Multi-product Multi-route Production Planning Simultaneously Considering Partial Backlogs and Lost Sales

Shan Lu, Hongye Su, Yue Wang, Lei Xie, Li Zhu

Abstract—A production planning approach is addressed in this paper for multi-period, multi-product and multi-route production environment considering backlogs and lost sales. When the production capacity or order demand fluctuates, it could lead to a low utilization of the production resource if considering the production system as a whole. Hence, each production stage and inventory is depicted separately, for which fluctuations can be alleviated by utilizing the spare capacity or even overtime in advance. Order demand can be fulfilled from end item inventory, however, if insufficient, both backlogs and lost sales are allowed. Trade-off between these two scenarios is interposed by a policy that lost sales occur when orders are backlogged over a time-phased threshold. The proposed planning problem can be exactly formulated by a mixed-integer nonlinear programming (MINLP) model that minimizes the corresponding cost, and then reformulated by a mixed-integer linear programming model (MILP) using linearization techniques. Consequently, a case study oriented from a steel rolling mill plant is introduced to illustrate the effectiveness of the proposed approaches.

I. INTRODUCTION

Determining appropriate production quantities (lot sizes) and schedules for a product manufactured from raw materials to the end-item contributes to smooth and cost-efficient operations of a plant. However, the customers’ desire for an increased variety of product variants with reliable due dates challenges in manufacturing industries. Specifically, for the industries in which production stages constitute a complex flow line network, e.g. in the steel rolling mill plant, the number of steel grades, alloy elements, dimensions and physical properties, as well as the number of production routes for different products has relatively increased, thus complicating the efforts to make decisions on lot sizes of each stage. Meanwhile, production capacity is not always constant due to the production environment, and demand orders may vary with market circumstances over time. These fluctuations would lead to an increase in backlogging or inventory levels, further resulting in delay in delivery or capacity waste. To handle these problems, decision makers need to consider all the interactions and constraints between adjacent production stages, such as production capacity, processing time, lead time, availability of raw materials, fulfillment, and so on. Besides, they should recognize the value that delivery flexibility provides in trade-off among the order requirements.

The proposed production planning problem can be considered as an extension and combination of dynamic lot-sizing problems (DLSP). Many researchers have reviewed DLSP since it’s a strong and useful tool in the field of production planning and inventory management [1], of which most literatures are focused on single-level capacitated lot-sizing problem (SLCLP) [2]. Specifically, the multi-stage production environment addressed can be related to multi-level capacitated lot sizing problem (MLCLSP) [3]. In multi-level production systems, there could be a certain time-phased or material relationship that mutually links the adjacent production stages [4]. Tempelmeier and Helber [5] developed a novel formulation for multi-item MLCLSP for general product structures under multiple capacity constraints. Afterwards, based on the previews model, Tempelmeier and Buschkühl [6] incorporated linked lotsizes that provides a good compromise between the big-bucket and small-bucket model. Lu and Su [7] addressed multi-product multi-stage flow line production systems in a make-to-order steel production environment. It should be noted, however, that the existing related work paid not much attention to handling the fluctuation on production system and demand. Some literatures considered the demand or capacity uncertainty using stochastic programming, robust optimization and fuzzy sets [8]. Distinctively, Mahdavi et al. [9] highlighted the insufficient utilization of available capacity when demand fluctuates and presented a novel model considering individual production routes, but they ignore the lead time and lost sale that were practical in industrial application.

This work is originally motivated by multi-route manufacturing characteristics and one form of customer response mechanism available in a steel rolling mill. A large number of end items are to be produced through a series of production stages in different routes depending on product specifications. Each stage of the whole production process can be regarded as a single planning unit that is dependent on its predecessor and successor stages, fundamentally ‘pulled’ by the order demand. The consideration of dividing the production planning problem into several single planning units rather than regarding the whole system as a time-phased bottleneck is reasonable. When the production capacity or order demand fluctuation occurs, non-bottleneck stages are able to utilize the spare capacity. We also incorporate lost sales to address limitation of backlogging level. Consequently, the decision maker must consider the trade-off between the cost of backlogs and lost sales.
Given such challenges, we develop a large-scale optimization model based on an exact MINLP problem and the approximate equivalent MILP problem, to address the above practical consideration in production planning. The objective is to find the optimal decision variables during planning horizon. The main characteristics of the proposed model are that it depicts the planning problem of the whole system as several interactional planning units so that each production stage can be tracked and managed by respective strategies to alleviate the fluctuation from orders; in the meanwhile, it also considers backlogs and partial lost sales simultaneously applied in a practical scenario in which to decision maker can exploit leeway to absorb the production capacity. As the proposed optimization model is originally formulated by a MINLP problem which is difficult to solve to optimality, we reformulate the nonlinear constraints through additional variables and boundaries, thus transforming the original MINLP problem into a MILP problem. Finally, the proposed approach is applied for a case study of steel rolling mill to validate the effectiveness.

II. PROBLEM STATEMENT AND MODEL DEFINITION

A. Problem Statement

For some large industries producing various products, the production system usually consists of a series production stages with different functionality. An end item could be produced from its raw materials through some given production stages that constitute a specific production route. When considering the production routes for multiple products, the system will construct a production network with starting and terminal points. The confirmed and forecast orders with delivery date are accumulated in advance and sent to the terminal of the production network at the beginning of each period. Since each stage differs in its dynamics and behaviors, it is more realistic to model them separately and then integrate them as a whole.

In a problem where production capacity and order demands fluctuate, the decision maker needs to take steps to reduce the fluctuation. If the production capacity of a certain stage is insufficient and assuming limited by planned maintenance policies during a future period, this limitation could be alleviated through making using of spare capacity or even overtime before period, to produce and store the items in advance. Then the successor stages do not need to follow its slowdown in production pace, further enhancing production utilization of the whole system. Although this time-phased strategy helps in fulfilling the demands on time, its effect is still limited when additional order demands or unexpected maintenance occurs so that the orders cannot be satisfied before delivery date. In this case, pending parts of the orders are backlogged or lost following the policies. However, the decision maker should weigh when to fulfil the pending parts or whether they are partially lost. We consider a practical scenario that the customers can accept a limited period for waiting the fulfillment after the delivery date. If the pending parts are not fulfilled within the ‘time tolerance’, they are lost, referred to as partial lost sales. Fig. 1 presents a scenario of which backlogs and lost sales occur simultaneously within a planning horizon. The policy allows the end item to be backlogged for a certain period but not over a threshold which is \( t + 3 \) in the figure. Backlogging horizon exceeding the threshold would trigger the scenario to lost sales.

![Figure 1. An illustration of considering backlogs and lost sales](image)

With the above brief analysis, the production planning problem addressed can be described as follows: given the order demands and relevant production parameters over the planning horizon, the planning procedure needs to obtain the optimal production lot sizes and schedule of each stage, inventory level, material schedule between adjacent stages, as well as the backlogging and lost sale level of the end items, so that the corresponding cost is minimized.

B. Production Network

To formulate the mathematical model, we use the following notations. It should be noted that material quantities are measured in tons, and time within a period is stated in hours.

Indices
- \( j \) index for production stages
- \( j' \) index for immediate successor production stages of \( j \)
- \( i \) index for items
- \( i' \) index for the immediate downstream items of \( i \)
- \( t \) index for time period
- \( t' \) index for the periods after \( t \)
- \( t^* \) index for the periods before \( t \)

Sets
- \( N \) set for all the items
- \( J \) set for all the production stages
- \( T \) set for time horizon
- \( B \) subset of \( N \) denoting raw materials
- \( W \) subset of \( N \) denoting work-in-process (WIP)
- \( P \) subset of \( N \) denoting end items
- \( C(i) \) set for the production stages that produce \( i \)
- \( O(i) \) set for the first production stage of \( i \)
M(j) set for the items produced by production stage \( j \)
\( \Pi(i,j) \) set for the items produced from \( i \) by stage \( j \)

**Parameters**
- \( \mu \) threshold of time horizon for backlogging
- \( \eta_i j \) production efficiency of stage \( j \) for producing \( i \)
- \( A_{t_k} \) hours in period \( t \)
- \( c_b i \) cost of supplying one unit of raw material \( i \)
- \( c_{b h} i \) cost of backlogging one unit of end item \( i \) for a period
- \( c_{w i} \) cost of holding one unit of WIP \( i \) for a period
- \( c_{p i} \) cost of holding one unit of end item \( i \) for a period
- \( c_l i \) cost of losing one unit of end item (order) \( i \)
- \( c_o i \) overtime cost per hour in period \( t \)
- \( c_{p i} \) cost of producing one unit of \( i \) at stage \( j \)
- \( c_s j \) setup cost of stage \( j \) for producing \( i \)
- \( H_{r,i}^{m a x} \) maximum inventory level of raw material warehouse
- \( L_{r,i}^{m a x} \) maximum inventory level of end item warehouse
- \( I_{r,i}^{m a x} \) maximum inventory level of WIP warehouse
- \( l_i \) lead time of supplying raw material \( i \in B \)
- \( l_j i \) lead time of producing at stage \( j \)
- \( m_{b,i} \) fixed supply batch size of raw material \( i \in B \)
- \( m_{b,j} i \) fixed production batch size of item \( i \in W \cup P \)
- \( m_{d} i \) demand for end item \( i \in P \) in period \( t \)
- \( m_{t} i \) regular production capacity of stage \( j \) (in hours)
- \( m_{t m} i \) maintenance time of stage \( j \) in period \( t \) (in hours)

**Variables**
- \( m_{b a,i,j} i \) backlogging quantity of end item \( i \in P \) in period \( t \) that will be fulfilled in period \( t' \)
- \( m_{t} i \) quantity of end item \( i \in P \) that fails to meet the order
- \( m_{p} i \) supply quantity of raw material \( i \in B \) in period \( t \)
- \( m_{x} i \) production quantity of item \( i \in W \cup P \) at stage \( j \)
- \( m_{o} i \) overtime production capacity at stage \( j \) in period \( t \)
- \( u_{b,i,j} i \) quantity of raw material \( i \in B \)
- \( u_{p,i,j} i \) quantity of end item \( i \in P \) shipped to the warehouse in period \( t' \) and will be stored to be used in period \( t' \)
- \( u_{w,i,j} i \) quantity of WIP \( i \in W \) shipped to the warehouse in period \( t' \) and will be stored to be used in period \( t' \)
- \( n_{b,i} i \) number of supply batch of raw material \( i \in B \)
- \( n_{b,i} i \) number of production batch of item \( i \in W \cup P \)
- \( x_{i} i \) binary variable that indicates whether item \( i \in W \cup P \) is produced at stage \( j \) in period \( t \)

(a) Material balance of raw material warehouse

\[
\sum_{i=t+1}^{T} \sum_{j=1}^{J} u_{w,i,j} = \sum_{i=t+1}^{T} \sum_{j=1}^{J} u_{b,i,j} + \sum_{i=1}^{m} u_{p,i,j} - \sum_{i=1}^{m} u_{b,i,j} \quad \forall i \in B, \forall t
\]

(b) Material balance of work-in-process warehouse

\[
\sum_{i=t+1}^{T} \sum_{j=1}^{J} u_{w,i,j} = \sum_{i=t+1}^{T} \sum_{j=1}^{J} u_{b,i,j} + \sum_{i=1}^{m} u_{p,i,j} - \sum_{i=1}^{m} u_{b,i,j} \quad \forall i \in W \cup P, \forall t
\]

(c) Inventory level constraints

\[
\sum_{i=t+1}^{T} \sum_{j=1}^{J} u_{w,i,j} \leq I_{r,i}^{m a x} \quad \forall t
\]

\[
\sum_{i=t+1}^{T} \sum_{j=1}^{J} u_{w,i,j} \leq L_{r,i}^{m a x} \quad \forall t
\]
Constraint (7) ensures that the accumulated inventory level of raw materials should not exceed the maximum inventory level. The same meaning can be given for (8) and (9), restricting the inventory level of WIP and end items.

(e) Production capacity constraints

\[
\sum_{i \in M(j)} \left( \frac{m x_{g_i}}{n_i} + t s_y \cdot x_{g_i} \right) \leq t a_{i} + t o_{i} \quad \forall i, \forall t
\]

(10)

\[
t a_{i} + t o_{i} \leq A u_{i} - t m_{o} \quad \forall j, \forall t
\]

(11)

\[
m x_{g_{i}} > 0 \Leftrightarrow x_{g_{i}} = 1 \quad \forall i \in W \cup P, \forall j, \forall t
\]

(12)

\[
m x_{g_{i}} = 0 \Leftrightarrow x_{g_{i}} = 0 \quad \forall i \in W \cup P, \forall j, \forall t
\]

(13)

Constraints (10) and (11) makes sure that the total utilized capacity consisting of processing and setup should not exceed the available capacity. Considering the setup mentioned is a binary variable, the relation between decision on setup and production quantity can be stated by (12) and (13).

(f) Batch production constraints

\[
m p_{a_{i}} = n b_{a_{i}} \cdot m b_{j} \quad \forall i \in B, \forall t
\]

(14)

\[
m x_{g_{i}} = n b_{g_{i}} \cdot m b_{j} \quad \forall i \in M(j), \forall j, \forall t
\]

(15)

The supply of raw materials takes place in batches of a fixed size, so the supply and production quantity can only be the integral multiple of batch size as shown in (14) and (15).

(g) Backlogging constraints

\[
\sum_{j \in R} w_{p_{i}, j} \cdot \sum_{j \in J} m b_{a_{i}, j} = 0 \quad \forall i \in P, \forall t
\]

(16)

For end items \(i \in P\), the policy allows them to be fulfilled in the future periods, as well as stored to be used in the future periods. However, these two situations cannot occur simultaneously, which is specified by (16). It is similar to the best I can do policy [10], that is, the decision maker should make the most of inventory to satisfy the order demands before considering backlogging.

(h) Lost sales constraints

\[
\Delta t = t' - t - \mu
\]

(17)

\[
\Delta t > 0 \Leftrightarrow m b_{a_{i}, j} = 0 \quad \forall i \in P, \forall t, \forall t' > t
\]

(18)

\[
\Delta t \leq 0 \Leftrightarrow m b_{a_{i}, j} \geq 0 \quad \forall i \in P, \forall t, \forall t' > t
\]

(19)

The order demand cannot be backlogged for indefinite time periods since the customers could have an agreement on the maximum time limitation of backlogging. Assume that \(\mu\) is the time threshold of backlogging from the current period \(t\), then \(t + \mu\) means the deadline for backlogging. We use a variable \(\Delta t\) to indicate the difference between assumptive backlogging time horizon and the deadline, which is calculated by (17). As given by (18) and (19), in the case that the assumptive backlogging time horizon exceeds the deadline, the backlogging level will be cleared to zero, and lost sale level will be available and obtained by (5).

(i) Decision variable constraints

\[
x_{g_{i}} \in \{0,1\} \quad \forall i \in W \cup P, \forall j, \forall t
\]

(20)

\[
b_{a_{i}} \in N \quad \forall i \in B, \forall t
\]

(21)

\[
b_{g_{i}} \in N \quad \forall i \in M(j), \forall j, \forall t
\]

(22)

Constraints (20) ~ (22) represent the binary and integral restrictions on the respective decision variables.

C. Objective Function

The objective function evaluated by the total cost that is composed of production, raw material, inventory holding, setup, overtime, backlogging and lost sale. The objective of the problem is to minimize the total cost during planning horizon \(T\). Consequently, the objective function of the production network can be given by (23).

\[
\text{Min } = \text{CH} + \text{CS} + \text{CO} + \text{CR} + \text{CI} + \text{CB} + \text{CL}
\]

(23)

\[
\text{CH} = \sum_{i \in M(j)} \sum_{j \in J} \sum_{t=1}^{T} m x_{g_{i}} \cdot c p_{g_{i}}
\]

(24)

\[
\text{CS} = \sum_{i \in P} \sum_{j \in C} \sum_{t=1}^{T} x_{g_{i}} \cdot c s_{g_{i}}
\]

(25)

\[
\text{CO} = \sum_{i \in M(j)} \sum_{j \in J} \sum_{t=1}^{T} t o_{j} \cdot c o_{j}
\]

(26)

\[
\text{CR} = \sum_{i \in I} \sum_{j \in J} \sum_{t=1}^{T} m p_{a_{i}} \cdot c b_{j}
\]

(27)

\[
\text{CI} = \sum_{i \in I} \sum_{j \in J} \sum_{t=1}^{T} (t' - t) \cdot u b_{a_{i}, j} \cdot c i b_{j}
\]

(28)

\[
\text{CB} = \sum_{i \in I} \sum_{j \in J} \sum_{t=1}^{T} (t' - t) \cdot m b_{a_{i}, j} \cdot c b a_{j}
\]

(29)

\[
\text{CL} = \sum_{i \in I} \sum_{j \in J} \sum_{t=1}^{T} m l_{a_{i}} \cdot c l_{i}
\]

(30)

Equation (24) ~ (26) represents the cost of production, setup cost and overtime associated with WIP and end item production. Equation (27) represents the cost of raw materials that is either purchased from the supplier or produced from the upstream plant. Equation (28) is the inventory holding cost at
raw material, WIP and end item warehouse, which take into account the storage time period. The similar explanation is given for backlogging penalty cost by (29). Equation (30) represents lost sale penalty cost that is relatively higher than backlogging cost.

III. MODEL REFORMULATION

The overall mathematical model with the aforementioned constraints (1) ~ (22), and the objective function (23), comprise a non-convex MINLP problem. However, solving an MINLP problem directly will result in inconsistency in solution quality and time [11]. Therefore, we will examine the characteristic of the terms firstly to enhance the formulation.

Since setup binary variable corresponds to the production quantity which can be regarded as continuous variables, constraints (12) and (13) can be stated by (31), where $m_{i}^{\text{min}}$ and $m_{i}^{\text{max}}$ are the lower and upper boundaries of $m_{i}$.

$$x_{i} \cdot m_{i}^{\text{min}} \leq x_{i} \cdot m_{i}^{\text{max}} \forall i \in W \cup P, \forall j, \forall t$$

Constraint (16) includes a bilinear term to restrict the backlogging and storage. To linearize the constraint, we use a positive number $\varepsilon$ and binary variables $\eta_{i,j,t}$ and $\xi_{i,j,t}$. Then the constraint can be linearized by (32) ~ (34).

$$Q \cdot E_{i} \geq \sum_{j \in J} \sum_{t \in T} u_{i,j,t} \forall i \in P, \forall t$$

$$Q \cdot G_{i} \geq \sum_{j \in J} \sum_{t \in T} \eta_{i,j,t} \forall i \in P, \forall t$$

$$E_{i} + G_{i} \leq 1 \forall i \in P, \forall t$$

Constraints (18) and (19) describe the relationship between two continuous variables. Alternatively, we use a binary variable $z_{i,j,t}$ to indicate whether the assumptive backlogging time horizon exceeds the deadline. Thus, these two constraints can be reformulated by (35) ~ (38).

$$z_{i,j,t} = 1 \iff \Delta t > 0 \forall i \in P, \forall t, \forall t' > t$$

$$z_{i,j,t} = 0 \iff \Delta t \leq 0 \forall i \in P, \forall t, \forall t' > t$$

$$z_{i,j,t} = 1 \iff m_{i}^{\text{ba}}_{i,j,t} = 0 \forall i \in P, \forall t, \forall t' > t$$

$$z_{i,j,t} = 0 \iff m_{i}^{\text{ba}}_{i,j,t} > 0 \forall i \in P, \forall t, \forall t' > t$$

It should be noted that, although the above four constraints represent the relationship between binary and continuous variables, they differ with those in (12) and (13) in restriction region. Then constraints (35) ~ (38) can be further transformed into a linear form, which is stated by (39) and (40), where $\Delta t^{\text{min}}$ and $\Delta t^{\text{max}}$ are the lower and upper boundaries of $\Delta t$; $\varepsilon$ is a small positive constant; $m_{i}^{\text{ba}}_{i,j,t}$ and $m_{i}^{\text{ba}}_{i,j,t}$ are the lower and upper boundaries of $m_{i}^{\text{ba}}_{i,j,t}$.

$$(1 - z_{i,j,t}) \cdot (\Delta t^{\text{min}} - \varepsilon) < \Delta t \leq z_{i,j,t} \cdot \Delta t^{\text{max}} \forall i \in P, \forall t, \forall t' > t$$

$$(1 - z_{i,j,t}) m_{i}^{\text{ba}}_{i,j,t} \leq m_{i}^{\text{ba}}_{i,j,t} \leq (1 - z_{i,j,t}) m_{i}^{\text{ba}}_{i,j,t} \forall i \in P, \forall t, \forall t' > t$$

Through the above model reformulation, the original MINLP model can be transformed into a MILP model. The lower and upper boundaries of the variables mentioned above can be specified easily based on physical insights into a given system [12]. In the reformulated multi-period MILP model, the objective function is cost minimization given by (23), and the constraints are specified by (1) ~ (11), (14) ~ (15), (17), (20) ~ (22), (31), (32) ~ (34), (39) and (40).

IV. CASE STUDY

The plant we investigate manufactures small and mid-size steel plates in a make-to-order environment. In summary, six production stages are considered in the plant: rolling, cooling, shearing, cracking off, repairing and re-heating. Raw materials are steel slabs supplied by the upstream production or purchased from suppliers. Then hundreds of different end items are produced through some or all the production stages upon the specific production routes. Each kind of raw material can produce multiple products. In the case study, we identify 10 product archetypes produced by 4 classes of raw materials. Fig. 2 depicts the typical production process and Fig. 3 presents the topology of material flow within the production network. In Fig. 2, the production stages mentioned above are stated as stage A ~ F respectively and WIP refers to the work-in-process warehouse. In Fig. 3, R1 ~ R3 refers to the raw materials; A1 ~ A5, B1 ~ B5, C1 ~C5, D1 ~ D3, E1 and E6 are the WIP, E1 ~ E8 and F1 ~ F2 are the end items. Note that E1 and E6 can act as both work-in-process and end items.

![Figure 2. Production system for the case study](image-url)
The production planning model addressed for the multi-product multi-route production system is applied to this case study. The planning horizon for the system is composed of 7 periods, and each period is 1 day. The case study for the proposed MILP model was formulated and solved using branch-and-bound algorithm by LINGO 11. The numerical experiments were implemented on an Int2 2.5 GHz personal computer with 4 GB RAM.

We apply two cases to compare the effectiveness of the proposed model (case 1). In case 2, backlogs are not allowed. In case 3, lost sales are not allowed. These two scenarios can be modelled by eliminating the terms related to backlogs or lost sales from the objective function and constraints.

Table I presents the optimization results of the three cases. It indicates that a 0.07% and 0.12% reduction of the total cost in case 1 can be achieved in comparison to the cost in case 2 and case 3 respectively. In case 2, the lost sale only policy results in a much higher cost of lost sales but lower cost of production and raw materials compared with those in case 1. It is due to the lost sale penalty is higher than the cost related to producing a product, and backlogging policy allows to fulfil the backlogs within a limited period which is more economic efficient. In case 3, the backlogging level is much higher than that in case 1 due to the policy can only handle the unfulfilled orders through backlogging them or producing overtime. However, working overtime is not always expected in practical. Within the three cases, simultaneously allowing backlogs and lost sales (in case 1) achieves least overtime.

V. CONCLUSION

In this paper, a multi-product multi-route production planning problem is studied, where both backlogs and lost sales are considered. We propose an exact MINLP model to address the problem. Raw material, WIP and end item warehouse are depicted separately for formulation of each production stage. This formulation strategy allows decision maker to handle the fluctuations from production capacity and orders by making the most of spare capacity and inventory in advance. Further, when the order demand still cannot be fulfilled, they are allowed to be backlogged or lost. A practical scenario is addressed to specify that backlogs will be lost if the backlogging level exceeds a time-phased threshold. To develop reliable decision tool, the exact model is reformulated to a MILP model. A case study is applied for the proposed approaches. The experimental results show that the proposed model is able to reduce the cost by balancing the backlogs and lost sales.

REFERENCES


