Expert Systems with Applications 39 (2012) 11001–11010

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

Modeling and solving a variant of the mixed-model sequencing problem with work overload minimisation and regularity constraints. An application in Nissan's Barcelona Plant

Joaquín Bautista *, Alberto Cano, Rocío Alfaro

Nissan Chair – Universitat Politècnica de Catalunya, Avda. Diagonal 647, 08028 Barcelona, Spain

ARTICLE INFO

Keywords:
Manufacturing
Sequencing
Just-in-time
Work overload
Linear programming
Production mix preservation

ABSTRACT

In this paper, we present an extension to the mixed-model sequencing problem with work overload minimisation (*MMSP-W*) for production lines with serial workstations and parallel homogeneous processors. The extension is intended to meet the industrial need to sequence various products so that the work required and completed at the workstations over time is as constant as possible. Several approaches are proposed to formulate the problem, giving rise to different models. Four models are selected, and their performances are compared after running the Gurobi solver, using instances from the literature and a case study of the Nissan powertrain plant in Barcelona.

1. Introduction

Production and assembly lines are the paradigm for product-oriented production systems. This type of manufacturing system is very common in the automotive industry, particularly in production environments based on manufacturing philosophies such as just-in-time (*JIT*) production and douki seisan (synchronised production), and is present throughout the manufacturing production chain (e.g., the body shop, paint shop, trim and chassis shop and engine plant).

These lines usually consist of a set (K) of workstations laid out in series. Each workstation ($k = 1,\ldots,|K|$) is characterised by the use of the human resources, tools and automated systems necessary to carry out the work assigned to the workstation. The set of tasks assigned to the workstation is called the workload, and the average time required to process these tasks at normal activity rates is called the workload time or the processing time.

An important attribute of these production lines is flexibility. The products (such as engines or car bodies) circulating through the lines are not completely identical. Although some of the products may be similar or of the same type, they may require different resources and components and therefore may require different processing times.

The desired flexibility of these mixed-product lines requires that the sequence in which the product types are manufactured follow two general principles: (1) to minimise the stock of components and semi-processed products and (2) to maximise the efficiency of the line, manufacturing the products in the least amount of time possible.

A classification of sequencing problems arising in this context was given in Boysen et al. (2009):

1. *Mixed-model sequencing*. The aim in this problem is to obtain sequences that complete the maximum work required by the work schedule.
2. *Car sequencing*. These problems are designed to obtain sequences that meet a set of constraints related to the frequency with which the workstations are required to incorporate special options (e.g., a sunroof, special seats or a larger engine) within the products.
3. *Level scheduling*. These problems focus on obtaining level sequences for the production and usage of components.

The mixed-model sequencing problem with workload minimisation (*MMSP-W*) falls in the first category of problems. Two approaches to this problem are: (1) the approach proposed by Scholl et al. (1998) with the objective of minimising work overload, and (2) the approach proposed by Yano and Rachamadugu (1991) which aims to maximise the completed work. The links between the stations on the line are not considered in either approach, and therefore, one must assume that the workstations are arranged in parallel.

The *MMSP-W* consists of sequencing T products, grouped into a set of I product types, of which d_i are of type i ($i = 1,\ldots,|I|$). A unit of product type i ($i = 1,\ldots,|I|$), when entering workstation k ($k = 1,\ldots,|K|$), requires a processing time equal to $p_{i,k}$ for each homogeneous processor (e.g., operator, robot or human–machine

* Corresponding author. Tel.: +34 93 4011703; fax: +34 934016054.
E-mail addresses: joaquin.bautista@nissanchair.com, joaquin.bautista@upc.edu (J. Bautista), alberto.cano-perez@upc.edu (A. Cano), rocio.alfaro@upc.edu (R. Alfaro).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2012.03.024

system) at normal activity, whereas the standard time granted at each station to work on an output unit is the cycle time, c. Sometimes, to complete part of the work on a product unit, that unit may be retained at station k for a time equal to l_k, which is called the time window and is longer than the cycle time ($l_k > c$). When this occurs, the time available for processing the next unit of output is reduced. When it is not possible to complete all of the work required, it is said that an overload is generated. The objective of *MMSP-W* is to maximise the total work completed, which is equivalent to minimising the total overload generated (see Theorem 1 in Bautista and Cano, 2011).

The overload or overuse is measured in units of time and corresponds to the work that cannot be completed on the line at the planned rate or with the level of activity established. To avoid this overload, the following types of actions can be taken:

I. The operator stops the line and completes the remaining work with a reinforcement of human resources (Okamura and Yamashina, 1979; Xiaobo and Ohno, 1997).
II. The operator, without having completed the work at the station, allows the unit to pass to the following stations under the assumption that it will be completed later or off the line. This unfinished work is called by different names: (1) work overload (Bautista and Cano, 2008, 2011; Scholl et al., 1998; Yano and Rachamadugu, 1991), (2) pending work (Bolat, 1997) or, (3) when it requires the intervention of other operators, utility work (Tsai, 1995).
III. The production capacity is increased above the standard through the intervention of reinforcement operators (Bautista and Cano, 2011; Boysen et al., 2011) or by the preprogramming of robotic systems to accelerate the timing of operations when needed.

In this study, we will consider the first two actions for the treatment of overloads.

The *MMSP-W* is an NP-hard problem (Yano and Rachamadugu, 1991), and several alternatives have been proposed to solve it, such as exact algorithms based on the branch-and-bound method (Bolat, 2003), dynamic programming (Bautista and Cano, 2011; Yano and Rachamadugu, 1991), local-search heuristic procedures (Bautista and Cano, 2008; Yano and Bolat, 1989), greedy algorithms based on priority rules (Bolat and Yano, 1992; Bautista and Cano, 2008), meta- and hyper-heuristics (Cano et al., 2010; Scholl et al., 1998) and beam search techniques (Erel et al., 2007). There are also precedents for various multi-criteria problems related to the *MMSP-W* (Aigbedo and Monden, 1997; Ding et al., 2006; Kotani et al., 2004; Rahimi-Vahed and Mirzaei, 2007).

Our research focuses on a case study of the Nissan engine plant in Barcelona. To carry out this study, it was necessary to adapt some models from the literature to be used as a starting point. The adaptation of the models took into account the following considerations:

- The engine line consists of workstations arranged in series and, therefore, must take into account the linkage between each pair of consecutive workstations (Bautista and Cano, 2011).
- An operation can be interrupted at any time between the moment of completion of a work cycle and the moment of completion marked by the time window associated with that cycle (Bautista et al., 2011).
- It is necessary to design product sequences so that the work required at all workstations in a workday is evenly distributed over time to avoid an undesirable overuse of human resources.

Taking into account the above considerations, this paper includes the following components:

(1) The formulation of two models for the *MMSP-W* with workstations arranged in series and homogeneous processors arranged in parallel. For this purpose, the *M3* and *M4* models proposed in Bautista and Cano (2011) are used.
(2) A discussion of whether to regulate the work required, the work completed or the overload at the stations throughout the workday. For this comparison, we study the new problem resulting from the different approaches.
(3) The proposal of different single – and multi-objective models that consider workload regularity for the new problem.
(4) The choice of two models that incorporate the concept of workload regularity. The optimal overload for 225 instances from the literature (Bautista and Cano, 2011) is determined from the two proposed models and the two models from the literature.
(5) An application of the proposed models to the powertrain production line of the Nissan plant in Barcelona, considering a set of examples representing different scenarios for a mixed production of engines.

The rest of the paper is organised as follows. In Section 2, we propose two *MMSP-W* models that minimise the total overload while considering the linkages between workstations in series, a parallel arrangement of homogeneous processors at the stations and the free interruption of operations. Section 3 presents three approaches (work required, work completed and overload) for achieving workload regularity on the production line. Section 4 presents a number of single-objective and multi-objective models that are consistent with the approaches of Section 3; these models incorporate the concept of a homogeneous distribution of the workload time on the line and are linked to the concept of regularity in the production mix, which is desirable in *JIT* and douki seisan manufacturing environments. In Section 5, using four models, a computational experiment is carried out, and the results are presented. Finally, we present the conclusions of this study in Section 6.

2. MMSP-W models with stations in series and parallel processors

For the *MMSP-W* with serial workstations, free interruption of the operations and time adjustments workload, we begin with models *M3* (*M1* extended) and *M4* (*M2* extended) proposed by Bautista and Cano (2011). Moreover, considering the equivalence of the objective functions of *M3* and *M4* (see Theorem 1, Bautista and Cano, 2011), we can combine the models in two ways if we consider the absolute time scales used in *M3* or the relative times scales used in *M4*.

Model $M\_3 \cup 4$ is focused on maximising the total work performed (i.e., minimising the total overload) and uses the absolute start instants of the units of the sequence at each station. The parameters and variables of this model are presented below.

To indiscriminately refer to workers, robots or workers-machines systems, the coefficient b_k represents the number of processors at station k ($k = 1, \ldots, |K|$) and not the number of workers, as in Yano and Rachamadugu (1991).

In this research, the presence of more than one processor at a given workstation does not imply that several substations exist in parallel in the station. In effect,

(1) If a workstation has b assigned homogeneous processors, the total load of the station is equally distributed to each processors, which collaboratively realises its work on a concrete unit of the work in progress *WIP*.

Parameters	
K	Set of workstations ($k = 1,\ldots,\lvert K\rvert$)
b_k	Number of homogeneous processors at workstation k
I	Set of product types ($i = 1,\ldots,\lvert I\rvert$)
d_i	Programmed demand of product type i
$p_{i,k}$	Processing time required by a unit of type i at workstation k for each homogeneous processor (at normal activity)
T	Total demand; obviously, $\sum_{i=1}^{\lvert I\rvert} d_i = T$
t	Position index in the sequence ($t = 1,\ldots,T$)
c	Cycle time, the standard time assigned to workstations to process any product unit
l_k	Time window, the maximum time that each processor at workstation k is allowed to work on any product unit, where $l_k - c > 0$ is the maximum time that the work in progress (*WIP*) is held at workstation k
Variables	
$x_{i,t}$	Binary variable equal to 1 if a product unit i ($i = 1,\ldots,\lvert I\rvert$) is assigned to the position t ($t = 1,\ldots,T$) of the sequence, and to 0 otherwise
$s_{k,t}$	Start instant for the tth unit of the sequence of products at station k ($k = 1,\ldots,\lvert K\rvert$)
$v_{k,t}$	Processing time applied to the tth unit of the product sequence at station k for each homogeneous processor (at normal activity)
$w_{k,t}$	Overload generated for the tth unit of the product sequence at station k for each homogeneous processor (at normal activity); measured in time

(2) If at any point along the line, b stations concur in parallel in the permanent regime, each station will contain b product units and each product unit will be allotted a time equal to b times the cycle ($b \times c$) to execute the assigned operations.

Under these conditions, we can define the following mathematical model:

Model $M\_3 \cup 4$:

$$Max\,V = \sum_{k=1}^{|K|}\left(b_k\sum_{t=1}^{T} v_{k,t}\right) \Longleftrightarrow Min\,W = \sum_{k=1}^{|K|}\left(b_k\sum_{t=1}^{T} w_{k,t}\right) \tag{1}$$

subject to :

$$\sum_{t=1}^{T} x_{i,t} = d_i \quad i = 1,\ldots,|I| \tag{2}$$

$$\sum_{i=1}^{|I|} x_{i,t} = 1 \quad t = 1,\ldots,T \tag{3}$$

$$v_{k,t} + w_{k,t} = \sum_{i=1}^{|I|} p_{i,k}x_{i,t} \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{4}$$

$$s_{k,t} \geqslant s_{k,t-1} + v_{k,t-1} \quad k = 1,\ldots,|K|; \quad t = 2,\ldots,T \tag{5}$$

$$s_{k,t} \geqslant s_{k-1,t} + v_{k-1,t} \quad k = 2,\ldots,|K|; \quad t = 1,\ldots,T \tag{6}$$

$$s_{k,t} + v_{k,t} \leqslant (t+k-2)c + l_k \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{7}$$

$$s_{k,t} \geqslant (t+k-2)c \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{8}$$

$$v_{k,t} \geqslant 0 \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{9}$$

$$w_{k,t} \geqslant 0 \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{10}$$

$$x_{i,t} \in \{0,1\} \quad i = 1,\ldots,|I|; \quad t = 1,\ldots,T \tag{11}$$

In the model, the equivalent objective functions (1) are represented by the total work performed (V) and the total overload (W). Constraint (2) requires that the programmed demand be satisfied. Constraint (3) indicates that only one product unit can be assigned to each position of the sequence. Constraint (4) establishes the relation between the processing times applied to each unit at each workstation and the overload generated in each unit at each workstation. Constraints (5)–(8) constitute the set of possible solutions for the start instants of the operations at the workstations and the processing times applied to the products in the sequence for each processor. Constraints (9) and (10) indicate that the processing times applied to the products and the generated overloads, respectively, are not negative. Finally, constraint (11) requires the assigned variables to be binary.

Model $M\_4 \cup 3$ is focused on minimising the total overload (i.e., maximising the total work performed) and uses relative start instants of the units of the sequence at each station. The additional variables of the model are described below.

Variables	
$\hat{s}_{k,t}$	Positive difference between the start instant and the minimum start instant of the tth operation at station k.

Model $M\_4 \cup 3$:

$$Min\,W = \sum_{k=1}^{|K|}\left(b_k\sum_{t=1}^{T} w_{k,t}\right) \Longleftrightarrow Max\,V = \sum_{k=1}^{|K|}\left(b_k\sum_{t=1}^{T} v_{k,t}\right) \tag{12}$$

subject to :

$$\sum_{i=1}^{T} x_{i,t} = d_i \quad i = 1,\ldots,|I| \tag{13}$$

$$\sum_{i=1}^{|I|} x_{i,t} = 1 \quad t = 1,\ldots,T \tag{14}$$

$$v_{k,t} + w_{k,t} = \sum_{i=1}^{|I|} p_{i,k}x_{i,t} \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{15}$$

$$\hat{s}_{k,t} \geqslant \hat{s}_{k,t-1} + v_{k,t-1} - c \quad k = 1,\ldots,|K|; \quad t = 2,\ldots,T \tag{16}$$

$$\hat{s}_{k,t} \geqslant \hat{s}_{k-1,t} + v_{k-1,t} - c \quad k = 2,\ldots,|K|; \quad t = 1,\ldots,T \tag{17}$$

$$\hat{s}_{k,t} + v_{k,t} \leqslant l_k \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{18}$$

$$\hat{s}_{k,t} \geqslant 0 \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{19}$$

$$v_{k,t} \geqslant 0 \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{20}$$

$$w_{k,t} \geqslant 0 \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{21}$$

$$x_{i,t} \in \{0,1\} \quad i = 1,\ldots,|I|; \quad t = 1,\ldots,T \tag{22}$$

$$\hat{s}_{1,1} = 0 \tag{23}$$

In the model, constraints (16)–(19) constitute the set of relative start instants of the operations at each station and the processing times applied to the products for each processor.

Obviously, models $M\_3 \cup 4$ and $M\_4 \cup 3$ can be unified because both models contain all of the information of the original models from Yano (*M1*) and Scholl (*M2*) relative to the *MMSP-W*. To do this, the following constraint must be added:

$$s_{k,t} - \hat{s}_{k,t} = (t+k-2)c \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{24}$$

3. Workload regularity on the production line

The occurrence of overload concentrations at certain times during the workday may be undesirable for some product sequences.

One way to avoid this occurrence is to program product sequences that achieve a constant workload over time on the production line. The consideration of this goal gives rise to a new problem, one of whose objectives is to minimise variations in the rates of the workload at all workstations during the manufacture of products.

Given this general consideration, we may choose different approaches to avoid excessive overload concentrations, both at each workstation and at specific times of the workday. Examples include the following approaches:

1. Work required. This approach is based on obtaining product sequences that regulate the cumulative time of work required at the workstations in all positions of the product sequence.
2. Completed work. This approach is based on obtaining product sequences for which the accumulated time of completed work is regulated at the workstations in all positions of the product sequence.
3. Work overload. Several alternatives are proposed: (a) levelling the accumulated time corresponding to the overload at the workstations in all positions of the product sequence, (b) minimising the corresponding overload for the most overloaded processor when all units have been manufactured and (c) minimising the overload corresponding to the position of the sequence that has the greatest overload when all units have been manufactured.

3.1. Work required

First, we consider the average time required at the kth workstation to process a product unit, which is the processing time for an ideal unit at workstation k. If $\dot{p}_k$ is the average time, then the ideal work rate for station k ($k = 1,\ldots,|K|$) is determined as follows:

$$\dot{p}_k = \frac{b_k}{T}\sum_{i=1}^{|I|} p_{i,k}\cdot d_i \quad k = 1,\ldots,|K| \tag{25}$$

Consequently, the ideal total work needed to complete t units of output at workstation k is

$$P^*_{k,t} = t\cdot\dot{p}_k \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{26}$$

The actual total work required at the kth workstation to process a total of t product units, of which $X_{i,t} = \sum_{\tau=1}^{t} x_{i,\tau}$ are of type i ($i = 1,\ldots,|I|$), is

$$P_{k,t} = b_k\sum_{i=1}^{|I|} p_{i,k}\cdot X_{i,t} = b_k\sum_{i=1}^{|I|} p_{i,k}\left(\sum_{\tau=1}^{t} x_{i,\tau}\right) \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{27}$$

The difference between the actual work and the ideal work required to process t units of output at station k is

$$\delta_{k,t}(P) = P_{k,t} - P^*_{k,t} \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{28}$$

One way to measure the irregularity of the work required at a set of workstations over the workday is to account for each of the differences and distances defined in (28):

$$\Delta_R(P) = \sum_{t=1}^{T}\sum_{k=1}^{|K|}|\delta_{k,t}(P)|, \quad \Delta_E(P) = \sum_{t=1}^{T}\sqrt{\sum_{k=1}^{|K|}\delta^2_{k,t}(P)},$$
$$\Delta_Q(P) = \sum_{t=1}^{T}\sum_{k=1}^{|K|}\delta^2_{k,t}(P) \tag{29}$$

where $\Delta_R(P)$, $\Delta_E(P)$ and $\Delta_Q(P)$ are the global discrepancies in the work required as measured by the rectangular, Euclidean and quadratic distances, respectively.

3.2. Completed work

Another way to achieve a uniform distribution of the work assigned in a workday over time is to consider the viewpoint of the operators on the line by analysing the work that can be completed at each station as the production of units progresses.

Let us define $V_{k,t}$ as the time required to complete the work at station k when processing a sequence of t product units:

$$V_{k,t} = b_k\sum_{\tau=1}^{t} v_{k,\tau} \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{30}$$

The work that should be completed at workstation k to process t product units is $P^*_{k,t} = t\cdot\dot{p}_k$.

Thus, the difference between the actual and ideal work completed at workstation k when processing a sequence of t units of output is

$$\delta_{k,t}(V) = V_{k,t} - P^*_{k,t} \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{31}$$

As in the previous approach, given a sequence of products, we can measure the global irregularity of the work completed using the following functions:

$$\Delta_R(V) = \sum_{t=1}^{T}\sum_{k=1}^{|K|}|\delta_{k,t}(V)|, \quad \Delta_E(V) = \sum_{t=1}^{T}\sqrt{\sum_{k=1}^{|K|}\delta^2_{k,t}(V)}, \quad \Delta_Q(V) = \sum_{t=1}^{T}\sum_{k=1}^{|K|}\delta^2_{k,t}(V) \tag{32}$$

3.3. Work overload

To use the concept of overload to define the peaks of excessive work during the manufacturing process, there are several alternatives.

The first alternative is to minimise the sum of the differences between the actual overload generated for each partial sequence of units at each workstation and the corresponding ideal overload, which must be zero.

Let us define $W_{k,t}$ as the overload accumulated at workstation k while processing the first t product units:

$$W_{k,t} = b_k\sum_{\tau=1}^{t} w_{k,\tau} = b_k\sum_{\tau=1}^{t}(p_{k,\tau} - v_{k,\tau}) = P_{k,t} - V_{k,t}$$
$$k = 1,\ldots,|K|; \quad t = 1,\ldots,T \tag{33}$$

Thus, we can define the following discrepancy functions:

$$\Delta_R(W) = \sum_{t=1}^{T}\sum_{k=1}^{|K|}W_{k,t}, \quad \Delta_E(W) = \sum_{t=1}^{T}\sqrt{\sum_{k=1}^{|K|}W^2_{k,t}},$$
$$\Delta_Q(W) = \sum_{t=1}^{T}\sum_{k=1}^{|K|}W^2_{k,t} \tag{34}$$

where $\Delta_R(W)$, $\Delta_E(W)$ and $\Delta_Q(W)$ are the global rectangular, Euclidean and quadratic functions, respectively, of the discrepancies in the work overload. Another way to limit the overload concentration at a given workstation is to minimise the overload generated by the most overloaded processor, which we assume to be located at station $m_{\max}$. This indicator is calculated as follows:

$$W_{m_{\max}} = \max_{1\leqslant k\leqslant|K|}\left\{\frac{W_{k,T}}{b_k}\right\} = \max_{1\leqslant k\leqslant|K|}\left\{\sum_{t=1}^{T} w_{k,t}\right\} \tag{35}$$

Finally, it is advisable to avoid having overload peaks among the workstations, which can be caused by a unit of product occupying a specific position in the production flow. These overload peaks increase the risk of stoppages at workstations and, therefore, stoppage of the assembly line. Thus, the objective is to minimise the overload corresponding to the position $t_{\max}$ that causes the greatest overload of all of the workstations:

$$W_{t_{\max}} = \max_{1 \leqslant t \leqslant T}\left\{\sum_{k=1}^{|K|} b_k \cdot w_{k,t}\right\} \tag{36}$$

3.4. Relationships between the approaches to facilitate workload regularity

To study the relationship between the approaches presented above, we will build on the properties derived from maintaining a production mix when manufacturing product units over time.

If $X^*_{i,t}$ is the number of units of product type i, of a total of t units, that should ideally be manufactured to maintain the production mix,

$$X^*_{i,t} = \frac{d_i}{T} \cdot t \quad i = 1,\ldots,|I|; \quad t = 1,\ldots,T \tag{37}$$

The first property of the ideal point $\vec{X}^* = \left(X^*_{1,1},\ldots,X^*_{|I|,T}\right)$ is that it levels the required workload.

Theorem 1. *For point* $\vec{X}^*$, *then* $\delta_{k,t}(P) = P_{k,t} - P^*_{k,t} = 0$ $(k = 1,\ldots,|K|;\ t = 1,\ldots,T)$.

Proof. In this case, we have:

$$P_{k,t} = b_k \sum_{i=1}^{|I|} p_{i,k} \cdot X^*_{i,t} \Longleftrightarrow P_{k,t} = b_k \sum_{i=1}^{|I|} \frac{p_{i,k} \cdot d_i \cdot t}{T} = t \cdot \left(\frac{b_k}{T}\sum_{i=1}^{|I|} p_{i,k} \cdot d_i\right)$$
$$= t \cdot \dot{p}_k = P^*_{k,t}$$

Therefore,

$$P_{k,t} - P^*_{k,t} = \delta_{k,t}(P) = 0 \quad k = 1,\ldots,|K|; \quad t = 1,\ldots,T. \tag{38}$$

Corollary 1. *For point* $\vec{X}^*$, $\Delta_R(P) = \Delta_E(P) = \Delta_Q(P) = 0$. *Therefore, the global rectangular, Euclidean and quadratic discrepancy functions of the required work are optimal.*

The second property of point $\vec{X}^*$ serves to link the global discrepancy functions of the completed work (V) to those corresponding to the work overload (W).

Theorem 2. *For point* $\vec{X}^*$, *the equalities* $\Delta_R(V) = \Delta_R(W)$, $\Delta_E(V) = \Delta_E(W)$ *and* $\Delta_Q(V) = \Delta_Q(W)$ *are satisfied.*

Proof. From Theorem 1, $P_{k,t} = P^*_{k,t} (k = 1,\ldots,|K|;\ t = 1,\ldots,T)$.

We can state that $|\delta_{k,t}(V)| = |P^*_{k,t} - V_{k,t}| = |P_{k,t} - V_{k,t}| = W_{k,t}$ $(k = 1,\ldots,|K|;\ t = 1,\ldots,T)$.

Therefore, $\Delta_R(V) = \Delta_R(W)$.

Additionally, we can state that $\delta^2_{k,t}(V) = \left(P^*_{k,t} - V_{k,t}\right)^2 = (P_{k,t} - V_{k,t})^2 = W^2_{k,t}$ $(k = 1,\ldots,|K|;\ t = 1,\ldots,T)$.

Therefore, $\Delta_E(V) = \Delta_E(W)$ and $\Delta_Q(V) = \Delta_Q(W)$.

Corollary 2. *For point* $\vec{X}^*$, *the levelling approaches based on the completed work and the work overload are equivalent.*

The third property of point $\vec{X}^*$ serves to relate the three approaches to workload levelling.

Theorem 3. *If* $\dot{p}_k \leqslant b_k \cdot c$ $(k = 1,\ldots,|K|)$, *the objective functions* $\Delta_R(W)$, $\Delta_E(W)$, $\Delta_Q(W)$, $W_{m_{\max}}$ *and* $W_{t_{\max}}$ *are optimal and are equal to zero for the point* $\vec{X}^*$.

Proof. The time required for one processor located at workstation k to complete an ideal product unit is $\dot{p}_k/b_k$ $(k = 1,\ldots,|K|)$. The minimum time available for a processor to complete a unit of output at any station is the cycle time c.

Because, according to our hypothesis, $\dot{p}_k \leqslant b_k \cdot c$ $(k = 1,\ldots,|K|)$, then $c \geqslant \dot{p}_k/b_k$ $(k = 1,\ldots,|K|)$. This means that, at any moment and at any workstation, any processor has an amount of time available to make an ideal product unit that is greater than or equal to the time required to complete that unit.

Under such conditions, $w_{k,t} = 0$ and $W_{k,t} = 0$ $(k = 1,\ldots,|K|;\ t = 1,\ldots,T)$, and consequently,

$$\Delta_R(W) = \Delta_E(W) = \Delta_Q(W) = W_{m_{\max}} = W_{t_{\max}} = 0.$$

Corollary 3. *If* $\dot{p}_k \leqslant b_k \cdot c$ $(k = 1,\ldots,|K|)$, *then for point* $\vec{X}^*, V_{k,t} = P_{k,t} = P^*_{k,t}$ $(k = 1,\ldots,|K|;\ t = 1,\ldots,T)$. *That is, for point* $\vec{X}^*$, *the completed work* ($V_{k,t}$) *coincides with the required work* ($P_{k,t}$) *and with the work ideally required* $\left(P^*_{k,t} = t \cdot \dot{p}_k\right)$ *at each workstation and for each position in the sequence.*

4. MMSP-W models for workload regularity on the line

The above approaches allow us to establish new models from the reference models to obtain sequences with regularity for the workload and the overload. There are at least two ways to achieve this:

(1) Add one or more objective functions representing the different approaches to the original models and thereby obtain a set of multi-objective problems from which to select the most suitable in each case.
(2) Add a set of constraints to the original models that favours a homogeneous distribution of the workload time and overload time generated by the production units at all workstations.

In this paper, we will employ the second option. To do this, we rely on the properties derived from maintaining the production mix during the manufacturing of units.

4.1. Bi-objective MMSP-W-workload rate variation models

Let $\Im_W$ be the set of objective functions corresponding to the different approaches to the workload regularity on the production line:

$$\Im_W = \{\Delta_R(P), \Delta_E(P), \Delta_Q(P), \Delta_R(V), \Delta_E(V), \Delta_Q(V), \Delta_R(W), \Delta_E(W), \Delta_Q(W), W_{m_{\max}}, W_{t_{\max}}\} \tag{39}$$

Based on the reference model for *M_3 ∪ 4* for the *MMSP-W*, we can state the following:

- For the *M*_3 ∪ 4_*WLRV* models,

$$(Max\,V \Longleftrightarrow Min\,W) \wedge (Min\,f(f \in \Im_W)) \tag{40}$$

subject to: (2)–(11) from *M*_3 ∪ 4

Similarly, based on the reference model for *M*_4 ∪ 3, we can state the following:

- For the *M*_4 ∪ 3_*WLRV* models,

$$(Min\,W \Longleftrightarrow Max\,V) \wedge (Min\,f(f \in \Im_W)) \tag{41}$$

subject to: (13)–(23) from *M*_4 ∪ 3.

4.2. Bi-objective MMSP-W-product rate variation models

Based on Theorems 1–3 and the conclusions derived from them, we can state that regularity can be achieved for both the required and the completed workload by obtaining sequences whose cumulative actual production $X_{i,t} = \sum_{\tau=1}^{t} x_{i,\tau}$ $(i = 1, \ldots, |I|;\ t = 1, \ldots, T)$ is as close as possible to the ideal cumulative production represented by point $\vec{X}^* = \left(X^*_{1,1}, \ldots, X^*_{|I|,T}\right)$, as defined in (37).

Let us consider the discrepancy between the actual and the ideal cumulative production for each product type i $(i = 1, \ldots, |I|)$ when t $(t = 1, \ldots, T)$ product units have passed through the line:

$$\delta_{i,t}(X) = X_{i,t} - X^*_{i,t} \quad i = 1, \ldots, |I|; \quad t = 1, \ldots, T \tag{42}$$

We can measure the irregularity of the accumulated production of a set of product types during the workday as follows:

$$\Delta_R(X) = \sum_{t=1}^{T}\sum_{i=1}^{|I|}|\delta_{i,t}(X)|, \quad \Delta_E(X) = \sum_{t=1}^{T}\sqrt{\sum_{i=1}^{|I|}\delta^2_{i,t}(X)},$$
$$\Delta_Q(X) = \sum_{t=1}^{T}\sum_{i=1}^{|I|}\delta^2_{i,t}(X) \tag{43}$$

Defining the set of functions $\Im_X = \{\Delta_R(X), \Delta_E(X), \Delta_Q(X)\}$, the resulting bi-objective models are as follows:

- For the $M\_3 \cup 4\_PRV$ models,

$$(Max\,V \Longleftrightarrow Min\,W) \wedge (Min\,f(f \in \Im_X)) \tag{44}$$

subject to: (2)–(11) from $M\_3 \cup 4$.

- For the $M\_4 \cup 3\_PRV$ models,

$$(Min\,W \Longleftrightarrow Max\,V) \wedge (Min\,f(f \in \Im_X)) \tag{45}$$

subject to: (13)–(23) from $M\_4 \cup 3$.

Note that the objective functions of the set $\Im_X$ are useful for maintaining a constant production mix on the line over time. This is a clear objective in the *JIT* (Toyota) and Douki Seisan (Nissan) philosophies, and in fact, there are references to this objective in the literature. Specifically, we refer to the production rate variation (*PRV*) problem, proposed by Miltenburg (1989), and to the output rate variation (*ORV*) problem, proposed by Monden (1983). *ORV* describes the manner of sequencing units (i.e., cars) used at Toyota to maintain a constant consumption of components in mixed product lines over time.

4.3. MMSP-W models with production mix restrictions (pmr)

Again using Theorems 1–3, another way to obtain regularity in both the workload and the overload in a sequence is to limit the values of the variables of cumulative production, $X_{i,t}$ $(i = 1, \ldots, |I|;$ $t = 1, \ldots, T)$, which must be whole integers, to be the integers closest to the ideal values $X^*_{i,t} = d_i \cdot t/T$ which, being rational numbers, may not have a practical application:

$$\left\lfloor \frac{d_i}{T} \cdot t \right\rfloor \leqslant X_{i,t} \leqslant \left\lceil \frac{d_i}{T} \cdot t \right\rceil \quad i = 1, \ldots, |I|; \quad t = 1, \ldots, T \tag{46}$$

Thus, from the reference models $M\_3 \cup 4$ and $M\_4 \cup 3$, we have the following:

- For model $M\_3 \cup 4\_pmr$

$$(Max\,V \Longleftrightarrow Min\,W) \tag{47}$$

subject to: (2)–(11) from $M\_3 \cup 4$

$$\sum_{\tau=1}^{t} x_{i,\tau} \geqslant \left\lfloor t \cdot \frac{d_i}{T} \right\rfloor \quad i = 1, \ldots, |I|; \quad t = 1, \ldots, T \tag{48}$$

$$\sum_{\tau=1}^{t} x_{i,\tau} \leqslant \left\lceil t \cdot \frac{d_i}{T} \right\rceil \quad i = 1, \ldots, |I|; \quad t = 1, \ldots, T \tag{49}$$

- For model $M\_4 \cup 3\_pmr$

$$(Min\,W \Longleftrightarrow Max\,V) \tag{50}$$

subject to: (13)–(23) from $M\_4 \cup 3$.

$$\sum_{\tau=1}^{t} x_{i,\tau} \geqslant \left\lfloor t \cdot \frac{d_i}{T} \right\rfloor \quad i = 1, \ldots, |I|; \quad t = 1, \ldots, T \tag{51}$$

$$\sum_{\tau=1}^{t} x_{i,\tau} \leqslant \left\lceil t \cdot \frac{d_i}{T} \right\rceil \quad i = 1, \ldots, |I|; \quad t = 1, \ldots, T \tag{52}$$

In this paper, we focus on comparing the behaviour of *pmr* models to their respective reference models.

5. Computational experiments

To study the behaviour of the $M\_3 \cup 4$, $M\_4 \cup 3$, $M\_3 \cup 4\_pmr$ and $M\_4 \cup 3\_pmr$ models, we performed two computational experiments using data from 225 instances in the literature and from 23 instances corresponding to different demand plans from the Nissan powertrain plant in Barcelona. In both computational experiments, given two models M and M', a set of functions $\Im_0 = \{W, \Delta_Q(P), \Delta_Q(V), \Delta_Q(W), \Delta_Q(X)\}$ and a set of instances E, we begin with the best solution, as determined by the minimum overload (W), offered by each model for each instance $\varepsilon \in \mathrm{E}$. The solutions for models M and M' are called $S^*_M(\varepsilon)$ and $S^*_{M'}(\varepsilon)$, respectively. Then, we measure the relative percentage deviations (*RPDs*) between the values of the function $f \in \Im_0$ for the solutions $S^*_M(\varepsilon)$ and $S^*_{M'}(\varepsilon)$. We define the following ratios:

$$RPD_1(f,\varepsilon) = \frac{f(S^*_{3\cup4}(\varepsilon)) - f\left(S^*_{3\cup4\_pmr}(\varepsilon)\right)}{f(S^*_{3\cup4}(\varepsilon))} \cdot 100 \quad (f \in \Im_0; \varepsilon \in \mathrm{E}) \tag{53}$$

$$RPD_2(f,\varepsilon) = \frac{f(S^*_{4\cup3}(\varepsilon)) - f\left(S^*_{4\cup3\_pmr}(\varepsilon)\right)}{f(S^*_{4\cup3}(\varepsilon))} \cdot 100 \quad (f \in \Im_0; \varepsilon \in \mathrm{E}) \tag{54}$$

$$RPD_3(f,\varepsilon) = \frac{f(S^*_{3\cup4}(\varepsilon)) - f(S^*_{4\cup3}(\varepsilon))}{f(S^*_{3\cup4}(\varepsilon))} \cdot 100 \quad (f \in \Im_0; \varepsilon \in \mathrm{E}) \tag{55}$$

$$RPD_4(f,\varepsilon) = \frac{f\left(S^*_{3\cup4\_pmr}(\varepsilon)\right) - f(S^*_{4\cup3}(\varepsilon))}{f\left(S^*_{3\cup4\_pmr}(\varepsilon)\right)} \cdot 100 \quad (f \in \Im_0; \varepsilon \in \mathrm{E}) \tag{56}$$

$$RPD_5(f,\varepsilon) = \frac{f(S^*_{3\cup4}(\varepsilon)) - f\left(S^*_{4\cup3\_pmr}(\varepsilon)\right)}{f(S^*_{3\cup4}(\varepsilon))} \cdot 100 \quad (f \in \Im_0; \varepsilon \in \mathrm{E}) \tag{57}$$

$$RPD_6(f,\varepsilon) = \frac{f\left(S^*_{3\cup4\_pmr}(\varepsilon)\right) - f\left(S^*_{4\cup3\_pmr}(\varepsilon)\right)}{f\left(S^*_{3\cup4\_pmr}(\varepsilon)\right)} \cdot 100 \quad (f \in \Im_0; \varepsilon \in \mathrm{E}) \tag{58}$$

5.1. Computational experiments with benchmark instances

To analyse the behaviour of the models when it is possible to obtain an optimal solution within a reasonable time frame, a set of 225 instances was selected from the literature (see Bautista and Cano, 2011, Tables 2 and 3).

These instances are built from 45 demand plans and five process time structures for the units produced at four workstations, representing the different workload conditions at the stations. Each demand plan has four product types and a total demand of 16 units. The demand plans are grouped into five blocks *(B1* to

B5), with each block corresponding to a different production mix. The cycle time for all of the instances considered is $c = 100$.

To obtain optimal solutions for the 225 instances from the four models under study, the Gurobi v4.5.0 solver was used, running on an Apple Macintosh iMac with an Intel Core i7 with 2.93 GHz and 8 GB of RAM and the MAC OS X 10.6.7 operating system.

Table 1 presents the minimum, maximum and average CPU times required by the models to obtain optimal solutions for the 225 instances.

Table 1 shows that the incorporation of restrictions to preserve the production mix in the original models ($M\_3 \cup 4$ and $M\_4 \cup 3$) reduces the average CPU time required to obtain the optimum solution by a factor of five. Additionally, the performance of the two $M\_4 \cup 3$ models, which use a relative measure of the operation starting time ($\hat{s}_{k,t}$), is better than that of the two $M\_3 \cup 4$ models, which use an absolute starting time ($s_{k,t}$), as the former models require only 2/3 of the average time needed by the latter models to obtain optimal solutions.

Regarding the quality of the optimal solutions, Table 2 shows the values of $RPD_1(f)$ and $RPD_2(f)(f \in \Im_0)$, which measure the impact of adding the constraints (46) to models $M\_3 \cup 4$ and $M\_4 \cup 3$, respectively, and the values of $RPD_6(f)(f \in \Im_0)$, through which the behaviour of model $M\_3 \cup 4\_pmr$ can be compared with that of $M\_4 \cup 3\_pmr$.

The most significant results presented in Table 2 are as follows:

- The incorporation of constraints (46) for preserving the production mix in the two reference models ($M\_3 \cup 4$ and $M\_4 \cup 3$) lowers the optimal values for the global overload (W) on human resources (see column RPD_1 for W). These values decline by an average of 1.34%. The differences are greatest in the *E1* process time structure (3.74%) and in the *B3* demand plan block (2.57%).
- For all structures and all blocks, the incorporation of these restrictions in the original models improves the regularity of the work required (see columns RPD_1 and RPD_2 for $\Delta_Q(P)$) and the regularity of production (see columns RPD_1 and RPD_2 for $\Delta_Q(X)$), which is desirable in the context of *JIT* manufacturing. Globally, the regularity improvements are approximately 18% for the work required and exceed 33% for maintaining the production mix.
- Models $M\_3 \cup 4\_pmr$ and $M\_4 \cup 3\_pmr$ are equivalent with respect to the minimum global overload (W) objective function, and the same optimal values are obtained for all instances ($RPD_6(W) = 0$). However, the average CPU time used by $M\_3 \cup 4\_pmr$ is 50% greater than that used by $M\_4 \cup 3\_pmr$ (17.19 s versus 11.78 s).
- On average, $M\_3 \cup 4\_pmr$ offers better results than $M\_4 \cup 3\_pmr$ for the following regularity functions: required work (column RPD_6 for $\Delta_Q(P)$), completed work (column RPD_6 for $\Delta_Q(V)$) and production (column RPD_6 for $\Delta_Q(X)$). However, $M\_4 \cup 3\_pmr$ exceeds $M\_3 \cup 4\_pmr$ in the work overload objective function (column RPD_6 for $\Delta_Q(W)$), both in the global average and in each block average.
- $M\_3 \cup 4\_pmr$ shows the most significant advantage over $M\_4 \cup 3\_pmr$ in the *E4* structure, with better performance in the regularity of required work (column RPD_6 for $\Delta_Q(P)$), the regularity of completed work (column RPD_6 for $\Delta_Q(V)$) and the regularity of production (column RPD_6 for $\Delta_Q(X)$). However, in this same structure, $M\_4 \cup 3\_pmr$ has significantly better overload regularity than $M\_3 \cup 4\_pmr$ (column RPD_6 for $\Delta_Q(W)$).

5.2. Case study of the Nissan powertrain plant in Barcelona

A production line in the engine plant at the Nissan Spanish Industrial Operations (*NSIO*) site located in Barcelona, Spain, was selected as an industrial case study.

This line consists of 21 workstations in series (Chica et al., 2011) and is used to produce nine types of engines ($p_1, \ldots, p_9$) with different characteristics. These engines are grouped into three families: (1) engines for 4 × 4 vehicles (p_1, p_2 and p_3), (2) engines for vans (p_4 and p_5) and (3) engines for commercial trucks of average tonnage (p_6 to p_9).

The operational processing times, $p_{i,k}$, for each product type ($i = 1, \ldots, 9$) and each workstation ($k = 1, \ldots, 21$) used in this experiment vary between 89 s and 185 s (see Table 5 in Bautista and Cano, 2011).

In the experiment, the cycle time c is equal to 175 s. The time window, l_k, is identical for all stations ($k = 1, \ldots, 21$) and is equal to 195 *s*. There is a single processor ($b_k = 1$) at each station ($k = 1, \ldots, 21$) that is composed of a two-operator team and all of

Table 1
Minimum, maximum and average CPU times needed to obtain optimal solutions for the 225 instances given by models $M\_3 \cup 4$, $M\_4 \cup 3$, $M\_3 \cup 4\_pmr$ and $M\_4 \cup 3\_pmr$.

	$M\_3 \cup 4$	$M\_4 \cup 3$	$M\_3 \cup 4\_pmr$	$M\_4 \cup 3\_pmr$
CPU_{min}	0.185	0.038	0.266	0.028
CPU_{max}	2004.24	2224.98	167.87	110.53
$\overline{CPU}$	93.66	59.95	17.19	11.78

Table 2
Values of $RPD_1(f)$, $RPD_2(f)$ and $RPD_6(f)$ $(f \in \Im_0)$, by structures, by blocks and globally, for the 225 instances.

	W	$\Delta_Q(P)$			$\Delta_Q(V)$			$\Delta_Q(W)$			$\Delta_Q(X)$		
	RPD_1	RPD_1	RPD_2	RPD_6	RPD_1	RPD_2	RPD_6	RPD_1	RPD_2	RPD_6	RPD_1	RPD_2	RPD_6
E1	−3.74	21.55	31.27	1.45	−2.75	−3.85	−8.53	4.72	8.39	2.02	48.67	46.89	1.87
E2	−1.25	22.71	21.75	0.06	−0.72	−2.69	−3.16	−3.79	−5.08	0.68	24.10	25.79	0.57
E3	−0.69	12.63	14.95	1.75	2.76	1.93	−2.86	−0.75	−1.46	0.37	38.48	34.24	−1.60
E4	−0.01	17.13	1.65	−23.73	8.86	−13.03	−32.27	−8.18	1.20	6.29	44.53	36.16	−9.85
E5	−1.00	20.32	19.28	−3.43	3.55	1.84	−6.37	−7.44	−5.24	−1.68	27.15	23.10	−3.19
B1	−0.04	10.40	11.70	−6.70	1.66	1.36	−6.22	−1.03	1.73	2.08	26.23	17.45	−1.33
B2	−1.31	15.13	20.77	−5.52	3.49	−3.17	−13.74	−3.90	−0.29	1.95	33.39	25.33	−2.56
B3	−2.57	26.44	27.31	−0.48	1.73	−0.92	−5.99	1.09	−0.10	0.91	41.92	45.49	0.73
B4	−1.84	15.56	20.52	−4.51	1.65	−2.63	−9.87	−0.74	−0.33	1.60	40.02	41.44	−3.04
B5	−1.12	19.55	14.81	−5.58	2.46	−4.66	−12.08	−4.84	−0.96	1.51	36.98	32.90	−3.42
Average	−1.34	18.87	17.78	−4.78	2.34	−3.16	−10.64	−3.09	−0.44	1.53	36.59	33.24	−2.44

Table 3
Values of the overload W and indicators $RPD_1(W)$ to $RPD_6(W)$ obtained from the four models applied to the 23 demand plans of the NISSAN-9ENG set.

	W				W					
	$3 \cup 4$	$3 \cup 4\_pmr$	$4 \cup 3$	$4 \cup 3\_pmr$	RPD_1	RPD_2	RPD_3	RPD_4	RPD_5	RPD_6
1	251	229	187	186	8.76	0.53	25.50	18.34	25.90	18.78
2	444	473	341	383	−6.53	−12.32	23.20	27.91	13.74	19.03
3	477	451	427	423	5.45	0.94	10.48	5.32	11.32	6.21
4	402	419	310	307	−4.23	0.97	22.89	26.01	23.63	26.73
5	754	763	633	661	−1.19	−4.42	16.05	17.04	12.33	13.37
6	525	544	413	478	−3.62	−15.74	21.33	24.08	8.95	12.13
7	818	774	742	731	5.38	1.48	9.29	4.13	10.64	5.56
8	228	338	139	160	−48.25	−15.11	39.04	58.88	29.82	52.66
9	824	817	732	751	0.85	−2.60	11.17	10.40	8.86	8.08
10	1208*	1208*	1208*	1208*	0.00	0.00	0.00	0.00	0.00	0.00
11	165	185	78	122	−12.12	−56.41	52.73	57.84	26.06	34.05
12	406	402	284	287	0.99	−1.06	30.05	29.35	29.31	28.61
13	383	475	286	336	−24.02	−17.48	25.33	39.79	12.27	29.26
14	500	478	420	423	4.40	−0.71	16.00	12.13	15.40	11.51
15	506	562	433	442	−11.07	−2.08	14.43	22.95	12.65	21.35
16	321	296	227	251	7.79	−10.57	29.28	23.31	21.81	15.20
17	550	618	478	488	−12.36	−2.09	13.09	22.65	11.27	21.04
18	673	717	605	619	−6.54	−2.31	10.10	15.62	8.02	13.67
19	949	945*	945*	945*	0.42	0.00	0.42	0.00	0.42	0.00
20	233	259	139	150	−11.16	−7.91	40.34	46.33	35.62	42.08
21	652	628	560	561	3.68	−0.18	14.11	10.83	13.96	10.67
22	1006	1003	987	984	0.30	0.30	1.89	1.60	2.19	1.89
23	188	193	140	121	−2.66	13.57	25.53	27.46	35.64	37.31
Average	–	–	–	–	−4.60	−5.79	19.66	21.83	16.08	18.66

the auxiliary systems and tools necessary to carry out the assigned operations.

We consider a total of 23 sample demand plans, which are characterised by their demands for the nine engine types, each with the same total demand (T) of 270 engines. This production corresponds to a workday with an effective time of 13.125 h, divided equally between the morning and afternoon shifts. The 23 demand plans correlate to different situations that may occur in practice and which affect the production mix of the nine engines (see Table 6, NISSAN-9ENG, in Bautista and Cano, 2011).

To implement the four models, the Gurobi v4.5.0 solver was used on a Apple Macintosh iMac computer with an Intel Core i7 2.93 GHz processor and 8 GB of RAM using MAC OS X 10.6.7. The solutions from this solver were obtained by allowing a CPU time of 7200 s for each model and for each of the 23 demand plans in the NISSAN-9ENG set.

Table 3 shows the results for the work overload (W) obtained in this calculation.

Based on the results shown in Table 3, we can conclude the following:

- Allowing a run time of 7200 s for each model and each instance, we could only guarantee the optimal solution for instances 10 and 19, which required CPU times of 143 s and 827 s, respectively.

Table 4
Values of RPD_1,RPD_2 and RPD_6 for the functions $\Delta_Q(P)$, $\Delta_Q(V)$, $\Delta_Q(W)$ and $\Delta_Q(X)$ for the 23 instances of the NISSAN-9ENG set.

	$\Delta_Q(P)$			$\Delta_Q(V)$			$\Delta_Q(W)$			$\Delta_Q(X)$		
	RPD_1	RPD_2	RPD_6	RPD_1	RPD_2	RPD_6	RPD_1	RPD_2	RPD_6	RPD_1	RPD_2	RPD_6
1	92.66	96.40	0.02	54.41	81.57	42.06	−4.07	7.91	41.12	95.45	97.73	0.00
2	91.73	89.70	3.26	18.68	18.53	25.96	−31.28	−24.02	36.56	94.34	92.19	1.87
3	91.62	89.24	5.66	19.21	14.91	3.67	3.09	11.95	14.66	93.22	91.48	0.76
4	94.85	92.35	−8.59	31.20	42.85	42.79	−15.86	4.97	48.57	96.60	94.59	−3.55
5	87.83	97.67	12.96	1.69	32.93	12.34	−2.57	−6.74	23.66	91.22	98.77	8.63
6	91.37	94.26	−1.33	5.13	22.54	19.77	−10.34	−26.97	23.03	93.11	95.78	−6.88
7	96.47	91.12	20.17	19.19	7.38	0.83	13.63	2.92	9.32	97.68	94.57	4.64
8	91.02	94.25	−1.07	−15.14	73.86	70.56	−96.65	−34.65	74.53	92.70	94.25	26.82
9	84.40	94.61	9.32	10.16	7.85	5.60	3.88	−2.33	12.27	87.63	96.46	3.26
10	81.11	87.04	17.68	2.78	1.66	−3.24	5.70	1.61	−3.17	88.78	89.59	1.87
11	93.07	95.25	7.52	71.35	84.98	38.91	−17.92	−266.70	45.48	94.65	96.06	−1.50
12	93.50	96.26	19.28	30.06	55.05	24.24	−9.13	2.75	48.62	97.17	97.28	2.42
13	91.62	95.59	14.20	−7.23	55.88	52.29	−50.85	−23.60	49.03	94.62	96.95	2.94
14	93.84	94.35	7.81	36.10	33.98	10.15	−4.93	−2.77	16.98	95.22	96.03	6.08
15	91.65	94.58	4.25	19.34	27.38	10.66	−30.10	6.21	33.73	94.58	96.47	0.25
16	89.20	90.42	7.83	49.00	43.14	25.57	0.71	−17.65	31.37	92.32	92.56	5.67
17	83.92	88.91	1.65	−1.83	13.27	15.21	−26.24	−2.72	36.00	89.90	92.34	−0.29
18	92.66	92.08	0.37	9.99	8.43	17.28	−15.50	−3.47	23.11	93.97	94.56	−2.51
19	88.69	94.96	−7.90	5.34	7.96	−4.15	3.37	0.38	−4.29	91.40	95.06	−4.79
20	91.11	96.31	13.23	28.27	79.92	57.79	−26.99	−41.39	67.72	92.12	96.44	8.71
21	92.15	86.39	−3.21	21.02	4.66	6.88	7.95	−3.47	16.78	93.29	88.88	1.30
22	91.93	90.16	20.29	6.57	4.39	−0.49	1.46	−3.96	5.25	96.55	94.27	2.37
23	89.98	86.55	8.16	57.47	48.07	7.34	−12.67	18.67	56.40	38.06	92.44	89.00
Average	90.71	92.54	6.59	20.56	33.53	20.96	−13.71	−17.53	30.73	91.07	94.55	6.39

- The reference model $M\_3 \cup 4$ achieves better results in overload than the corresponding model ($M\_3 \cup 4\_pmr$) with the constraints of maintaining the production mix (a difference of 4.60% in $RPD_1(W)$) when we consider the average of the 23 instances.
- Similarly, the reference model $M\_4 \cup 3$ achieves a better average overload than $M\_4 \cup 3\_pmr$ (a difference of 5.79% in $RPD_2(W)$) on the set of 23 instances.
- The minimal overload results for the $M\_4 \cup 3$ model in the 23 instances are better than or equal to those of the $M\_3 \cup 4$ (see column $RPD_3(W)$) and $M\_3 \cup 4\_pmr$ models (see column $RPD_4(W)$). The average improvement offered by the first model over the other two is 19.66% and 21.83%, respectively.
- Similarly, the performance of the $M\_4 \cup 3\_pmr$ model is better than or equal to that of the $M\_3 \cup 4$ (column RPD_5 (W)) and $M\_3 \cup 4\_pmr$ (column $RPD_6(W)$) models for all instances, with average improvements of 16.08% and 18.66%, respectively.

Table 4 shows the main results for the regularity of work (both required and completed), the overload and the production mix.

In this case, the most significant results are as follows:

- The incorporation of constraints (46) for maintaining the production mix in the two reference models, $M\_3 \cup 4$ and $M\_4 \cup 3$, produces a significant improvement in the following functions for the 23 instances: (1) regularity of the required work (columns RPD_1 and RPD_2 for $\Delta_Q(P)$), with average improvements of 90.71% and 92.54% over the $M\_3 \cup 4$ and $M\_4 \cup 3$ models, respectively, and (2) regularity of production (columns RPD_1 and RPD_2 for $\Delta_Q(X)$), with improvements of 91.07% and 94.55%.
- Regarding the regularity of completed work, $M\_4 \cup 3\_pmr$ is better than $M\_4 \cup 3$ (column RPD_2 for $\Delta_Q(V)$) in the 23 instances, with an average improvement of 33.53%, and $M\_3 \cup 4\_pmr$ improves upon the results of $M\_3 \cup 4$ by an average of 20.56% (column RPD_1 for $\Delta_Q(V)$).
- Logically, due to the increased global overload caused by the constraints (46), the $M\_3 \cup 4$ model achieves better average results for regularity of the work overload than the $M\_3 \cup 4\_pmr$ model, with a difference of 13.71% (column RPD_1 for $\Delta_Q(W)$). Similarly, $M\_4 \cup 3\_pmr$ is worse than $M\_4 \cup 3$, with an average difference of 17.53% (column RPD_2 for $\Delta_Q(W)$).
- Finally, $M\_4 \cup 3\_pmr$ produces better results than $M\_3 \cup 4\_pmr$ for the regularity of completed and required work, regularity of the overload and the regularity of production (columns RPD_6 for $\Delta_Q(P)$, $\Delta_Q(V)$, $\Delta_Q(W)$ and $\Delta_Q(X)$), with average improvements of 6.59%, 20.96%, 30.73% and 6.39%, respectively.

6. Conclusions

We have presented two models for the *MMSP-W*, $M\_3 \cup 4$ and $M\_4 \cup 3$, that produce sequences that minimise the overall overload in mixed-product assembly lines with workstations in series and homogeneous processors in parallel. The sequences include the freedom to interrupt the operations at any time while retaining the work in progress. These models are both an extension and a unification of the *M3* and *M4* models proposed in Bautista and Cano (2011).

In a case study of the Nissan engine plant in Barcelona, we considered the human resources assigned to the production line in the creation of engine sequences to distribute the work required of the operators at all stations during the workday as evenly as possible over time. This consideration, analysed from different approaches, led to the development of various single- and multi-objective models for the problem. Of the models, two were selected, $M\_3 \cup 4\_pmr$ and $M\_4 \cup 3\_pmr$, that incorporate a set of constraints into the original models, $M\_3 \cup 4$ and $M\_4 \cup 3$, respectively, to preserve the production mix while the engines are being manufactured.

These four models were implemented using the Gurobi v4.5.0 solver to find the optimal solutions for 225 examples from the literature. The optima were computed with average CPU times of 93.66 s, 59.95 s, 17.19 s and 11.78 s for the $M\_3 \cup 4$, $M\_4 \cup 3$, $M\_3 \cup 4\_pmr$ and $M\_4 \cup 3\_pmr$ models, respectively. The models using a relative measure of the operation starting time ($M\_4 \cup 3$ and $M\_4 \cup 3\_pmr$) performed better than the two corresponding models ($M\_3 \cup 4$ and $M\_3 \cup 4\_pmr$) using an absolute starting time. Additionally, by incorporating constraints to maintain the production mix in the reference models, $M\_3 \cup 4$ and $M\_4 \cup 3$, the average CPU time required to obtain the optimal solution was reduced by a factor of five. These faster models give results that are, on average, 1.34% worse than the reference models in terms of the overall work overload, but they provide an 18% average improvement in the regularity of required work and an improvement of over 33% in maintaining the production mix. These are both desirable properties in the context of *JIT* manufacturing and douki seisan.

Additionally, we applied the four models to a case study of the Nissan engine plant in Barcelona using 23 sample demand plans with a total demand of 270 engines, corresponding to a production day with two shifts. Each of the four models was allowed a run time of 7200 s for each instance, but the optimum was found for only two examples within that time frame. Regarding the total overload, the following can be observed: (1) $M\_3 \cup 4$ achieves 4.60% better average results than $M\_3 \cup 4\_pmr$; (2) the performance of $M\_4 \cup 3$ exceeds that of $M\_4 \cup 3\_pmr$ by 5.79%; (3) the results of $M\_4 \cup 3$ are better than or equal to those of $M\_3 \cup 4$ (with an average difference of 19.66%) and $M\_3 \cup 4\_pmr$ (with an average difference of 21.83%) for the 23 instances; and (4) the performance of $M\_4 \cup 3\_pmr$ is better than or equal to that of $M\_3 \cup 4$ (by 16.08%) and $M\_3 \cup 4\_pmr$ (by 18.66%) for the 23 instances. Focusing on the regularity functions, we can observe the following: (1) $M\_3 \cup 4\_pmr$ and $M\_4 \cup 3\_pmr$ produce better results than $M\_3 \cup 4$ and $M\_4 \cup 3$ for the 23 instances in regulating the required work (with improvements of 90.71% and 92.54%) and in regulating production (with improvements of 91.07% and 94.55%); (2) when regulating the completed work, $M\_4 \cup 3\_pmr$ performs better than $M\_4 \cup 3$, with an average improvement of 33.53%, and $M\_3 \cup 4\_pmr$ improves upon the average results of $M\_3 \cup 4$ by 20.56%; and (3) $M\_4 \cup 3\_pmr$ performs better than $M\_3 \cup 4\_pmr$ in regulating the completed work, the required work, the overload and the production, with average differences of 6.59%, 20.96%, 30.73% and 6.39%, respectively.

For future work, we propose the following research areas: (1) the design and implementation of other solution methods for the models proposed in this research, (2) the incorporation of realistic alternatives that facilitate the completion of required work for the proposed models and (3) the incorporation of mechanisms to modulate the work required and the work completed by human resources throughout the workday in the proposed models.

Acknowledgment

The authors greatly appreciate the collaboration of Nissan Spanish Industrial Operations (NSIO) as well as the Nissan Chair UPC for partially funding this research. This work was also partially funded by project PROTHIUS-III, DPI2010-16759 including EDRF funding from the Spanish government.

References

Aigbedo, H., & Monden, Y. (1997). A parametric procedure for multcriterion sequence scheduling for just-in-time mixed-model assembly lines. *International Journal of Production Research, 35*, 2543–2564.

Bautista, J., & Cano, J. (2008). Minimizing work overload in mixed-model assembly lines. *International Journal of Production Economics, 112*(1), 177–191.

Bautista, J., & Cano, A. (2011). Solving mixed model sequencing problem in assembly lines with serial workstations with work overload minimisation and interruption rules. *European Journal of Operational Research, 210*(3), 495–513.

Bautista, J., Cano, A., & Alfaro, R. (2011). A bounded dynamic programming algorithm for the MMSP-W considering workstation dependencies and unrestricted interruption of the operations, In *Proceedings (CD) 11th international conference on intelligent systems design and applications (ISDA 2011)*, Córdoba, ISBN: 978-1-4577-1675-1, Spain.

Bolat, A. (1997). Efficient methods for sequencing minimum job sets on mixed model assembly lines. *Naval Research Logistics, 44*(5), 419–437.

Bolat, A. (2003). A mathematical model for sequencing mixed models with due dates. *International Journal of Production Research, 41*(5), 897–918.

Bolat, A., & Yano, C. A. (1992). Scheduling algorithms to minimize utility work at a single station on paced assembly line. *Production Planning and Control, 3*(4), 393–405.

Boysen, N., Fliedner, M., & Scholl, A. (2009). Sequencing mixed-model assembly lines: Survey, classification and model critique. *European Journal of Operational Research, 192*(2), 349–373.

Boysen, N., Kiel, M., & Scholl, A. (2011). Sequencing mixed-model assembly lines to minimise the number of work overload situations. *International Journal of Production Research, 49*(16), 4735–4760.

Cano, J., Ríos, R., & Bautista, J. (2010). A scatter search based hyper-heuristic for sequencing a mixed-model assembly line. *Journal of Heuristics, 16*(6), 749–770.

Chica, M., Cordón, O., Damas, S., & Bautista, J. (2011). Including different kinds of preferences in a multiobjective ant algorithm for time and space assembly line balancing on different Nissan scenarios. *Expert Systems with Applications, 38*(1), 709–720.

Ding, F., Zhu, J., & Sun, H. (2006). Comparing two weighted approaches for sequencing mixed-model assembly lines with multiple objectives. *International Journal of Production Economics, 102*(1), 108–131.

Erel, E., Gocgun, Y., & Sabuncuoglu, I. (2007). Mixed-model assembly line sequencing using beam search. *International Journal of Production Research, 45*(22), 5265–5284.

Kotani, S., Ito, T., & Ohno, K. (2004). Sequencing problem for a mixed-model assembly line in the Toyota production system. *International Journal of Production Research, 42*(23), 4955–4974.

Miltenburg, J. (1989). Scheduling mixed-model assembly lines for just-in-time production systems. *Management Science, 35*(2), 192–207.

Monden, Y. (1983). Toyota production system. Industrial engineering and management press, Institute of industrial engineers, Atlanta, GA.

Okamura, K., & Yamashina, H. (1979). A heuristic algorithm for the assembly line model-mix sequencing problem to minimize the risk of stopping the conveyor. *International Journal of Production Research, 17*(3), 233–247.

Rahimi-Vahed, A., & Mirzaei, A. H. (2007). A hybrid multi-objective shuffled frog-leaping algorithms for a mixed-model assembly line sequencing problem. *Computers and Industrial Engineering, 53*(4), 642–666.

Scholl, A., Klein, R., & Domschke, W. (1998). Pattern based vocabulary building for effectively sequencing mixed-model assembly lines. *Journal of Heuristics, 4*(4), 359–381.

Tsai, L. H. (1995). Mixed-model sequencing to minimize utility work and the risk of conveyor stoppage. *Management Science, 41*(3), 485–495.

Xiaobo, Z., & Ohno, K. (1997). Algorithms for sequencing mixed models on assembly line in a JIT production system. *Computers and Industrial Engineering, 31*(1), 47–56.

Yano, C. A., & Bolat, A. (1989). Survey, development, and application of algorithms for sequencing paced assembly lines. *Journal of Manufacturing and Operations Management, 2*(3), 172–198.

Yano, C. A., & Rachamadugu, R. (1991). Sequencing to minimize work overload in assembly lines with product options. *Management Science, 37*(5), 572–586.