t^{\text{short-term}}$ for any time unit t in the planning horizon:

$$J_t^{\text{short-term}} = \left\{ j \in J : \sum_{s=1}^{S} x_{jst} + I_{j(t-1)} > 0 \right\}, \tag{25}$$

that, in words, represents the subset of product options that are available at the beginning of time unit t.

4. Model with backorders

In this section the model proposed in Section 3 is extended in order to allow backorders. Thus, in any time unit t, if a product is not in the *short-term* product assortment or if it has run out of stock in that time unit, whether the customer substitutes the product with another one or a backorder is placed.

4.1 *Notation*

The notation proposed in Section 3 is here retained, except for the parameters referring to the lost sales. Then, new parameters referring to backorders are introduced as follows:

- the variable xb_{mjt}, indicating the amount of product option j backordered at the m-th level of substitution in time unit t;
- the backorder rate w_j^b of any product option j, with $w_j^b = 1 - \sum_{\substack{k=1\\k\neq j}}^{J} w_{kj}$;
- the unit backlogging cost c_{mj}^{back} at the m-th level of substitution from product option j.

4.2 *Model formulation*

The MILP model is formulated as follows:

$$\min \quad TC = TCP + TCH + TCS + TCB \tag{26}$$

subject to

Constraints $(2) - (4)$,

$$TCB = \sum_{t=1}^{T} \sum_{m=1}^{M} \sum_{j=1}^{J} xb_{mjt}\, c_{mj}^{\text{back}}, \tag{27}$$

Constraints $(6) - (7)$,

$$x0_{j1} + \sum_{m=1}^{M} \left(\sum_{\substack{k=1\\k\neq j}}^{J} xs_{mkj1} + xb_{mj1} \right) = d_{jt}\,, \quad \forall j, \tag{28}$$

$$x0_{jt} + \sum_{m=1}^{M} \left(\sum_{\substack{k=1\\k\neq j}}^{J} xs_{mkjt} + xb_{mjt} \right) = d_{jt} + \sum_{m=1}^{M} xb_{mj,t-1}\,, \quad \forall j, \forall t > 1\,, \tag{29}$$

$$xs_{1jkt} \leq \left(d_{kt} + \sum_{m=1}^{M} xb_{mk,t-1} - x0_{kt} \right) w_{jk}\,, \quad \forall j, k, \text{ with } k \neq j, \forall t\,, \tag{30}$$

$$xb_{1kt} \leq \Big(d_{kt} + \sum_{m=1}^{M} xb_{mk,t-1} - x0_{kt}\Big)\, w_k^b\,, \quad \forall k, \forall t\,, \tag{31}$$

$$xs_{2jkt} \leq \Big(d_{kt} + \sum_{m=1}^{M} xb_{mk,t-1} - x0_{kt} - \sum_{\substack{r=1\\ r\neq j,k}}^{J} xs_{1rkt} - xb_{1kt}\Big) \sum_{\substack{r=1\\ r\neq j,k}}^{J} w_{rk}w_{jr}\,,$$

$$\forall j,k, \text{ with } k \neq j, \forall t\,, \tag{32}$$

$$xb_{2kt} \leq \Big(d_{kt} + \sum_{m=1}^{M} xb_{mk,t-1} - x0_{kt} - \sum_{\substack{r=1\\ r\neq k}}^{J} xs_{1rkt} - xb_{1kt}\Big) \sum_{\substack{r=1\\ r\neq k}}^{J} w_{rk}w_r^b\,,$$

$$\forall k, \forall t\,, \tag{33}$$

$$\text{Constraints } (13) - (18)\,,$$

$$xb_{mjt} \geq 0\,, \quad \forall m\,,\ \forall j\,,\ \forall t\,, \tag{34}$$

$$\text{Constraints } (19) - (24)\,,$$

where the new constraints are as follows:

- constraint (27) defines the total backlogging cost;
- constraints (28)-(29) are the balance equations, modified to take into account that the actual request for product option j in any time unit t is given by both its demand in that time unit and the backorders of the previous time unit. Then, backorders can be placed for the current time unit also;
- substitution inequalities (30) and (32) have been modified (with respect to Section 3) to take into account the backorders of the previous time unit and the backorders of the previous substitution levels;
- constraint (31) (similar to constraint 30) and constraint (33) (similar to constraint 32) represent the backorder inequalities. For any level of substitution, the amount of product option k that is backordered is less than or equal to a certain proportion of the unsatisfied demand of product k. Similarly as for the substitution inequalities, this proportion depends on the substitution rates and the backorder rate of that product. Also the backorder inequalities are written for each of the M levels of substitution. Once again, only the first two backorder inequalities are provided;
- constraint (34) defines xb_{mjt} as a non-negative variable.

The objective function is still in the form of cost minimization where the total backlogging cost (TCB) replaces the total lost-sale cost in the model of Section 3.

Similarly as in Section 3, by solving this MILP model, we obtain the optimal schedule of the production orders and the *short-term* product assortment over the planning horizon. In particular, note that Eq. (25) remains valid here.

5. Numerical examples

The mixed-integer linear programming models proposed in Section 3 and Section 4 for the production scheduling problem under product substitution (with lost sales and with backorders) are here applied to a set of random instances, with the objective of evaluating the models' behaviour in a wide variety of operative scenarios.

The set of random instances was generated similarly as in the experimental analysis proposed by Yücel et al. (2009) for the product assortment problem in the retail sector.

In particular, we discretized the planning horizon into 10 time units. The *long-term* product assortment is composed of 5 different product options that can be produced on 4 capacitated stations. We considered 3 levels of substitution, that is a reasonable value since substitution rates tend to become very small at higher levels.

Then, we adopted the same functions used in Yücel et al. (2009) to define the substitution costs. Specifically, the substitution costs for each product option j are assumed to be linear functions of the substitution level, m, and the margin of the very product, mg_j. Thus,

$$c_{mj}^{\text{sub}} = \theta \cdot m \cdot mg_j \, , \tag{35}$$

where $\theta > 0$ is a parameter to be set by the decision maker depending on the category under consideration and the customer expectations.

We introduced also the lost-sale or backorder costs as:

$$c_{mj}^{\text{lost}} / c_{mj}^{\text{back}} = \beta \cdot mg_j \, , \tag{36}$$

where $\beta = \max\{3\theta; 1\}$ in order to guarantee that the penalty for lost sales or backorders is at least equal to that of the last (e.g., third) substitution level.

The remaining model parameters are reported in Table 1, where those parameters that we have in common with Yücel et al. (2009) (i.e., w_{jk}, d_j, c^{setup}, c_j^{h}, mg_j) have been generated according to the same distributions. The capacity of each station C_s has been generated by the same distribution that Yücel et al. (2009) used for the order quantity in an application to the retail sector.

[PUT HERE TABLE 1]

The experimental analysis has been carried out by generating 100 random data sets according to the provided distributions. The 100 data sets have been tested by considering both the operative situations investigated in this paper (i.e. lost sales and backorders) and by varying the parameter θ (and, consequently β), in order to obtain different scenarios with increasing substitution costs.

The average of the investigated values over these 100 data sets are provided as test results. All runs were performed using the mixed-integer programming solver ILOG-CPLEX 12.6 on an Intel Core-i7 3.0 GHz PC.

5.1 *A single illustrative instance*

In the sequel, the proposed mathematical models are applied to one of the random instances from the data set generated in the present experimental analysis.

As explained above, we consider a planning horizon of 10 time units and a *long-term* product assortment of 5 different product options (called P1, P2, . . . , P5). The demand of each option per time unit is reported in Table 2.

[PUT HERE TABLE 2]

For each couple of options the matrix of substitution rates is given in Table 3, along with the lost-sale or backorder rates.

[PUT HERE TABLE 3]

The capacities, expressed in number of units, of the four production resources (denoted as R1, R2, R3 and R4) are as in Table 4, while the remaining input data, generated as described above, are not reported due to space limitations.

[PUT HERE TABLE 4]

For both the analysed operative situations (lost sales and backorders), the solution can be discussed in terms of production schedule and *short-term* product assortment.

5.1.1 *Lost sales*

By applying the MILP model of Section 3, we obtain the optimal production schedule depicted in Figure 1 case a), where it can be seen that products properly rotate over the planning horizon.

[PUT HERE FIGURE 1]

For a brief discussion about the *short-term* product assortment, we can focus on a particular time unit, e.g., time unit $t = 4$ (when a production order started in $t = 2$ is completed).

As explained in more detail below, the *short-term* product assortment in $t = 4$ is composed of only 3 product options (i.e., P1, P4, P5) out of the five constituting the *long-term* product assortment. In particular, Table 5 reports the product options on the rows while on the columns we have the demand in $t = 4$ (taken from Table 2 and repeated here for the sake of clearness) and the optimal values of the following variables:

- The inventory level at the beginning of time unit $t = 4$ (i.e., the inventory level at the end of the previous time unit, $I_{P_i\,3}$).
- The quantity that becomes available in that time unit from all the production resources (depending on the scheduling of production orders), where $X_{P_i\,4} = \sum_{s=1}^{4} x_{P_i, R_s, t=4}$ according to the notation of Section 3. In Figure 1 (case a) we can see that a production order of option P4, launched at time $t = 2$ on R2, completes in the time unit of interest, while there is neither inventory nor production for options P2 and P3 that, consequently, do not belong to the *short-term* product assortment in $t = 4$.
- The quantity of other options that substitutes for each P_i in $t = 4$, where

$$Xs^{\text{from}}_{P_i\,4} = \sum_{m=1}^{3} \sum_{\substack{j=1 \\ j \neq i}}^{5} xs_{m, P_j, P_i, t=4} \,,$$

  according to the notation of Section 3. For example, the demand of P2 (250 units) is partially satisfied (207 units) by the other substitutive options, according to the substitution rate matrix, while the remaining unsatisfied demand ($250 - 207 = 43$ units) are lost sales. A similar reasoning can be applied to option P3.
- The quantity of each option P_i used to substitute other demands in $t = 4$, where

$$Xs^{\text{to}}_{P_i\,4} = \sum_{m=1}^{3} \sum_{\substack{j=1 \\ j \neq i}}^{5} xs_{m, P_i, P_j, t=4} \,,$$

  according to the notation of Section 3. In this case, there are positive values for P1, P4 and P5 that substitute for P2 and P3.
- The inventory level at the end of the time unit, $I_{P_i\,4}$.
- The lost sales in $t = 4$, where $Xl_{P_i\,4} = \sum_{m=1}^{3} xl_{m, P_i, t=4}$. In this case, we have a certain amount of lost sales only for options P2 and P3.

Finally, it can be noted that the flow-balance relations are satisfied for each product options.

[PUT HERE TABLE 5]

Similar solutions are obtained for each time unit of the planning horizon demonstrating the ability of the proposed model to take into consideration both inventory and product assortment issues in case of demand substitution.

5.1.2 Backorders

The model of Section 4 is suitable for the operative situation where backorders are allowed.

In this case, the optimal production schedule is depicted in Figure 1, case b).

Similarly as in Section 5.1.1, we can discuss the *short-term* product assortment in a specific time unit. By selecting $t = 4$ again, it is composed of four out of the five options (P1, P3, P4, P5) and, specifically, Table 6 reports the values of the variable of interest in the optimal solution. The columns of Table 6 are the same as in Table 5, except for the second and the last columns that refer to the amount of backorders, where $Xb_{P_i\,t} = \sum_{m=1}^{3} xb_{m,\,P_i,\,t}$ for each option P_i.

By reasoning similarly as in Section 5.1.1 for lost sales, we can here limit our explanation to one of the product options. For example, let us consider option P3 which shows the most interesting behaviour. Since there are no backorders from the previous time unit ($Xb_{P_3\,3} = 0$), the quantity of P3 requested in $t = 4$ is equal to the demand in that time unit (6 700 units). The initial inventory (2 940 units) is used in large part to satisfy its own demand (for 2 396 units), but a portion (544 units) is also used to substitute for P2. Thus, the amount of demand of P3 that is not satisfied by the very option is 4 304 units (6 700 - 2 396). A significant portion is satisfied by the substitutive options (3 812 units) and the remaining part is a backorder (492 = 4 304 - 3 812).

[PUT HERE TABLE 6]

5.2 *Experimental results*

Table 7 reports the costs that have been minimized by applying both the models proposed in this paper, with lost sales and with backorders. These values are obtained by averaging the results from the 100 random data sets discussed above over different scenarios which differ from each other by the operative situation (i.e., lost sales or backorders) and by the level of the substitution and lost-sale/backorder costs. Both in the case where backorders are allowed and in the case with lost sales, the parameters θ and β have been varied in order to understand the behaviour as the substitution costs, together with the backorder/lost-sale costs, increase. Note that, for a given value of θ, the value of β is the same for both backorders and lost sales, meaning that their weight is the same in the two situations.

[PUT HERE TABLE 7]

Table 7 shows that scenarios with backorders generally involve lower setup costs but higher inventory and substitution costs. This is because the unsatisfied demand of any time unit is not lost and must be eventually fulfilled (and, consequently, the material flows along the planning horizon are higher than in the scenarios with lost sales).

Then, it can be noted that as substitution costs increase, the optimal solution extends the *short-term* product assortment. This result is inferred by observing the increase of both the total setup costs and the total inventory costs. In particular, if we consider the sum of setup and inventory costs we have:

- Backorders:
  - if $\theta = 0$, C(Setup+Inventory)= 469 935 + 731 362 = 1 201 296
  - if $\theta = 1$, C(Setup+Inventory)= 478 113 + 881 530 = 1 359 643
- Lost Sales:
  - if $\theta = 0$, C(Setup+Inventory)= 470 369 + 726 819 = 1 197 189

- if $\theta = 1$, C(Setup+Inventory)= 479 287 + 866 253 = 1 354 540

Thus, the sum of setup and inventory costs increases of more than 10% as θ passes from 0 to 1 in both situations, backorders and lost sales.

Moreover, we can see that the extension of the *short-term* product assortment can even lead to a reduction of the amount of backorders, whose costs reduces from 21 208 € to 18 173 € in our experimental analysis.

In order to better understand how demand substitution can help reducing backorders and lost sales (and, consequently, increasing customer satisfaction) it is now interesting to reason not only in terms of costs, but in terms of satisfied demand and amount of backorders or lost sales.

Table 8 shows, for both the considered situations and increasing substitution and lost-sale/backorder costs, the following information:

- Scenarios where substitution is allowed (and, consequently the proposed models can be applied in the form of Section 3 and Section 4):
  - %ds is the first-choice demand satisfied over the planning horizon, i.e., the percentage of the product demand that has been satisfied by offering to the customers that very option;
  - %sub is the percentage of the product demand that has been satisfied with substitutive options (by satisfying, even if partially, the customers which are willing to substitute);
  - %ls/%bk is the percentage of backorders or lost sales;
- Scenarios where substitution is not allowed (and, consequently a simpler version of the models where only the balance equations are retained can be applied): %ds and %ls/%bk have the same meaning as above, but the factor %sub is no more applicable (since the demand of a certain product option can be satisfied by that very option only)
- Difference Δ of backorders/lost sales when substitution is allowed and when it is not allowed (negative values mean a reduction of the percentage of backorders or lost sales when substitution is allowed).

As expected, as the substitution and lost-sale/backorder costs increase the percentage of first-choice demand satisfied increases as well, while backorders and lost sales decrease. But the most interesting result is that allowing substitution always helps in reducing backorders or lost sales. Such a reduction can even be of about 30% when the substitution costs are low in our experimental analysis.

Hence we can conclude that it is important to study the customers' buying behaviour and understand if they are willing to substitute between different product options. If so, it is possible to obtain significant benefits in terms of reduction of backorders or lost sales.

[PUT HERE TABLE 8]

6. Conclusions

In this paper two mixed-integer linear programming models for the production scheduling under customer-driven demand substitution are proposed. In the first model, if a product is not available in that time unit, whether the customer substitutes the product with another one or a lost sale occurs. In the second model backorders are allowed.

The proposed models could support the production planning especially when it is not economically or technically suitable to have all the product options of the *long-term* product assortment available at the same time. The subset of products that can be offered to the customers at a time is here called *short-term* product assortment.

An extensive experimental analysis has underlined the importance of including demand substitution issues in the production planning. In particular, we have found that when the customers are willing to substitute between different product options it is possible to obtain significant benefits in terms of reduction of backorders or lost sales.

Further possible extensions of the models are related to the possibility of considering different product categories (i.e., different sets of substitutable product options). The introduction of additional production stages could also lead to more realistic models.

References

Bassok, Y., R. Anupindi, and R. Akella (1999). Single-period multiproduct inventory models with substitution. *Operations Research 47*(4), 632–642.

Bitran, G. R. and S. Dasu (1992). Ordering policies in an environment of stochastic yields and substitutable demands. *Operations Research 40*(5), 999–1017.

Drezner, Z., H. Gurnani, and B. A. Pasternack (1995). An EOQ model with substitutions between products. *Journal of the Operational Research Society 46*(7), 887–891.

Gallego, G., K. Katircioglu, and B. Ramachandran (2006). Semiconductor inventory management with multiple grade parts and downgrading. *Production Planning & Control 17*(7), 689–700.

Gebennini, E., L. Zeppetella, A. Grassi, and B. Rimini (2015). Production scheduling to optimize the product assortment in case of constrained capacity and customer-driven substitution. *IFAC-PapersOnLine 48*(3), 1954–1959.

Hsu, V. N., C.-L. Li, and W.-Q. Xiao (2005). Dynamic lot size problems with one-way product substitution. *IIE Transactions 37*(3), 201–215.

Jans, R. and Z. Degraeve (2008). Modeling industrial lot sizing problems: a review. *International Journal of Production Research 46*(6), 1619–1643.

Kahn, E. B. (1998). Dynamic relationships with customers: High-variety strategies. *Journal of the Academy of Marketing Science 26*(1), 45–53.

Kök, A. G. and M. L. Fisher (2007). Demand estimation and assortment optimization under substitution: Methodology and application. *Operations Research 55*(6), 1001–1021.

Mantrala, M. K., M. Levy, B. E. Kahn, E. J. Fox, P. Gaidarev, B. Dankworth, and D. Shah (2009). Why is assortment planning so difficult for retailers? A framework and research agenda. *Journal of Retailing 85*(1), 71–83.

Netessine, S. and N. Rudi (2003). Centralized and competitive inventory models with demand substitution. *Operations Research 51*(2), 329–335.

Rao, U. S., J. M. Swaminathan, and J. Zhang (2004). Multi-product inventory planning with downward substitution, stochastic demand and setup costs. *IIE Transactions 36*(1), 59–71.

Smith, S. A. and N. Agrawal (2000). Management of multi-item retail inventory systems with demand substitution. *Operations Research 48*(1), 50–64.

Stavrulaki, E. (2011). Inventory decisions for substitutable products with stock-dependent demand. *International Journal of Production Economics 129*(1), 65–78.

Tripathy, A., H. Süral, and Y. Gerchak (1999). Multidimensional assortment problem with an application. *Networks 33*(3), 239–245.

Yücel, E., F. Karaesmen, F. S. Salman, and M. Türkay (2009). Optimizing product assortment under customer-driven demand substitution. *European Journal of Operational Research 199*(3), 759–768.

List of Figures

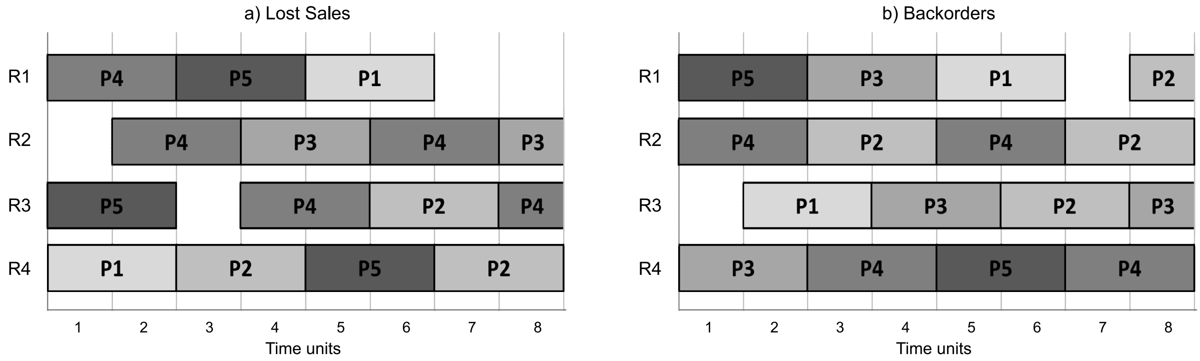

Figure 1. Optimal production schedule.

List of Tables

Table 1. Model parameters.

Parameter	Distribution
w_{jk}	Uniform distribution, where $\sum_{k=1}^{J} w_{jk} = 1$ and $0 \leq w_{jk} \leq 1$, $\forall j, k$
$w_j^l\ w_j^b$	$w_j^l = w_j^b = 1 - \sum_{k=1}^{J} w_{jk} = 1$, $\forall j$
d_{jt}	$d_{jt} = \alpha_{jt} \cdot 10\,000$, where α_{jt} has uniform distribution, with $0 \leq \alpha_{jt} \leq 1$, $\forall j$
C_s	Uniform distribution, where $4\,000 \leq C_s \leq 34\,000$, $\forall s$
c^{setup}	Uniform distribution, where $30\,000 \leq c^{\text{setup}} \leq 50\,000$ (it is the same for all stations in a given instance)
c_j^{h}	$c_j^{\text{h}} = h_j \cdot mg_j$, where h_j has an uniform distribution, with $0.3 \leq h_j \leq 1$, and mg_j is the margin of product j, $\forall j$
mg_j	Normal distribution with mean 6 and deviation 2, $\forall j$
LT_j	2 time units, $\forall j$
T^r	6 time units
I_j^0	Initial inventory able to cover the demand of each product for 2 time units (e.g., the production lead time). So, $I_j^0 = \sum_{t=1}^{2} = d_{jt}$

Table 2. Demand per time unit [units].

	1	2	3	4	5	6	7	8	9	10
P1	8300	9700	3600	8200	5700	7400	2900	6200	3300	200
P2	3100	7400	5600	250	9000	3600	8000	9800	17900	9500
P3	7200	1700	12100	6700	12100	15000	7300	1100	2800	13520
P4	6300	2700	7700	12700	4500	5500	4100	12600	7000	8100
P5	7000	5300	300	3900	3800	4900	12000	1500	1000	3400

Table 3. Substitution rates matrix.

	P1	P2	P3	P4	P5	lost-sale/backorder
P1	-	0.17	0.31	0.28	0.15	0.09
P2	0.37	-	0.15	0.15	0.2	0.13
P3	0.25	0.45	-	0	0.18	0.12
P4	0.32	0.21	0.09	-	0.23	0.15
P5	0.11	0.13	0.48	0.18	-	0.1

Table 4. Capacities of the production resources.

	R1	R2	R3	R4
Capacity [units]	9947	22314	15491	19071

Table 5. Optimal solution for $t = 4$.

	$d_{P_i\,4}$	$I_{P_i\,3}$	$X_{P_i\,4}$	$x0_{P_i\,4}$	$Xs^{\text{from}}_{P_i\,4}$	$Xs^{\text{to}}_{P_i\,4}$	$I_{P_i\,4}$	$Xl_{P_i\,4}$
P1	82 00	10 438	0	8 200	0	1 430	808	0
P2	250	0	0	0	207	0	0	43
P3	6 700	0	0	0	5 778	0	0	922
P4	12 700	0	22 314	12 700	0	1 081	8 533	0
P5	3 900	7 375	0	3 900	0	3 475	0	0

Table 6. Optimal solution for $t = 4$.

	$d_{P_i\,4}$	$Xb_{P_i\,3}$	$I_{P_i\,3}$	$X_{P_i\,4}$	$x0_{P_i\,4}$	$Xs^{\text{from}}_{P_i\,4}$	$Xs^{\text{to}}_{P_i\,4}$	$I_{P_i\,4}$	$Xb_{P_i\,4}$
P1	8 200	478	0	15 491	8 678	0	1 841	4 972	0
P2	250	846	0	0	0	1 086	0	0	10
P3	6 700	0	2940	0	2 396	3 812	544	0	492
P4	12 700	0	12 715	0	12 700	0	15	0	0
P5	3 900	0	7 700	0	3 900	0	2 498	1 302	0

Table 7. Optimal solutions (model parameters generated according to Table 1).

			Costs [€]					
	Teta	Beta	Setup	Inventory	Substitution	Lost-sale	Backorder	Total
Backorders	0	1	469 935	731 362	-	-	21 208	1 222 504
	0.2	1	466 798	750 940	97 186	-	18 921	1 333 844
	0.4	1.2	469 297	772 613	163 855	-	20 297	1 426 061
	0.6	1.8	473 424	811 635	202 079	-	19 860	1 506 999
	0.8	2.4	474 175	845 530	235 767	-	19 290	1 574 762
	1	3	478 113	881 530	256 002	-	18 173	1 633 818
Lost Sales	0	1	470 369	726 819	-	21 991	-	1 219 180
	0.2	1	468 162	741 633	95 831	21 900	-	1 327 527
	0.4	1.2	470 929	764 935	157 434	25 196	-	1 418 494
	0.6	1.8	475 208	801 185	198 346	26 534	-	1 501 273
	0.8	2.4	476 286	835 319	230 500	29 041	-	1 571 146
	1	3	479 287	866 253	258 927	28 189	-	1 632 656

Table 8. Comparison with the case where substitution is not allowed.

			with substitution			no substitution		Delta
	Teta	Beta	%ds	%subs	%ls /%bk	%ds	%ls /%bk	%ls/%bk
Backorders	0	1	68.04%	31.96%	1.72%	100.00%	29.89%	-28.17%
	0.2	1	73.36%	26.64%	1.52%	100.00%	29.89%	-28.36%
	0.4	1.2	76.66%	23.34%	1.36%	100.00%	24.70%	-23.34%
	0.6	1.8	80.22%	19.78%	0.91%	100.00%	16.24%	-15.34%
	0.8	2.4	82.19%	17.81%	0.66%	100.00%	13.13%	-12.47%
	1	3	84.36%	15.64%	0.52%	100.00%	11.24%	-10.72%
Lost Sales	0	1	67.54%	30.72%	1.74%	66.21%	33.79%	-32.05%
	0.2	1	72.41%	25.85%	1.74%	66.21%	33.79%	-32.05%
	0.4	1.2	76.03%	22.31%	1.66%	72.33%	27.67%	-26.01%
	0.6	1.8	79.50%	19.29%	1.21%	83.80%	16.20%	-14.99%
	0.8	2.4	81.48%	17.51%	1.00%	88.28%	11.72%	-10.72%
	1	3	83.37%	15.84%	0.79%	91.46%	8.54%	-7.75%