International Journal of Production Research,
Vol. 45, No. 16, 15 August 2007, 3579–3597

Production smoothing in just-in-time manufacturing systems: a review of the models and solution approaches

MESUT YAVUZ*† and ELIF AKÇALI‡

†Research and Engineering Education Facility, University of Florida, FL, USA
‡Supply Chain and Logistics Engineering (SCALE) Center, Department of Industrial and Systems Engineering, University of Florida, FL, USA

(*Revision received January 2007*)

Production smoothing is one of the most important tactical planning activities for the efficient operation of mixed-product just-in-time (JIT) manufacturing systems. As a result, increased research attention has focused on this topic. However, a closer examination of the analytical literature reveals that the majority of the existing work concentrates on synchronized assembly line systems, which is, in part, due to the fact that JIT philosophy originated in an assembly line environment. This limits the applicability of the results of the analytical research in practical settings. Although a relatively recent line of work considers alternative manufacturing environments, an incomplete understanding of the practical and modelling challenges associated with production smoothing hinders the wider adoption of JIT philosophy in a variety of manufacturing environments. Therefore, this paper first discusses the practical and modelling challenges that arise in production smoothing in the context of JIT manufacturing. Then, an extensive review of the existing literature that focuses on analytical models and solution algorithms developed in the field is given. Finally, the gaps in the current body of knowledge are highlighted and several areas where further research is needed in production smoothing are identified.

Keywords: Production smoothing; Toyota production system; Just-in-time manufacturing

1. Introduction

The Toyota Production System (TPS) is an integrated set of tools and methods that focus on the identification and elimination of waste, and, hence, the improvement of productivity. 'Just-in-time' (JIT) philosophy, named after a phrase originated at Toyota Motor Company, recommends designing and controlling the manufacturing processes such that the required items are produced in the quantity needed when they are needed. To this end, TPS suggests that production should be triggered by demand, advocating the use of a pull system for production control. In particular, when pull production control is in effect, the production schedule for the last stage

*Corresponding author. Email: yavuz@reef.uf1.edu

International Journal of Production Research
ISSN 0020–7543 print/ISSN 1366–588X online
http://www.tandf.co.uk/journals
DOI: 10.1080/00207540701223410

of manufacturing operations is propagated through all stages of manufacturing operations. The objective of *production smoothing*, which is a tactical level planning decision also referred to as *Heijunka* or *level scheduling*, is to reduce the variability of the production rate at the final stage of manufacturing operations so as to create a stable demand stream for the other manufacturing operations at the preceding stages. Therefore, production smoothing is a key element of TPS, and, hence, a key component of the JIT philosophy (Walleigh 1986, Coleman and Vaghefi 1994, Monden 1998).

Due to customer expectations for increased product variety, manufacturing companies have been expanding their product mix to include a larger number of end-products each with several different variants. Hence, mixed-product systems, where manufacturing resources are shared among a family of multiple products each with possibly several variants, have become more prevalent in manufacturing industry and have been widely studied (Thomopoulos 1967, 1970, Macaskill 1972, Chakravarty and Shtub 1985, Merengo *et al.* 1999, Kim *et al.* 2000, Miltenburg 2002, Kara *et al.* 2004). In some industries such as the electronics industry, this increase in product variety leads to a high variance in demand. This, in turn, requires the enhancement of the capabilities of manufacturing systems to respond to the increased variability, possibly via the adoption of flexible manufacturing and agile manufacturing principles (Yusuf *et al.* 1999). However, in some industries such as the industrial equipment industry, although the product variety is high, the demand variability may still be relatively low as the companies make an effort to stabilize the demand for end-products using effective supply management strategies. In such industries, the use of JIT manufacturing principles is still a viable option. With the increased complexity of product structures and the level of diversification in product configurations, manufacturing operations have become increasingly more complex, rendering production smoothing for mixed-product JIT systems a considerably challenging problem (Sumichrast *et al.* 1992, Xiaobo and Ohno 1994, 1997, Bolat 1997, Hyun *et al.* 1998, Scholl *et al.* 1998, Celano *et al.* 1999, Ponnambalam *et al.* 2003, Kotani *et al.* 2004, Bock *et al.* 2006).

An increasing number of companies from various industries are interested in adopting JIT philosophy to improve their productivity, and, thereby, their competitiveness in the marketplace. In fact, a recent study reports that JIT philosophy has been adopted by numerous manufacturing companies from a gamut of industries, including electronics, industrial machinery, food and textile, among others (Fullerton and McWatters 2001). Adoption of JIT philosophy requires companies to restructure their manufacturing operations. Walleigh (1986) emphasizes the importance of production smoothing and claims that it is likely to be one of the first steps in transformation to JIT manufacturing. Lummus (1995) conducts a simulation study of a three-station assembly line (that withdraws sub-assemblies from three sub-lines with one, two and three stations, respectively) where different products have different setup and processing time requirements, and shows that the schedule obtained by the production smoothing methods primarily designed for synchronized assembly lines is as good as a randomly generated schedule for this system. Therefore, in order to facilitate the wider adoption of JIT manufacturing, there is a clear need to develop a set of analytical models and solution algorithms that address the production smoothing problem (PSP) in a variety of realistic manufacturing environments.

In this paper, our objective is to provide a critical review of the current analytical literature on production smoothing for mixed-product JIT manufacturing systems. Alternative analysis tools, such as simulation, are cited as relevant, however are beyond the scope of this review, since our interest lies in placing emphasis on analytical modelling and algorithm development for decision-making support. Upon investigating the degree to which the existing literature addresses the PSP in a variety of manufacturing environments that are encountered in practice, we will review the existing modelling and solution approaches and identify new avenues of research. The remainder of the paper is organized as follows. In section 2 we discuss practical and modelling issues in production smoothing. Sections 3–5 review the papers that are concerned with the PSP in assembly-line, single-machine and flow-shop environments, respectively. Finally, section 6 summarizes the literature reviewed throughout the paper, proposes several future research directions and concludes the paper.

2. A discussion of the practical and modelling issues in production smoothing

2.1 *Practical issues*

Monden's (1983) seminal work refers to production smoothing as ''the cornerstone of the TPS'', thereby stimulating a stream of research on production smoothing for mixed-product JIT systems, where the final manufacturing stage is a synchronized assembly line. Kubiak (1993) provides a comprehensive review of this analytical literature until 1993. However, a closer examination of the manufacturing environments inherent in industries today where JIT manufacturing is becoming increasingly more widespread reveals that the final stage of manufacturing operations is not necessarily comprised of a synchronized assembly line; it may be (i) a single machine shop, (ii) a flow shop, or (iii) a job shop, as well. Each of these manufacturing environments poses different analytical modelling challenges for the PSP, where the benefit to be obtained from production smoothing is the same and very important. That is, withdrawing sub-assemblies from upstream stages smoothly, thereby creating a stable demand for sub-assemblies. Motivated by automotive and electronics contract manufacturing plants, a recent stream of research addresses production smoothing for mixed-product JIT systems, where the final stage of manufacturing operations is a unit-processing single machine shop (Yavuz *et al*. 2006a, Yavuz and Tufekci 2006, 2007a) or a flow shop (Yavuz *et al*. 2006b, Yavuz and Tufekci 2007b).

Production smoothing at Toyota focuses on reducing the variability of the consumption rates of the sub-assemblies used at the final stage (Monden 1983). Miltenburg and Sinnamon (1989) extend Monden's (1983) approach by considering smoothing both end-product production rates and sub-assembly consumption rates at the preceding stages of the manufacturing system. Miltenburg (1989) is concerned with reducing the variability of the production rate for end-products at the final stage only. More specifically, Miltenburg and Sinnamon's (1989) approach focuses on controlling how frequently the sub-assemblies that are required for the end-products are pulled as well as the end-products are completed, whereas Miltenburg's (1989) approach concentrates only on controlling how frequently the end-products are completed. Kubiak (1993) refers to the former approach as the *output rate variation*

(ORV) and the latter as the *product rate variation* (PRV). In fact, a rich line of research addresses the PRV approach, whereas the stream that investigates the ORV approach is relatively limited in scope, possibly due in part to the complexity of the underlying planning problem.

2.2 *Modelling issues*

Monden (1983) identifies smoothing the usage of the sub-assemblies and the loading of resources as two important objectives in JIT manufacturing. *Usage goal* concentrates on the production rates of end-products, as well as the consumption rates of sub-assemblies that go into the end-products. The usage goal is generally measured as a function of the deviations of *actual* (cumulative) production/consumption amounts from pre-specified *ideal* (cumulative) production/consumption amounts. *Loading goal* concentrates on processing requirements and captures the deviation of actual workload levels on the production resources from the ideal workload levels.

Assuming that the demand rate for an end-product is constant and continuous, the 'ideal' cumulative production quantity for an *end-product* over time is commonly modeled in the literature using a linear function (figure 1). However, since production resources cannot process multiple different end-products simultaneously, achieving this ideal production rate is not feasible in practice. The actual cumulative production quantity for an end-product in any feasible schedule can be represented by a (discontinuous) piecewise-linear function, i.e. linearly increasing when the end-product is being produced and non-increasing otherwise, which is also depicted in figure 1. Then, the (shaded) area between the linear and the discontinuous piecewise-linear functions represents the deviation of the actual production quantities

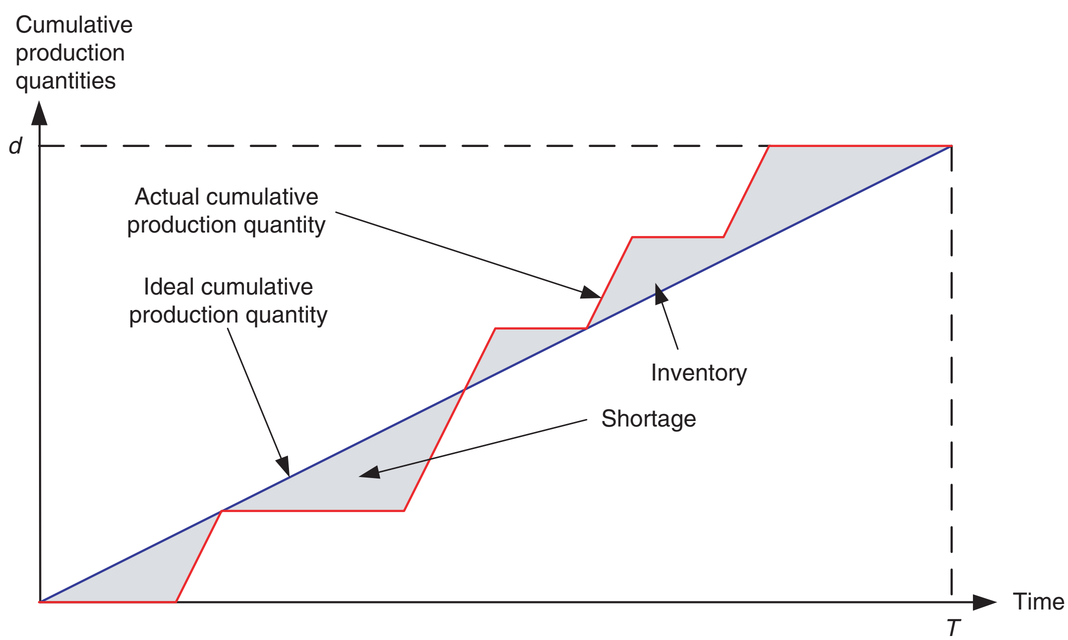

Figure 1. The ideal and actual cumulative production quantities.

from the ideal quantities. Clearly, the smaller this total area over the planning horizon, the better, i.e. *smoother*, the actual schedule is.

For both the usage and loading goals, both positive and negative deviations can be observed. Therefore, in formulating an objective function, we can either take the square or the absolute value of these deviations, which are referred to as the *squared* or *absolute value* objective functions, respectively. Finally, the optimization problem may be formulated to minimize either the total deviation or the maximum deviation, which are referred to as the *minsum* and *minmax* objective functions, respectively.

The modelling assumption that pertains to the characteristics of the setup times impacts the difficulty of the underlying optimization problems. In the literature, there is a significant amount of work that assumes *zero* setup times, which may be difficult to justify in practice, as they are typically *arbitrary* and *non-zero*, and possibly sequence-dependent. Similarly, the assumption that pertains to the characteristics of the processing times also impacts the complexity of the underlying optimization problems. The earlier work in the area assumes *unit* processing times, whereas the most recent work focuses on *arbitrary* processing times.

In our review of the literature, we pay close attention to these four identifying characteristics that relate to the practical and modelling issues, namely (i) the characteristics of the final stage of manufacturing operations; (ii) the scope of production smoothing activity; (iii) the characterization and formulation of the objective function; and (iv) the assumptions that relate to setup and processing times. As we have discussed above, each of these characteristics has an impact on the practical relevance of the underlying system and the computational complexity of the underlying optimization model.

3. PSP in assembly line systems

The majority of the PSP literature, primarily motivated by Monden's (1983) book, is concerned with synchronized assembly lines. The processing time required for each end-product at each station defines the cycle time of a synchronized assembly line. That is, at intervals of equal length, defined by the duration of the cycle time, an end-product leaves the assembly line as a finished product, all the units on the line complete their processing at a station and move to the immediate downstream station, and a unit of an end-product is released to the first station on the assembly line. Moreover, the setup times incurred to switch between different end-products are assumed to be negligible.

3.1 *Preliminaries*

The mixed-product JIT system is assumed to consist of L stages of manufacturing operations, referred to as the levels, indexed by ℓ. The final stage of manufacturing operations is an assembly line, and it is referred to as the first level (i.e. $\ell = 1$). Similarly, the first stage of the manufacturing operations is referred to as the Lth level (i.e. $\ell = L$). Each level ℓ processes n_ℓ different items. For instance, the first level processes n_1 different end-products, whereas intermediary stages process n_ℓ different sub-assemblies for $\ell = 2, \ldots, L$. The quantity of a sub-assembly i at level ℓ required to assemble a unit of end-product h is given by $b_{\ell,i,h}$. Demand for item i at level ℓ is

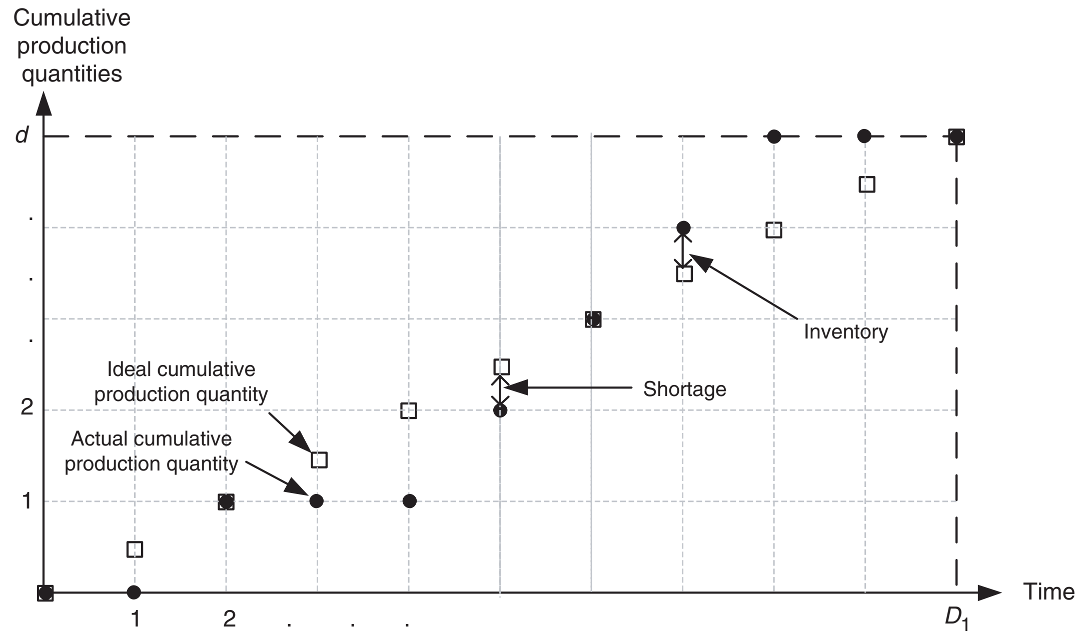

Figure 2. The ideal and actual cumulative production quantities in discrete time.

denoted by $d_{\ell,i}$ for $\ell = 1, \ldots, L$ and $i = 1, \ldots, n_\ell$. Finally, D_ℓ denotes the *total demand of level* ℓ, i.e. $D_\ell = \sum_{i=1}^{n_\ell} d_{\ell,i}$, and the *demand rate* for each item i at each level ℓ is denoted by $r_{\ell,i} = d_{\ell,i}/D_\ell$.

The production schedule for the first level is denoted by σ. σ consists of D_1 stages in total, and at each stage a single unit of an end-product can be processed, i.e. $\sigma_k \in \{1, \ldots, n_1\}$ for $k = 1, \ldots, D_1$. Let $x_{1,i,k}$ denote the cumulative quantity of end-product i *produced* over the first k stages of the sequence for $i = 1, \ldots, n_1$, and $k = 0, \ldots, D_1$. We have $x_{1,i,0} = 0$ and $x_{1,i,k} - x_{1,i,k-1} \in \{0, 1\}$ for $k = 1, \ldots, D_1$. Similarly, $x_{\ell,i,k}$ denotes the cumulative quantity of sub-assembly i *consumed* over the first k stages of the sequence for $\ell = 2, \ldots, L$, $i = 1, \ldots, n_\ell$, and $k = 0, \ldots, D_1$, and we have $x_{\ell,i,0} = 0$ and $x_{\ell,i,k} = x_{\ell,i,k-1} + b_{\ell,i,\sigma_k}$. Here, the cumulative consumption of a sub-assembly is initially zero and at each stage increases by the quantity needed to assemble the end-product processed at that stage.

Since the total demand for the final stage of manufacturing operations, i.e. the first level, is given by D_1, a production schedule for this level consists of D_1 stages. Therefore, it becomes possible to model the actual cumulative production/consumption quantities as discontinuous functions, as depicted in figure 2, and the deviation is not measured as the area between the two curves, but as the difference between the two functions at discrete points in time, which correspond to the end of each stage.

3.2 *PRV on assembly lines*

Kubiak (2003) shows that cyclic schedules are optimal for the PSP problem under a PRV approach. That is, if a schedule is known to be optimal for a given instance of the PRV, then repeating it c times will be the optimal production sequence for a larger problem where demand for the products is c times the original demands. In other words, if the demand values for different end-products have a common divisor, then they can all be divided into the greatest common divisor and a smaller problem can be solved. In this section, without loss of generality, we assume that the demands do not have a common divisor greater than one.

Miltenburg (1989) focuses on the usage goal for the end-products at the final stage of the manufacturing system (i.e. $\ell = 1$), and formulates the problem as an integer quadratic optimization model with an objective function of

$$\min \sum_{k=1}^{D_1} \sum_{i=1}^{n_1} (x_{1,i,k} - kr_{1,i})^2,$$

and constraints assuring that a single unit of an end-product is processed at each stage throughout the planning horizon. Aigbedo (2000) studies the structural properties of this formulation, and shows that $\sum_{i=1}^{n_1}(D_1^2 - d_{1,i}^2)/(12D_1)$ yields a tight lower bound on the objective function value of the optimal solution to the problem.

Early solution methods for Miltenburg's (1989) model include computationally expensive exact methods (Miltenburg 1989, Miltenburg *et al.* 1990) and constructive heuristics (Miltenburg 1989, Sumichrast *et al.* 1990, Ding and Cheng 1993, Cheng and Ding 1996).

The most efficient exact solution of the problem is due to Kubiak and Sethi (1991, 1994) who note that, for each individual unit (i.e. copy) of an item, it is possible to define an ideal position in the sequence. It is also possible to define a cost function that increases if a particular copy of an item deviates from its ideal position. Kubiak and Sethi's (1991) cost definition leads to a reformulation of Miltenburg's model as an assignment problem with D_1 elements, and, hence, is solvable in $O(D_1^3)$ time. Moreover, this reformulation can be used when the usage goal is formulated in the form $\sum_{k=1}^{D_1}\sum_{i=1}^{n_1} F_i(\cdot)$, where $F_i(\cdot)$ is a unimodal convex function that has a minimum of $F_i(0) = 0$ (Kubiak and Sethi 1991, 1994). Note that this definition covers both the squared and absolute value type objective functions, and can be generalized to cases where weights are associated with the end-products.

Using a similar approach, Inman and Bulfin (1991) define ideal positions for each copy of each end-product to be produced. Instead of measuring actual usage of the products in each stage of the sequence and then comparing them to the ideal usages, they consider measuring the deviation of their actual positions in the sequence from those ideal positions. This modified problem is solved with an efficient earliest-due-date (EDD) approach that also finds good solutions for the original formulation of Miltenburg (1989).

Steiner and Yeomans (1993) use the minmax absolute objective function and show that this formulation is reducible to the Release Date/Due Date Decision Problem, which can be solved to optimality with an EDD algorithm, in $O(D_1)$ time. The formulation using the minmax squared objective function is solvable in $O(n_1 D_1)$ time (Brauner and Crama 2004).

Existing models considering the loading goal under the PRV approach use a weighted objective function of both the usage and loading goals, where w_{U} and w_{L} denote the respective weights of these goals. Miltenburg *et al.* (1990)

denote the time to produce end-product i by t_i and then formulate the loading goal using

$$\min \sum_{k=1}^{D_1} \sum_{i=1}^{n_1} t_i^2 (x_{1,i,k} - kr_{1,i})^2.$$

The authors develop a dynamic programming (DP) procedure that solves the problem with the weighed objective function, which, however, has an exponential time complexity ($O(n_1 \prod_{i=1}^{n_1} (d_{1,i} + 1))$). Korkmazel and Meral (2001) distinguish between the processing time requirements on different stations, and penalize the difference between the actual production time that is spent for product i on workstation m and the ideal production time that should have been spent for product i on workstation m, in the first k positions of the sequence:

$$\min \sum_{k=1}^{D_1} \sum_{i=1}^{n_1} \sum_{m} W_m \left(\frac{t_{i,m}}{\bar{t}_m} \right)^2 \left(x_{1,i,k} - \gamma k \frac{d_{1,i}}{T_m} \right)^2,$$

where W_m is the weight associated with station m, $t_{i,m}$ is the processing time of end-product i on station m, $\bar{t}_m$ and T_m are the average and total processing times on station m, respectively, and γ is the cycle time. The authors state that the problem with the weighted objective function can be reduced to an assignment problem and solved efficiently. Closer observation reveals that the key to this transformation is the decomposition of the total workload into pieces of workload created by each end-product. The formulations which take the total cumulative workload as a whole, as will be seen in the next subsection, cannot be solved efficiently.

McMullen (1998) is concerned with setups that can be performed in negligible time, which, however, incur a significant cost, and utilizes the total number of setups as an additional objective. The minimization of a weighted sum of the minsum squared formulation of the usage goal and the number of setups is considered and implementations of tabu search (McMullen 1998), simulated annealing (McMullen and Frazier 2000, Cho *et al.* 2005), and genetic algorithms (McMullen *et al.* 2000) are presented in the literature. McMullen (2001a) uses the two objectives in a bi-criteria formulation and compares the efficient frontiers obtained by tabu search, simulated annealing and genetic algorithms. Later, for the same bi-criteria model, ant colony optimization (McMullen 2001b), artificial neural networks (McMullen 2001c), beam search (McMullen and Tarasewich 2005), and multi-objective genetic algorithms (Mansouri 2005) were proposed. Hyun *et al.* (1998) develop a multi-objective model that simultaneously aims to minimize the incomplete work, which is referred to as the total utility work in the literature, the minsum absolute deviation usage goal and the summation of sequence-dependent setup costs. The authors propose a multi-objective genetic algorithm to obtain an approximate set of Pareto-optimal sequences in the multi-objective setting.

Ventura and Radhakrishnan (2002) introduce batch processing for PSP on assembly lines using the PRV approach. The authors assume given batch sizes for the end-products. In this case, setup times are easily incorporated into the (integer) batch processing times that vary among the products. As different products require different times to process, the authors state that the emerging optimization problem is difficult and propose efficient heuristic procedures for its solution. This work is an important contribution to the PSP literature, in that it allows batch processing.

However, the assumption that the batch sizes are given limits the scope of the problem. Yavuz and Tufekci (2004a) consider batch processing, and extend Aigbedo's (2000) lower bound on the objective function to account for batch sizes.

A final important, and practically relevant, variant of the PSP on assembly lines under the PRV approach was studied by Drexl and Kimms (2001), where the PSP is considered in conjunction with the so-called *car sequencing problem* (CSP). The CSP is based on options, i.e. properties that a car may or may not possess. In formulating the CSP, one pre-processes the processing time requirements of cars with an option on the assembly station that installs that option, and generates constraints (of the form $H_o : N_o$) such that at most H_o cars in any subsequence of N_o cars can possess that option. For instance, processing time requirements may mandate that at most three cars with the sun-roof option are sequenced in any sub-sequence of five cars, i.e. constraint sun-*roof*:3:5 is added to the model. Drexl and Kimms (2001) use Inman and Bulfin's (1991) objective function and constraints in addition to the ($H_o : N_o$ type) constraints introduced by the CSP. For exact solution of the combined problem a column generation method is presented in Drexl and Kimms (2001) and a branch-and-bound algorithm in Drexl *et al.* (2006).

3.3 *ORV on assembly lines*

Monden (1998) formulates the ORV approach using an objective function that is formed by summing up squared deviations in sub-assembly consumption rates over the sub-assemblies and stages:

$$\min \sum_{k=1}^{D_1} \sum_{i=1}^{n_2} \left(x_{2,i,k} - k \frac{d_{2,i}}{D_1} \right)^2.$$

Miltenburg and Sinnamon (1989) generalize Monden's formulation of the problem by considering the deviation between the ideal and actual schedules at four levels. Their objective function incorporates all the four levels, with respect to their weights w_ℓ, $\ell = 1, \dots, 4$, and defines the ideal consumption quantity using the total consumption amount ($X_{\ell,k} = \sum_{i=1}^{n_\ell} x_{\ell,i,k}$) at a certain level, up to a certain stage in the sequence:

$$\min \sum_{k=1}^{D_1} \sum_{\ell=1}^{4} \sum_{i=1}^{n_\ell} w_\ell (x_{\ell,i,k} - X_{\ell,k} r_{\ell,i})^2.$$

Kubiak (1993) develops a generalization of ORV, which includes both the Monden and Miltenburg and Sinnamon models as special cases. He also shows that this general formulation is NP-hard. Complementing his work, Kubiak *et al.* (1997) show that the ORV approach with a minmax squared/absolute value objective function is strongly NP-hard. They develop a DP procedure that is capable of handling both the minsum/minmax squared/absolute value objective functions, running in $O(n_1 (\sum_{\ell=1}^{L} n_\ell) \prod_{i=1}^{n_1} (d_{1,i} + 1))$ time. Although this complexity is prohibitive, the authors note that

$$\prod_{i=1}^{n_1} (d_{1,i} + 1) \leq \left(\frac{D_1 + n_1}{n_1} \right)^{n_1},$$

and, therefore, the growth rate is polynomial in D_1. Consequently, the authors argue that the DP can be effective for those cases in which n_1 is small, even when the total product demand, D_1, is large. Considering the number of feasible sequences, $D_1!/(d_{1,1}!d_{1,2}!\ldots d_{1,n_1}!)$, the DP requires less computational effort than explicit enumeration.

As the PSP under the ORV approach is a difficult optimization problem, heuristic solution procedures that find good solutions with reasonable computational effort are desired to solve real-life instances of the problem (Inman and Bulfin 1992, Miltenburg and Sinnamon 1992, 1995, Cakir and Inman 1993, Morabito and Kraus 1995, Bautista *et al.* 1996, Duplaga *et al.* 1996, Aigbedo and Monden 1997, Duplaga and Bragg 1998, Zeramdini *et al.* 2000, Ding *et al.* 2006). Two widely recognized heuristics developed at Toyota are Monden's (1983) Goal Chasing Method (GCM-I) and Goal Chasing Method II (GCM-II) methods. Miltenburg and Sinnamon (1989) propose two constructive (namely, one- and two-stage) heuristics for ORV. The first (respectively, second) heuristic considers the next stage (respectively, next two stages) of the sequence and assigns the end-product that minimizes the total deviation at that stage (respectively, those two stages). The authors show that Toyota's goal chasing methods (GCM-I and GCM-II) are special cases of the one-stage heuristic that can be obtained by setting $w_1 = w_3 = w_4 = 0$ and $w_2 = 1$.

Several researchers have implemented meta-heuristic methods including genetic algorithms (Leu *et al.* 1996), beam search (Leu *et al.* 1997), ant colony optimization (Sun and Sun 2005), and a multi-agent method (Caridi and Sianesi 2000) on the ORV approach in order to obtain better solutions than the (relatively) simple heuristic approaches mentioned above. Ponnambalam *et al.* (2003) develop a genetic algorithm for the Miltenburg and Sinnamon model. They also present a multi-objective genetic algorithm to solve an extension of the problem that aims at minimizing total utility work, the minsum squared ORV usage goal and summation of sequence-dependent setup costs also studied by Hyun *et al.* (1998).

A challenging variant of the ORV approach is the consideration of batch processing. In this case, not only the sequences of the sub-assemblies and end-products, but also their respective processing batch sizes need to be determined. Yavuz and Tufekci (2004a) generalize the Monden and Miltenburg and Sinnamon models to incorporate the batch sizes. They study the structural properties of these formulations to develop lower bounds on the objective function values. However, to the best of our knowledge, there is no paper that adopts the ORV approach and develops sequencing methods (optimization or heuristic algorithms) for the final assembly line with batch processing in the literature.

In the majority of the quantitative work that focuses on ORV, where the loading goal is also studied (Miltenburg and Goldstein 1991, Tamura *et al.* 1999), the loading goal has a single-level structure, i.e. considers smoothing the workload of production resources at the final stage only. An exception is due to Aigbedo and Monden (1997), who consider the sub-assembly level as well as the final assembly level, and formulate the loading goal to capture the variation in the station loads over those two levels. The authors propose a solution method that requires the user to prioritize usage and loading goals for the two levels, and finds a sequence that satisfies all four goals simultaneously, breaking ties according to the priorities provided by the user.

Kurashige *et al*. (2002) address the problem of sequencing mixed-product assembly lines, where processing time requirements vary among the end-products on the assembly stations. They present a heuristic named *time-based goal-chasing* (TBGC), that evaluates the end-products and selects the one with the minimum total sub-assembly usage deviation at completion time to processing time ratio. In other words, they aim to minimize deviation captured by the shaded area in figure 1, and with the TBGC method they construct a sequence by assigning the end-product that minimizes the average deviation incurred during its processing.

As the ORV approach yields a challenging optimization problem, some researchers have considered simplifying assumptions to obtain efficiently solvable special cases of ORV. For instance, Steiner and Yeomans (1996) are concerned with sub-assemblies that are not commonly used by the end-products. That is, a sub-assembly is needed for only one of the end-products. The authors show that the optimization problem can be transformed into an assignment problem and solved efficiently. Similarly, Duplaga *et al*. (1996), based on the approach in use at Hyundai, focus on the usage of the options instead of sub-assemblies, and for each option they use the most critical sub-assembly. While their scheduling method is essentially the same as Toyota's GCM-I, the authors argue that by reducing the number of sub-assemblies considered in formulating the objective function, significant savings in computational effort are obtained.

All the quantitative work on ORV we discuss above assume that when an end-product is released to the first assembly station, all the sub-assemblies required should be ready at the first station of the assembly line. Considering the actual consumption (pull) times of sub-assemblies, Sumichrast and Clayton (1996) show that the models and solution approaches above may not yield better solutions than randomly generated sequences. Xiaobo *et al*. (1999) give an example from the Motomachi Factory of the Toyota Motor Corporation that assembles 'Crown' series automobiles, where an end-product passes through the assembly line in approximately seven hours. Therefore, if a particular sub-assembly is needed in the last station of the assembly line, it would be a waste to get that sub-assembly ready when the end-product that the sub-assembly will go into starts its assembly at the first station. The authors propose a modified GCM that constructs a sequence by selecting the end-product that minimizes the total deviation in sub-assembly usage over all the stations at each stage. Also, Xiaobo and Zhou (1999) develop a simulated annealing approach for the problem, and conclude that simulated annealing yields optimal solutions for small-sized problems within acceptable computation times.

4. PSP in single-machine systems

A recent line of work on production smoothing focuses on systems where the final stage of manufacturing operations is performed on a single machine. Motivated by an automotive parts manufacturing facility, Yavuz and Tufekci (2004b) adopt the PRV approach of the PSP in a special setting, where the final stage of manufacturing is a single machine shop with arbitrary setup and processing time requirements. The presence of non-zero setup times hinders the implementation of the ideal one-piece-flow, when the total available time is not sufficient to make one setup for each unit of each product. In this environment, both the number of setups required for each end-product as well as the sequencing of the different setups should be chosen. The authors consider a system where a fixed-length time-bucket is defined that can be dedicated to the processing of a particular type of end-product. That is, the time available in a bucket can be utilized for the necessary setup, processing of one or more units of the end-product, and, possibly, some idle time. The authors propose a two-phase solution method, where the number of setups for each product (number of batches), batch sizes, and the length of the time-bucket are determined in the first phase and the batches are sequenced in the second phase. Using the time-bucket makes the second phase problem relatively easy, as the problem is a discrete-time PRV approach and can be efficiently solved as an assignment problem. The first phase problem, on the other hand, is shown to be NP-hard (Yavuz and Tufekci 2004b). The authors propose an enumerative procedure that can solve small-sized instances of the problem and a parametric heuristic procedure to solve medium-sized instances.

Later, Yavuz and Tufekci (2006) developed a bounded DP procedure that can be used to obtain the optimal solution for instances that are of relatively larger size. However, due to the combinatorial nature of the problem, exact optimization approaches are prohibitively time consuming for the large-sized instances. Yavuz *et al*. (2006a) implement three meta-heuristic methods, namely strategic oscillation, scatter search and path re-linking, for the problem. Their computational analysis shows that the path re-linking approach provides high-quality solutions with relatively small computational effort.

5. PSP in flow-shop systems

A flow shop consists of a set of machines placed in series, where a number of products is processed on all of the machines following the same processing sequence. Although the sequence of the operations is common among the products, the products do not need to be processed in the same order on all of the machines. That is, each machine can select the next product to process from the set of products that have already completed their previous operations and arrived at that machine. In that case, the number of possible schedules grows exponentially with the number of machines in the flow shop. A common approach to managing flow shops is to use permutation schedules, i.e. all the machines process the products in the order they arrive. Note that, if we ignore the setup times and assume unit processing times for each product on each machine, then the permutation flow shop is identical to a synchronized assembly line. The PSP on flow shops without these restrictive assumptions has received limited interest in the literature.

McMullen (2002) introduces the concept of production smoothing for manufacturing systems where the final stage of manufacturing operations is a flow shop with sequence-dependent setup times for the end-products. He considers the usage goal for the PRV approach in conjunction with minimizing the makespan (which is the completion time of the last product in the flow shop, i.e. the time required to produce all the products). McMullen (2002) presents a bi-criteria model, and uses simulated annealing to approximate the efficient frontier.

Yavuz and Tufekci (2007b) formulate the PRV approach for flow shops with arbitrary setup and processing times that vary among the end-products as well as the machines. The authors extend the time-bucket approach (originally developed for single-machine shops) and also the two-phase solution approach to this flow shop environment. They show that once the batching decisions are made and the length of the time-bucket is established, the batches can be sequenced with relative ease. They show that the batching problem is NP-hard, and propose a bounded DP procedure to obtain the optimal solution for the problem. Yavuz *et al*. (2006b) develop a hybrid meta-heuristic approach that inherits components from strategic oscillation and path re-linking and yields high-quality solutions with little computational effort.

6. Discussion and future research directions

The papers reviewed in sections 3–5 contribute to our understanding of the modelling and algorithmic challenges associated with the solution of the production smoothing problem that arises in the context of JIT manufacturing. Our review of the literature reveals that the existing work remains short of (1) addressing the PSP in a variety of manufacturing environments that are encountered in practice and (2) placing an emphasis on realistic considerations that pertain to the characteristics of these manufacturing environments.

For assembly line systems, for instance, with a few exceptions, the majority of existing work assumes unit processing times and negligible setup times. This assumption obviously leads to a situation where an end-product completes its processing on the assembly line at equal intervals. That is, the main research focus to date is placed on synchronized assembly lines. Moreover, in the ORV models, i.e. where sub-assembly consumptions are considered, it is assumed that all the required sub-assemblies to assemble an end-product are needed at the beginning of the assembly of that end-product. That is, the assembly sequence of the end-products is the same as the consumption sequence of the sub-assemblies. Similarly, for single-machine and flow-shop systems, the research is still in its infancy, as the papers reviewed in this section are rather limited in scope, since only the PRV approach has been considered for these types of manufacturing systems. Hence, the existing analytical literature analyses the PSP under rather restrictive assumptions, which may not hold in practice. Therefore, to help bridge the gap between the academic literature and industry practice, there is a need to build models that consider different manufacturing environments and incorporate the practical issues we discuss in section 2 and develop effective algorithms to obtain solutions for these models. Below, we indicate several specific directions for future research in production smoothing to address the needs of industry, and, thereby, help facilitate the broader dissemination of JIT manufacturing.

First and foremost, the practical challenges that have not been considered to date must be addressed. As emphasized throughout the paper, primarily motivated by the TPS, the PSP has been studied mainly on synchronized assembly lines with restrictive assumptions regarding the processing and setup times. While these assumptions simplify the problem on hand and enable researchers to develop very efficient solution methods, it limits the applicability of this line of analytical work in practice, where the final stage of manufacturing operations does not constitute a synchronized assembly line. Therefore, there is a need to consider the problem in different manufacturing contexts, and some recent work considers single-machine and flow-shop environments as the final-stage of a multi-stage JIT manufacturing system. In fact, Cruickshanks *et al*. (1984) discuss a PSP in job shops which aims at finding the optimal production and inventory levels over a finite planning horizon. However, to the best of the authors' knowledge, the idea of mixed-product scheduling for smooth workload and part usages in job shops has not been studied. Furthermore, the scheduling literature considers a rich variety of manufacturing environments including open shops and cycle shops in addition to those mentioned thus far (see, for example, Leung (2004)). Also, the manufacturing system under consideration may include batch processors capable of processing multiple units of the same or different products at the same time and/or not necessarily identical parallel machines to perform certain operations on the products. Hence, the consideration of alternative manufacturing systems is an important research direction. Moreover, in the limited literature on the PSP for single-machine and flow-shop systems, only the PRV approach has been taken. Therefore, the ORV approach should also be studied in various manufacturing systems.

Another important practical challenge pertains to the setup time requirements. A recent line of work in the production smoothing literature considers arbitrary setup time requirements incurred when switching between end-products, whereas these requirements are generally overlooked in earlier work. Despite its importance in the JIT philosophy, reduction of setup times may not be achieved in all industries. Therefore, in manufacturing systems where significant, and possibly sequence-dependent, setup times are unavoidable, the production smoothing problem remains a challenge, and, hence, a key research direction for the broader dissemination of JIT manufacturing.

One of the cornerstones of the JIT philosophy is equipment maintenance. The JIT philosophy advocates practicing preventive maintenance in order to reduce the number of machine failures and total lost productive time. However, the production smoothing literature assumes uninterrupted availability of production resources. Therefore, both proactive and reactive approaches need to be developed for smooth production in the case of failures. For instance, in a manufacturing system with significant setup times, one may temporarily increase batch sizes to make up for the time lost (repair time) and catch up with the ideal production quantities. A fast analytical tool to update batch sequences and their sequences, in this regard, would be very practical. Moreover, preventive, and possibly predictive, maintenance activities can be scheduled with an integrated *production and maintenance smoothing* approach.

From a modelling perspective, in a majority of the existing papers, the sub-assemblies are assumed to be required at the moment an end-product starts its processing at the final stage. However, in reality, an end-product may require a certain sub-assembly at the last station of a multi-station assembly line or at the last machine of a flow shop. Therefore, models that take actual consumption points into account can help further reduce the inventory accumulations before the final stage (Xiaobo and Zhou 1999, Xiaobo *et al*. 1999). A further elaboration of this point is that in manufacturing environments where processing times vary among the products, the consumption of sub-assemblies depends on the processing times, in addition to the sequence of products at the final stage. Therefore, a scheduling approach should be adopted instead of the current sequencing approach. We propose the term *continuous-time production smoothing* to address this important future research direction in production smoothing.

In today's competitive environment, efficiencies are sought at not only the manufacturing company level but at the supply chain level, rendering transitions to JIT supply chains critical. Despite the fact that the supply chain management field has developed a tremendous literature, the work considering smooth product flows in JIT supply chains is limited. Kubiak (2005) introduces the concept of mixed-product supply chains and presents a model that aims to smooth the production sequence at the final stage of a supply chain. Aigbedo (2004) is also concerned with minimizing variation in sub-assemblies usage in mixed-product supply chains. He emphasizes that, in many situations, suppliers deliver required parts in a constant order cycle, variable order quantity framework, and proposes several new metrics to evaluate the variation in parts usage. Optimizing the production schedules at the final stage of the supply chain with respect to the variation and (delivery, etc.) costs incurred by the suppliers is another important area for future research.

Focusing on the delivery of finished products, we note that the existing production smoothing literature assumes a constant and continuous distribution of end-products to the customers throughout the planning horizon, i.e. ignores delivery times that may be specified by the customers. In the case of firm (on-the-book) orders for one or more of the end-products, the existing models and solution approaches cannot be used, thereby creating a need to develop new approaches for production smoothing (Sinnamon and Milner 1995, Inman and Schmeling 2003). Furthermore, with the firm orders from the customers in mind, the distribution of the finished goods to customers can also be incorporated into a more comprehensive model. We propose the term *production and distribution smoothing* for this integrated approach.

References

Aigbedo, H., Some structural properties for the just-in-time level schedule problem. *Prod. Plann. Contr.*, 2000, **11**, 357–362.

Aigbedo, H., Analysis of parts requirements variance for a JIT supply chain. *Int. J. Prod. Res.*, 2004, **42**, 417–430.

Aigbedo, H. and Monden, Y., A parametric procedure for multicriterion sequence scheduling for just-in-time mixed-model assembly lines. *Int. J. Prod. Res.*, 1997, **35**, 2543–2564.

Bautista, J., Companys, R. and Corominas, A., Heuristics and exact algorithms for solving the Monden problem. *Eur. J. Oper. Res.*, 1996, **88**, 101–113.

Bock, S., Rosenberg, O. and van Brackel, T., Controlling mixed-model assembly lines in real-time by using distributed systems. *Eur. J. Oper. Res.*, 2006, **168**, 880–904.

Bolat, A., Efficient methods for sequencing minimum job sets on mixed model assembly lines. *Nav. Res. Logist.*, 1997, **44**, 419–437.

Brauner, N. and Crama, Y., The maximum deviation just-in-time scheduling problem. *Discr. Appl. Math.*, 2004, **134**, 25–50.

Cakir, A. and Inman, R.R., Modified goal chasing for products with non-zero/one bills of material. *Int. J. Prod. Res.*, 1993, **31**, 107–115.

Caridi, M. and Sianesi, A., Multi-agent systems in production planning and control: an application to the scheduling of mixed-model assembly lines. *Int. J. Prod. Econ.*, 2000, **68**, 29–42.

Celano, G., Fichera, S., Grasso, V., La Commare, U. and Perrone, G., An evolutionary approach to multi-objective scheduling of mixed model assembly lines. *Comput. Ind. Engng*, 1999, **37**, 69–73.

Chakravarty, A.K. and Shtub, A., Balancing mixed model lines with in-process inventories. *Mgmt Sci.*, 1985, **31**, 1161–1174.

Cheng, L. and Ding, F.Y., Modifying mixed-model assembly line sequencing methods to consider weighted variations for just-in-time production systems. *IIE Trans.*, 1996, **28**, 919–927.

Cho, H.S., Paik, C.H., Yoon, H.M. and Kim, H.G., A robust design of simulated annealing approach for mixed-model sequencing. *Comput. Ind. Engng*, 2005, **48**, 753–764.

Coleman, B.J. and Vaghefi, M.R., Heijunka (?): a key to the Toyota Production System. *Prod. Invent. Mgmt J.*, 1994, **35**, 31–35.

Cruickshanks, A.B., Drescher, R.D. and Graves, S.C., A study of production smoothing in a job shop environment. *Mgmt Sci.*, 1984, **30**, 368–380.

Ding, F.Y. and Cheng, L., A simple sequencing algorithm for mixed-model assembly lines in just-in-time production systems. *Oper. Res. Lett.*, 1993, **13**, 27–36.

Ding, F.Y., Zhu, J. and Sun, H., Comparing two weighted approaches for sequencing mixed-model assembly lines with multiple objectives. *Int. J. Prod. Econ.*, 2006, **102**, 108–131.

Drexl, A. and Kimms, A., Sequencing JIT mixed-model assembly lines under station-load and part-usage constraints. *Mgmt Sci.*, 2001, **47**, 480–491.

Drexl, A., Kimms, A. and Matthiessen, L., Algorithms for the car sequencing and the level scheduling problem. *J. Schedul.*, 2006, **9**, 153–176.

Duplaga, E.A. and Bragg, D.J., Mixed-model assembly line sequencing heuristics for smoothing component parts usage: a comparative analysis. *Int. J. Prod. Res.*, 1998, **36**, 2209–2224.

Duplaga, E.A., Hahn, C.K. and Hur, D., Mixed-model assembly line sequencing at Hyundai motor company. *Prod. Invent. Mgmt J.*, 1996, 20–25.

Fullerton, R.R. and McWatters, C.S., The production performance benefits from JIT implementation. *J. Oper. Mgmt*, 2001, **19**, 81–96.

Hyun, C.J., Kim, Y. and Kim, Y.K., A genetic algorithm for multiple objective sequencing problems in mixed model assembly lines. *Comput. Oper. Res.*, 1998, **25**, 675–690.

Inman, R.R. and Bulfin, R.L., Sequencing JIT mixed-model assembly lines. *Mgmt Sci.*, 1991, **37**, 901–904.

Inman, R.R. and Bulfin, R.L., Quick and dirty sequencing for mixed-model multi-level JIT systems. *Int. J. Prod. Res.*, 1992, **30**, 2011–2018.

Inman, R.R. and Schmeling, D.M., Algorithm for agile assembling-to-order in the automotive industry. *Int. J. Prod. Res.*, 2003, **41**, 3831–3848.

Kara, Y., Ozcan, U. and Peker, A., A simulated annealing approach for balancing and sequencing of mixed-model U-lines, in *Yoneylem Arastirmasi-Endustri Muhendisligi XXIV Ulusal Kongresi*, 2004, pp. 238–240.

Kim, Y.K., Kim, S.J. and Kim, J.Y., Balancing and sequencing mixed-model U-lines with a co-evolutionary algorithm. *Prod. Plann. Contr.*, 2000, **11**, 754–764.

Korkmazel, T. and Meral, S., Bicriteria sequencing methods for the mixed-model assembly line in just-in-time production systems. *Eur. J. Oper. Res.*, 2001, **131**, 188–207.

Kotani, S., Ito, T. and Ohno, K., Sequencing problem for a mixed-model assembly line in the Toyota Production System. *Int. J. Prod. Res.*, 2004, **42**, 4955–4974.

Kubiak, W., Minimizing variation of production rates in just-in-time systems: a survey. *Eur. J. Oper. Res.*, 1993, **66**, 259–271.

Kubiak, W., Cyclic just-in-time sequences are optimal. *J. Glob. Optim.*, 2003, **27**, 333–347.

Kubiak, W., Balancing mixed-model supply chains. In *Graph Theory and Combinatorial Optimization*, Vol. 8, Chapter 6, edited by D. Avis, A. Hertz and O. Marcotte, 2005 (Springer: New York).

Kubiak, W. and Sethi, S., A note on 'Level schedules for mixed-model assembly lines in just-in-time production systems'. *Mgmt Sci*., 1991, **37**, 121–122.

Kubiak, W. and Sethi, S.P., Optimal Just-in-Time schedules for flexible transfer lines. *Int. J. Flex. Mfg Syst*., 1994, **6**, 137–154.

Kubiak, W., Steiner, G. and Yeomans, S., Optimal level schedules for mixed-model, multi-level just-in-time assembly systems. *Ann. Oper. Res*., 1997, **69**, 241–259.

Kurashige, K., Yanagawa, Y., Miyazaki, S. and Kameyama, Y., Time-based goal chasing method for mixed-model assembly line problem with multiple work stations. *Prod. Plann. Contr*., 2002, **13**, 735–745.

Leu, Y., Huang, P.Y. and Russell, R.S., Using beam search techniques for sequencing mixed-model assembly lines. *Ann. Oper. Res*., 1997, **70**, 379–397.

Leu, Y., Matheson, L.A. and Rees, L.P., Sequencing mixed-model assembly lines with genetic algorithms. *Comput. Engng*, 1996, **30**, 1027–1036.

Leung, J.Y.-T. (editor), *Handbook of Scheduling: Algorithms, Models and Performance Analysis*, 2004 (Chapman and Hall/CRC: New York).

Lummus, R.R., A simulation analysis of sequencing alternatives for JIT lines using kanbans. *J. Oper. Mgmt*, 1995, **13**, 183–191.

Macaskill, J.L.C., Production-line balances for mixed-model lines. *Mgmt Sci*., 1972, **19**, 423–434.

Mansouri, S.A., A multi-objective genetic algorithm for mixed-model sequencing on JIT lines. *Eur. J. Oper. Res*., 2005, **167**, 696–716.

McMullen, P.R., JIT sequencing for mixed-model assembly lines with setups using tabu search. *Prod. Plann. Contr*., 1998, **9**, 504–510.

McMullen, P.R., An efficient frontier approach to addressing JIT sequencing problems with setups via search heuristics. *Comput. Ind. Engng*, 2001a, **41**, 335–353.

McMullen, P.R., An ant colony optimization approach to addressing a JIT sequencing problem with multiple objectives. *Artif. Intell. Engng*, 2001b, **15**, 309–317.

McMullen, P.R., A Kohonen self-organizing map approach to addressing a multiple objective, mixed-model JIT sequencing problem. *Int. J. Prod. Econ*., 2001c, **72**, 59–71.

McMullen, P.R., The permutation flow shop problem with just in time production considerations. *Prod. Plann. Contr*., 2002, **13**, 307–316.

McMullen, P.R. and Frazier, G.V., A simulated annealing approach to mixed-model sequencing with multiple objectives on a just-in-time line. *IIE Trans*., 2000, **32**, 679–686.

McMullen, P.R. and Tarasewich, P., A beam search heuristic mixed-model scheduling with setups. *Int. J. Prod. Econ*., 2005, **96**, 273–283.

McMullen, P.R., Tarasewich, P. and Frazier, G.V., Using genetic algorithms to solve the multi-product JIT sequencing problem with set-ups. *Int. J. Prod. Res*., 2000, **38**, 2653–2670.

Merengo, C., Nava, F. and Pozzetti, A., Balancing and sequencing manual mixed-model assembly lines. *Int. J. Prod. Res*., 1999, **37**, 2835–2860.

Miltenburg, J., Level schedules for mixed-model assembly lines in just-in-time production systems. *Mgmt Sci*., 1989, **35**, 192–207.

Miltenburg, J., Balancing and scheduling mixed-model U-shaped production lines. *Int. J. Flex. Mfg Syst*., 2002, **14**, 119–151.

Miltenburg, J. and Goldstein, T., Developing production schedules which balance part usage and smooth production loads for just-in-time production systems. *Nav. Res. Logist*., 1991, **38**, 893–910.

Miltenburg, J. and Sinnamon, G., Sequencing mixed-model multi-level just-in-time production systems. *Int. J. Prod. Res*., 1989, **27**, 1487–1509.

Miltenburg, J. and Sinnamon, G., Algorithms for scheduling multi-level just-in-time production systems. *IIE Trans*., 1992, **24**, 121–130.

Miltenburg, J. and Sinnamon, G., Revisiting the mixed-model multi-level just-in-time scheduling problem. *Int. J. Prod. Res*., 1995, **33**, 2049–2052.

Miltenburg, J., Steiner, G. and Yeomans, S., A dynamic programming algorithm for scheduling mixed-model just-in-time production systems. *Math. Comput. Model*., 1990, **13**, 57–66.

Monden, Y., *Toyota Production System*, 1983 (Industrial Engineering & Management Press).

Monden, Y., *Toyota Production System: An Integrated Approach to Just-In-Time*, 3rd ed., 1998 (Engineering & Management Press).

Morabito, M.A. and Kraus, M.E., A note on Scheduling mixed-model multi-level just-in-time production systems . *Int. J. Prod. Res.*, 1995, **33**, 2061–2063.

Ponnambalam, S.G., Aravindan, P. and Rao, M.S., Genetic algorithms for sequencing problems in mixed model assembly lines. *Comput. Ind. Engng*, 2003, **45**, 669–690.

Scholl, A., Klein, R. and Domschke, W., Pattern based vocabulary building for effectively sequencing mixed-model assembly lines. *J. Heurist.*, 1998, **4**, 359–381.

Sinnamon, G. and Milner, S., Just-in-time schedules for the small make-to-order shop. *Can. J. Admin. Sci.*, 1995, **12**, 340–351.

Steiner, G. and Yeomans, S., Level schedules for mixed-model, just-in-time processes. *Mgmt Sci.*, 1993, **39**, 728–735.

Steiner, G. and Yeomans, S., Optimal level schedules in mixed-model, multi-level JIT assembly systems with pegging. *Eur. J. Oper. Res.*, 1996, **95**, 38–52.

Sumichrast, R.T. and Clayton, E.R., Evaluating sequences for paced, mixed-model assembly lines with JIT component fabrication. *Int. J. Prod. Res.*, 1996, **34**, 3125–3143.

Sumichrast, R.T., Russell, R.S. and Taylor, B.W., Evaluating mixed-model assembly line sequencing heuristics for just-in-time production systems. *J. Oper. Mgmt*, 1990, **9**, 371–390.

Sumichrast, R.T., Russell, R.S. and Taylor, B.W., A comparative analysis of sequencing procedures for mixed-model assembly lines in a just-in-time production system. *Int. J. Prod. Res.*, 1992, **30**, 199–214.

Sun, X.-Y. and Sun, L.-Y., Ant colony optimization algorithms for scheduling the mixed model assembly lines, in *ICNC (3)*, 2005, pp. 911–914 (http://dx.doi.org/10.1007/11539902_112, http://dblp.uni-trier.de).

Tamura, T., Long, H. and Ohno, K., A sequencing problem to level part usage rates and work loads for a mixed-model assembly line with a bypass subline. *Int. J. Prod. Econ.*, 1999, **60/61**, 557–564.

Thomopoulos, N.T., Line balancing-sequencing for mixed-model assembly. *Mgmt Sci.*, 1967, **14**, B59–B75.

Thomopoulos, N.T., Mixed model line balancing with smoothed station assignments. *Mgmt Sci.*, 1970, **16**, 593–603.

Ventura, J.A. and Radhakrishnan, S., Sequencing mixed-model assembly lines for a just-in-time production system. *Prod. Plann. Contr.*, 2002, **13**, 199–210.

Walleigh, R.C., What's your excuse for not using JIT. *Harv. Bus. Rev.*, 1986, **64**, 38–54.

Xiaobo, Z. and Ohno, K., A sequencing problem for a mixed model assembly line in a JIT production system. *Comput. Ind. Engng*, 1994, **27**, 71–74.

Xiaobo, Z. and Ohno, K., Algorithms for sequencing mixed models on an assembly line in a JIT production system. *Comput. Ind. Engng*, 1997, **32**, 47–56.

Xiaobo, Z. and Zhou, Z., Algorithms for Toyota's goal of sequencing mixed models on an assembly line with multiple stations. *J. Oper. Res. Soc.*, 1999, **50**, 704–710.

Xiaobo, Z., Zhou, Z. and Asres, A., A note on Toyota's goal of sequencing mixed models on an assembly line. *Comput. Ind. Engng*, 1999, **36**, 57–65.

Yavuz, M. and Tufekci, S., Some lower bounds on the mixed-model level-scheduling problems, in *10th International Conference on Industry, Engineering and Management Systems*, 2004a, pp. 385–395.

Yavuz, M. and Tufekci, S., The single-level batch production smoothing problem: an analysis and a heuristic solution. Department of Industrial and Systems Engineering, University of Florida, 05, 2004b.

Yavuz, M. and Tufekci, S., Bounded dynamic programming solution to the batching problem in mixed-model just-in-time manufacturing systems. *Int. J. Prod. Econ.*, 2006, **103**, 841–862.

Yavuz, M. and Tufekci, S., An analysis and solution to the single-level batch production smoothing problem. *Int. J. Prod. Res.*, 2007a, **17**, forthcoming.

Yavuz, M. and Tufekci, S., Dynamic programming solution to the batching problem in just-in-time flowshops. *Comput. Ind. Engng*, 2007b, **51**, 416–432.

Yavuz, M., Akcali, E. and Tufekci, S., Optimizing production smoothing decisions via batch selection for mixed-model just-in-time manufacturing systems with arbitrary setup and processing times. *Int. J. Prod. Res.*, 2006a, **44**, 3061–3081.

Yavuz, M., Akcali, E. and Tufekci, S., A hybrid meta-heuristic for the batching problem in just-in-time flow shops. *J. Math. Model. Algor.*, 2006b, **5**, 371–393.

Yusuf, Y.Y., Sarhadi, M. and Gunasekaran, A., Agile manufacturing: the drivers, concepts and attributes. *Int. J. Prod. Econ.*, 1999, **62**, 33–43.

Zeramdini, W., Aigbedo, H. and Monden, Y., Bicriteria sequencing for just-in-time mixed-model assembly lines. *Int. J. Prod. Res.*, 2000, **38**, 3451–3470.