PERGAMON

Computers & Operations Research 27 (2000) 1325–1345

computers & operations research

www.elsevier.com/locate/dsw

New lot-sizing formulations for less nervous production schedules

Osman Kazan, Rakesh Nagi, Christopher M. Rump*

Department of Industrial Engineering, 342 Bell Hall, State University of New York at Buffalo, Buffalo, NY 14260, USA

Received 1 January 1999; received in revised form 1 June 1999

Abstract

Previously scheduled production plans frequently need to be updated because of demand uncertainty. After making a comprehensive definition of nervousness which includes costs for changes in production schedule and quantity, we suggest three methodologies. Two methods are modified versions of very well-known methods: the Wagner–Whitin algorithm and the Silver–Meal heuristic. However, our definition of nervousness and its consequences for altering predetermined production volumes make the well-known property of producing either zero or a sum of several periods' demand suboptimal. Therefore a third method, a new mixed integer linear programming formulation, is proposed which is shown to be more effective in some cases. Numerical analyses are carried out for a wide range of possible cases, through which we provide insights to the most appropriate algorithm in a parameterized space.

Scope and purpose

Uncertainty in demand forecasts and a rolling horizon create volatility in lot-sizing results. This volatility is characterized by frequent changes in predetermined production schedules and is highly undesirable for production managers. It causes nervousness in the system in terms of canceling existing setups, introducing new setups, and altering the production volumes. In this paper, we propose new cost structures for these changes, and offer several models that identify less nervous production schedules in a rolling horizon basis. For practitioners, this work identifies the most preferable algorithm for a variety of system parameters. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Production; Inventory; Lot sizing; Material requirements planning; Nervousness

* Corresponding author. Tel.: + 716-645-2357; fax: + 716-645-3302.

E-mail address: crump@eng.buffalo.edu (C.M. Rump)

PII: S0305-0548(99)00076-3

1. Introduction

Manufacturing companies determine their production schedules based on the forecasts for future demand. It is a commonly recognized fact that accurate forecasts are generally available over a given horizon for only a few initial periods. For the rest of the periods such figures become progressively blurred. Yet companies have to determine a production policy that is the most robust to any kind of such uncertainties. Due to new information, previous schedules may need to be updated regularly. Because of the dynamic environment, updated schedules may be quite different than previous ones. These differences may cause the following changes in the schedule: assigning new production setups for some periods, calling off some previously scheduled setups for some other periods, and altering the production volumes of previously scheduled setups. Such changes in schedules are referred to as *nervousness*. It is quite reasonable to anticipate that the changes mentioned above would introduce some costs. In the following, we provide a review of the developments in single-level, uncapacitated lot-sizing and in particular those that have considered some form of nervousness in a rolling horizon.

Wagner and Whitin [1] first proposed an optimal algorithm to solve the single item, single-level, uncapacitated economic lot size problem. In their model, demand figures for future periods were assumed to be deterministic. The algorithm is based upon three theorems that give some important clues about the structure of optimal solutions:

1. Initial inventory can always be assigned to zero.
2. At optimality, a production volume is either zero or a sum of demands for several periods.
3. A setup results in a production quantity that satisfies all demand until the next production setup.

The last condition of optimality encouraged researchers to suggest several simpler heuristic methods. The Silver–Meal heuristic [2], in particular, tries to identify the production setup points by including demand figures one by one in the order. Such a straightforward approach constitutes myopic behavior. Nevertheless, its effectiveness is observed to be as attractive as its simplicity.

Steele [3] and Mather [4] approached the nervousness problem from a managerial point of view. The causes of nervousness were listed as: master production schedule (MPS) changes, unexpected changes in previously made customer orders, parameter (lead time, safety stock, etc.) changes, forecast changes, vendor plant fall-down, scrap and spoilage, engineering changes, record errors, and unplanned transactions.

Although optimal for a single horizon, the Wagner–Whitin algorithm is not optimal in a rolling horizon environment. Despite that, Baker [5] showed that rolling Wagner–Whitin schedules produce effective results that are very close to the optimal solutions when demand is certain and there are no costs for nervousness. However, when nervousness costs are considered, these schedules may be less attractive.

Carlson et al. [6] defined nervousness as the difficulty encountered in shifting of previously scheduled production setups because of new information obtained in a rolling horizon. They introduced a schedule change cost (SCC), which consists of the cost of scheduling a *new setup*. After the introduction of SCC, the model obtained can easily be converted into the one proposed by Wagner and Whitin, i.e., both models are structurally identical. For this reason, the dynamic programming solution methodology proposed by Wagner and Whitin is applicable.

Based upon the exact schedule change cost definition, Kropp et al. [7] established two modified versions of the Silver–Meal heuristic and one modified Part-period balancing heuristic as well. For the first time, different methodologies including the one proposed by Carlson et al. [6] were tested and compared under a rolling horizon. Numerical analyses showed that the modified Silver–Meal approach was only slightly more costly than the modified Wagner–Whitin algorithm.

De Bolt and Van Wassenhove [8] illustrated that the cost figure would increase due to demand uncertainty in a dynamic rolling-schedule environment. Their work is one of the first to insert forecast errors into the material requirement planning (MRP) lot-sizing research. They also suggested buffering against the forecast errors. The simulation analysis conducted showed that the Silver–Meal heuristic with buffering against forecast errors might generate good solutions.

Blackburn et al. [9] examined the effectiveness of alternative strategies in multi-level production processes. A series of simulation experiments was conducted to test the effectiveness of the strategies. According to the findings, cost for schedule changes and freezing the schedule within the horizon are significantly effective on schedule stability. Their work encouraged researchers to focus on those aspects of MRP system nervousness.

Ho and Carter [10] analyzed several other dampening techniques; static, dynamic, and cost-based procedures. The cost-based dampening utilized the exact definition of the schedule change cost suggested by Carlson et al. [6]. They claimed that a proper dampening procedure together with a lot-sizing rule may result in system improvement.

Aull and LaForge [11] investigated allowable limits on predetermined production figures without changing the timing of setups established by the previous schedule. This is a different approach than that of Carlson et al. [6] in which the question was the setup points. Their approach motivated us to give more emphasis on consequences of altering production volume. The Incremental Part-Period Algorithm was used for this purpose.

Recent studies have focused on other detailed aspects of MRP system nervousness: stochastic demand [12,13], supply and process uncertainty [14], forecast error distribution [15], and detecting minimal forecast window [16].

Here in our work, we question the classical definition of the nervousness in MRP systems. As researchers have previously stated, system performance strongly depends on such a definition. Thus, incorporating a more complete definition of nervousness costs, we examine the performance of several algorithms in a rolling horizon.

In the following section, after presenting the definitions of the problem arguments, we provide modified versions of the Wagner–Whitin algorithm and the Silver–Meal heuristic, and a mixed integer linear programming method. They are all designed to be employed in a rolling schedule environment. The experimental design to test the effectiveness of those methodologies is discussed in Section 3. Finally, the paper concludes with the discussion of the results.

2. Problem environment

This section contains mathematical formulations of the models that are believed to produce less nervous production schedules. Through the rest of this paper the terms *forecast window* and *production window* (or *horizon*) are used interchangeably as are the terms *schedule change cost* and *nervousness cost*.

2.1. Definitions

Through the following definitions, index i stands for production periods, whereas index k stands for the order of a period within a production window. Thus, k may have the values between 1 and the length of forecast window.

N length of forecast window
$\mathcal{N}_i$ set of new setup points offered by schedule i that were not scheduled in schedule $i-1$
$\mathcal{O}_i$ set of setup points cancelled from schedule $i-1$ by schedule i
$\mathcal{A}_i$ set of periods where setup decision is unaltered by schedule i
$d_{i,k}$ demand forecast for the kth period at the beginning of horizon i
$x_{i,k}$ production volume suggested by schedule i for the kth period of the horizon; so $x_{i-1,k+1}$ refers to production amount suggested by the previous schedule for the same period
$\Delta^+_{i,k}$ increase in the production volume suggested by schedule i for the kth period ($k \neq N$); clearly, $\Delta^+_{i,k} = x_{i,k} - x_{i-1,k+1}$ if $x_{i,k} > x_{i-1,k+1}$, 0 otherwise
$\Delta^-_{i,k}$ decrease in the production volume suggested by schedule i for the kth period ($k \neq N$); clearly, $\Delta^-_{i,k} = x_{i-1,k+1} - x_{i,k}$ if $x_{i,k} < x_{i-1,k+1}$, 0 otherwise
s_k setup cost at the kth period
h_k holding cost at the kth period
n_k cost of assigning a new setup to the kth period; Logically, $n_1 \geqslant n_2 \geqslant \cdots \geqslant n_N$ can be assumed. That also holds for the following costs
o_k cost of canceling a setup that was previously scheduled to the kth period
a^+_k cost of increasing production volume by 1 unit at the kth period of the horizon
a^-_k cost of decreasing production volume by 1 unit at the kth period of the horizon
$I_{i,k}$ ending inventory at the kth period in schedule i
C^{S}_i total setup cost of schedule i
C^{H}_i total holding cost of schedule i
C^{N}_i total nervousness cost of schedule i
C_i total cost of schedule i

No cost of increasing/decreasing is incurred if the change in production volume is a result of assigning a new setup or canceling a setup.

2.2. Dynamic programming: a modified version of Wagner–Whitin algorithm

Wagner and Whitin proposed the well-known algorithm based on the following cost structure:

$$C_i = C^{\mathrm{S}}_i + C^{\mathrm{H}}_i, \quad \forall i, \tag{1}$$

where

$$C^{\mathrm{S}}_i = \sum_{k=1}^{N} s_k \delta(x_{i,k}), \quad \forall i, \tag{2}$$

$$C^{\mathrm{H}}_i = \sum_{k=1}^{N} h_k I_{i,k}, \quad \forall i \tag{3}$$

with $\delta(x_{i,k})$ being an indicator function defined as

$$\delta(x_{i,k}) = \begin{cases} 1 & \text{if } x_{i,k} > 0, \\ 0 & \text{if } x_{i,k} \leqslant 0. \end{cases}$$

It can easily be seen that Eq. (1) does not capture the nervousness concept because the algorithm was designed for finding the optimal schedule only for a single horizon. Carlson et al. [6] introduced nervousness into the system for the first time. Here, we claim that schedule change cost that they offered cannot wholly represent the nervousness in a manufacturing environment because not only assigning new setups, but also canceling, or even changing production volumes may incur some costs to companies. For this reason, a more comprehensive model must include the costs of canceling a previously scheduled setup (o_i) and altering production volume (a_i^+, a_i^-). For a *single* planning horizon such a cost function can be written as

$$C_i = C_i^{\mathrm{S}} + C_i^{\mathrm{H}} + C_i^{\mathrm{N}}, \quad \forall i, \tag{4}$$

where

$$C_i^{\mathrm{N}} = \sum_{k \in \mathcal{N}_i} n_k + \sum_{k \in \mathcal{O}_i} o_k + \sum_{k \in \mathcal{A}_i} (a_k^+ \Delta_{i,k}^+ + a_k^- \Delta_{i,k}^-), \quad \forall i. \tag{5}$$

Our next task is to prove that total cost of a schedule for a *single* horizon, say i, is structurally identical to the one used by Wagner and Whitin.

Theorem. *Cost of a schedule for a single horizon, C_i, is structurally identical to Eq.* (1).

Proof. When a production schedule is generated at the beginning of period i, the following relationships will already be established:

$$k \in \mathcal{N}_i \Rightarrow \delta(x_{i,k}) = 1,\ \delta(x_{i-1,k+1}) = 0,$$

$$k \in \mathcal{O}_i \Rightarrow \delta(x_{i,k}) = 0,\ \delta(x_{i-1,k+1}) = 1,$$

$$k \in \mathcal{A}_i \Rightarrow \delta(x_{i,k}) = \delta(x_{i-1,k+1}).$$

Eq. (4) can be rewritten as

$$\begin{aligned} C_i &= C_i^{\mathrm{S}} + C_i^{\mathrm{H}} + \sum_{k \in \mathcal{N}_i} n_k + \sum_{k \in \mathcal{O}_i} o_k + \sum_{k \in \mathcal{A}_i} (a_k^+ \Delta_{i,k}^+ + a_k^- \Delta_{i,k}^-) \\ &= \sum_{k=1}^{N} s_k \delta(x_{i,k}) + C_i^{\mathrm{H}} + \sum_{k \in \mathcal{N}_i} n_k \delta(x_{i,k}) + \sum_{k \in \mathcal{O}_i} o_k \delta(x_{i-1,k+1}) + \sum_{k \in \mathcal{A}_i} (a_k^+ \Delta_{i,k}^+ + a_k^- \Delta_{i,k}^-) \delta(x_{i,k}) \\ &= \sum_{k \in \mathcal{N}_i} s_k \delta(x_{i,k}) + \sum_{k \in \mathcal{A}_i} s_k \delta(x_{i,k}) + C_i^{\mathrm{H}} + \sum_{k \in \mathcal{N}_i} n_k \delta(x_{i,k}) + \sum_{k \in \mathcal{O}_i} o_k \delta(x_{i-1,k+1}) \\ &\quad + \sum_{k \in \mathcal{A}_i} (a_k^+ \Delta_{i,k}^+ + a_k^- \Delta_{i,k}^-) \delta(x_{i,k}) \\ &= \sum_{k \in \mathcal{N}_i} (s_k + n_k) \delta(x_{i,k}) + \sum_{k \in \mathcal{O}_i} o_k \delta(x_{i-1,k+1}) + \sum_{k \in \mathcal{A}_i} (s_k + a_k^+ \Delta_{i,k}^+ + a_k^- \Delta_{i,k}^-) \delta(x_{i,k}) + C_i^{\mathrm{H}}. \end{aligned}$$

Let

$$\bar{x}_{i,k} = \begin{cases} x_{i,k} & \text{if } k \in (\mathcal{N}_i \cup \mathcal{A}_i), \\ x_{i-1,k+1} & \text{if } k \in \mathcal{O}_i \end{cases}$$

and

$$\bar{s}_k = \begin{cases} s_k + n_k & \text{if } k \in \mathcal{N}_i, \\ o_k & \text{if } k \in \mathcal{O}_i, \\ s_k + a_k^+ \Delta_{i,k}^+ + a_k^- \Delta_{i,k}^- & \text{if } k \in \mathcal{A}_i \end{cases}$$

yielding

$$C_i = \sum_{k=1}^{N} \bar{s}_k \delta(\bar{x}_{i,k}) + \sum_{k=1}^{N} h_k I_{i,k},$$

which is structurally equivalent to Eq. (1). □

A direct conclusion from the above theorem is that the Wagner–Whitin algorithm can be used to find a new schedule based on a previous one. In order to perform such a task, first $\bar{x}_{i,k}$'s and $\bar{s}_k$'s must be computed according to the relationships given above. Since it is quite straightforward – running time of $O(N)$ – there will not be any change in the complexity of the algorithm.

Recursive Expression of the Dynamic Program. The following backward recursive expression may be utilized to find a new schedule at the beginning of period i:

$$\bar{C}_{N+1} = 0,$$

$$\bar{C}_j = \min_{m: 1 \leqslant j < m \leqslant N+1} \left\{ s_j + \sum_{k=j}^{m-2} h_k I_{i,k} + (1-\delta_j) n_j + \sum_{k=j+1}^{m-1} \delta_k o_k + \delta_j (a_j^+ \Delta_{i,j}^+ + a_j^- \Delta_{i,j}^-) + \bar{C}_m \right\}$$

where $j = N, \ldots, 1$, and $\delta_j = \delta(x_{i-1,j+1})$.

At the end, $\bar{C}_1$ will be the last term obtained from the recursion. The corresponding production sequence will be the schedule offered by the algorithm and cost of the new schedule will be $\bar{C}_1$ ($C_i = \bar{C}_1$). Since it is in the exact structure of the original algorithm, its running time will be $O(N^2)$.

Now the question arises: for a fixed planning horizon, is the schedule offered by the modified Wagner–Whitin algorithm optimal? The answer to this question can be investigated by means of the three conditions of optimality provided in the introduction. One of the most essential conditions is that a production batch must equal the demand for an integral number of periods. Evidently, this condition of optimality does not hold for our case since altering the production volume is not a binary decision. We also have to decide the change in production volume. The dynamic programming approach implicitly enumerates possible production volumes that are equal to demand for an integral number of future periods. If an optimal production batch for a period does not equal one of those possible values, then the recursive approach will simply skip that figure. The result will be a suboptimal schedule. This approach establishes the following theorem.

Theorem. *The schedule identified by the modified Wagner–Whitin algorithm cannot guarantee optimality for a fixed single planning horizon.*

One of the most ironic facts about the lot-sizing problem is that a simple heuristic, such as the one by Silver and Meal, can often outperform an optimal approach, such as the Wagner–Whitin algorithm, in a rolling horizon. For this reason, the optimal schedule for a *single* horizon may not be so attractive. In the following section, we present a modified version of the Silver–Meal heuristic for the nervousness case as has been defined in this work. The reason behind this choice is its myopic structure which may offer interesting production schedules.

2.3. Modified Silver–Meal heuristic

We modify the Silver–Meal heuristic for nervousness as follows: Let the first setup covers m periods' production where $1 \leqslant m \leqslant N$ is determined by the following conditions:

$$\frac{s_1 + H_k + N_1 + O_k + A_k + L_{k+1}}{k} \leqslant \frac{s_1 + H_{k-1} + N_1 + O_{k-1} + A_{k-1} + L_k}{k-1}, \quad 2 \leqslant k \leqslant m,$$

$$\frac{s_1 + H_{m+1} + N_1 + O_{m+1} + A_{m+1} + L_{m+2}}{m+1} > \frac{s_1 + H_m + N_1 + O_m + A_m + L_{m+1}}{m},$$

where

$$H_k = \sum_{j=1}^{k-1} h_j I_{i,j},$$

$$N_1 = (1 - \delta(x_{i-1,2}))n_1,$$

$$O_k = \sum_{j=2}^{k} \delta(x_{i-1,j+1}) o_j,$$

$$A_k = \delta(x_{i-1,2})(\Delta_{i,1}^{+} a_1^{+} + \Delta_{i,1}^{-} a_1^{-}),$$

$$L_{k+1} = (1 - \delta(x_{i-1,k+2}))n_{k+1}.$$

If no such $m < N$ exists the complete horizon is covered by just one setup. If $m < N$ exists, evidently $x_{i,m+1} > 0$. The same procedure is now applied to the remaining problem over periods $m+1, m+2, \ldots, N$. Period $m+1$ will be assumed to be period 1 in the next iteration. Note the difference between N_1 and A_k. We always have to pay N_1 regardless of the production volume. However, A_k is determined by the change in the production volume. Although the expression does not have a k term, $\Delta_{i,1}^{+}$ or $\Delta_{i,1}^{-}$ will certainly depend on k. The procedure to identify the setup point is identical to the heuristic. The only difference is the expressions for total cost per period. Although the new expressions seem a bit more complex than the original ones, they do not increase computational complexity, because it is straightforward to compute $N_1 + A_k + O_k + L_{k+1}$ when a previous schedule is on hand. Here, L_{k+1} was introduced into the model to capture the status of the period that would be the next candidate production point. That made the heuristic more effective, though it would not be so suitable to view the expression as a "*total cost per period*" anymore.

As in the modified Wagner–Whitin algorithm case, the modified version of the heuristic cannot guarantee the optimality for a single horizon. To remedy this we present a mixed integer linear programming (MILP) model, that would provide us optimality for a single horizon in the following section.

2.4. Mixed integer linear programming model

The following MILP provides an optimal new schedule for a single horizon at the beginning of period i based on a previous known schedule. In order to represent the status of a period in the previous schedule, $\bar{\delta}_k$ is introduced through the expression. So $\bar{\delta}_k = \delta(x_{i-1,k+1})$ always holds. It can easily be inferred that $\bar{\delta}_k$'s are known binary constants ($\bar{\delta}_k \in \{0, 1\}$) at the beginning of the new planning horizon:

$$\text{minimize:} \quad \sum_{k=1}^{N} s_k y_k + \sum_{k=1}^{N} h_k I_{i,k} + \sum_{k=1}^{N} (1 - \bar{\delta}_k) n_k y_k + \sum_{k=1}^{N} \bar{\delta}_k o_k (1 - y_k) + \sum_{k=1}^{N} (a_k^+ \Delta_{i,k}^+ + a_k^- \Delta_{i,k}^-)$$

$$\text{subject to:} \quad I_{i,k-1} + x_{i,k} - I_{i,k} = d_{i,k}, \quad k = 1, 2, \ldots, N, \tag{6}$$

$$I_{i,0} = 0, \quad I_{i,N} = 0, \tag{7}$$

$$M_i y_k \geqslant x_{i,k}, \quad k = 1, 2, \ldots, N, \tag{8}$$

$$\Delta_{i,k}^+ \geqslant (x_{i,k} - x_{i-1,k+1}) - M_i(2 - y_k - \bar{\delta}_k), \quad k = 1, 2, \ldots, N, \tag{9}$$

$$\Delta_{i,k}^- \geqslant (x_{i-1,k+1} - x_{i,k}) - M_i(2 - y_k - \bar{\delta}_k), \quad k = 1, 2, \ldots, N, \tag{10}$$

$$x_{i,k} \geqslant 0, \quad \Delta_{i,k}^- \geqslant 0, \quad \Delta_{i,k}^+ \geqslant 0, \quad k = 1, 2, \ldots, N, \tag{11}$$

$$y_k \in \{0, 1\}, \quad k = 1, 2, \ldots, N. \tag{12}$$

Notice that there are N binary decision variables y_k, $k = 1, \ldots, N$, which are the same as $\delta(x_{i,k})$'s, so

$$y_k = \begin{cases} 1 & \text{if a production setup is assigned at the } k\text{th period,} \\ 0 & \text{if not.} \end{cases}$$

The constant M_i is a large number; $M_i = \sum_{k=1}^{N} d_{i,k}$ will be sufficient. Constraints (9) and (10) are constructed to assign suitable values of production change. Through the use of M_i, these constraints ensure that no cost of increasing or decreasing production volume is incurred if the change in production volume is a result of assigning a new setup or canceling a setup. Note that if $\Delta^{-}_{i,k} > 0$ then $\Delta^{+}_{i,k} = 0$ will hold, and vice versa.

Although the MILP contains several binary variables, it is not prohibitive to solve with a standard solver. The model's size depends only on N, the forecast horizon, which is unlikely to be larger than 20 periods.

3. Numerical analysis

In order to identify the most effective procedure for the lot-sizing problem in the long run, several numerical experiments were conducted. This study compared the conventional and nervousness costs generated by several production scheduling algorithms. This comparison was conducted for a variety of demand patterns and demand forecast error patterns with the intent of examining both horizon length effects and forecast error effects.

3.1. The effects

It is assumed that during the process we have a fixed horizon length. That is to say, once length of the forecast window is fixed, it will not be changed until the end of the experiment. The following values are chosen for the horizon length 4, 6, 8, 10, 12, 14, and 16. The length of an entire experiment is 1040 periods. For each case, a single experiment is executed for six different replications of data files generated from different seed numbers.

The horizon effect. In this effect, it was assumed that the forecast figures within the window are not subject to change. Once a forecast figure is set up for a particular period, it never changes until the end of the experiment. Such a forecasting process is perfectly stable. The only uncertainty involved is due to the rolling horizon when a new demand figure for the very last period of the horizon appears. This new information may cause schedule changes that cause nervousness. In fact, this is the reason why the term "*horizon effect*" is used for this case.

The forecast effect. Contrary to the previous effect, here all demand figures within the window are subject to change. In such a case, it would not be difficult to guess that the models would be less successful to identify solutions close to optimality. Since every single forecast figure is subject to change, we may be forced to reschedule the previous schedules more often. Obviously, this will result in much higher nervousness costs. The details of demand figure generation will be presented later.

3.2. The models

The Wagner and Whitin model (WW). The original algorithm is coded without any nervousness. It is aimed to observe differences between a method that completely ignores the nervousness concept and others that do not. When a new production schedule is derived, the costs of nervousness are not taken into account. On the other hand, they are computed subsequently for comparison purposes.

The Carlson, Jucker, and Kropp model (CJK). As mentioned in the introduction, Carlson et al. [6] proposed a dynamic programming model that included the cost of assigning a new production setup as nervousness. In order to observe the efficiency of the model and to create a basis for comparison, it is included in this experimental study. As in the Wagner and Whitin case, after determining the best schedule according to the algorithm, all components of the nervousness are calculated subsequently.

The modified Wagner and Whitin model (MWW). This is the model that is proposed in Section 2.2 of this paper. In this case, all components of the nervousness concept are taken into account in the decision-making step.

The modified Silver and Meal heuristic (MSM). This model is included in the study in order to observe the performance of a heuristic. Notice that, previous models (WW, CJK, and MWW) have a running time of $O(N^2)$. Theoretically, this version of the heuristic runs in $O(N)$ time.

The mixed integer linear programming model (MILP). This is the only model that guarantees to provide optimal production schedules for a single production window. It contains an integer variable for each period of the production window. For our experimental analysis, there would be at most 16 integer variables. It will obviously be the most expensive in CPU usage. LINDO was chosen as the MILP solver.

3.3. The demand distributions

The following demand distributions are considered: U1: Uniform(0, 40), U2: Uniform(20, 40), N1: Normal(20, 6), N2: Normal(30, 12), B1: Uniform(20, 60) w.p. 0.6, or 0 w.p. 0.4, and B2: Uniform(45, 6) w.p. 0.4, or 0 w.p. 0.6. The bimodal distributions (B1 and B2) model less frequent demand patterns with some intermittent periods having no demand. All of the above distributions are used for the horizon effect case. However, for the forecast error case, only normal distributions (N1 and N2) are used. That is because it is more complicated to generate demand figures for the forecast effect case. N1 and N2 are termed as *external* distributions. Details are discussed in the following section.

3.4. Demand generation in forecast effect

Whenever a production period enters into the production window for the very first time, a forecast figure is generated from the external distribution (either N1 or N2). Subsequently for the next decision period, a new forecast figure for that production period is generated from a normal distribution that has a mean as the previous forecast figure, and a standard deviation that is specified by the period index. This distribution is termed as the *internal* distribution. To illustrate the situation, let us suppose that the production window is equal to 4, and that N1 is selected as the external distribution. Further suppose that 24, 10, 18, and 15 are the previous forecast figures. New forecast figures and the demand encountered in the next decision period are summarized in Table 1. The series of $\sigma_1, \sigma_2, \ldots, \sigma_{N-1}$ constitutes a standard deviation pattern. Three of such patterns were established. Table 2 presents these patterns. It may seem that the most interesting standard deviation pattern is the decreasing one, because such a pattern represents the case in which the imminent future has more variation than the distant future. This pattern may be relevant in a highly volatile and hard to forecast situation where the planners do not expend much effort revising distant forecasts. In such a situation, demand figures for only the imminent periods will tend to deviate.

Table 1
Demand generation in the forecast effect

Period	1	2	3	4	5
Previous forecast	24	10	18	15	N/A
New forecast	$\boxed{24}$	N(10, σ_1)	N(18, σ_2)	N(15, σ_3)	N(20, 6)
	Demand	(*Internal*)	(*Internal*)	(*Internal*)	(*External*)

Table 2
Standard deviation patterns

	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7	σ_8	σ_9	σ_{10}	σ_{11}	σ_{12}	σ_{13}	σ_{14}	σ_{15}
Increasing	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Constant	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
Decreasing	7	6.5	6	5.5	5	4.5	4	3.5	3	2.5	2	1.5	1	0.5	0

Table 3
An example of nervousness costs

Period (k)	1	2	3	4	5	6	7	8	9	10	11	12
New (n_k)	20	18	16	14	12	10	8	6	4	2	0	0
Canceled (o_k)	10	9	8	7	6	5	4	3	2	1	0	0
Altered (a_k)	0.67	0.60	0.53	0.47	0.40	0.33	0.27	0.20	0.13	0.07	0	0

3.5. Cost structures

Conventional holding cost (h) was fixed at 1. Relative to this cost, setup costs (S) were varied as 20, 40, 60, and 80. For nervousness costs, the following relationships were established: $n_k = S(11-k)/20$ for $k = 1, 2, \ldots, 10$, or 0 otherwise, $o_k = n_k/2$ for any k, and $a_k^+ = a_k^- = n_k/30$ for any k. These formulas represent a linearly decreasing relationship between the period index and each nervousness cost. We argue that such a relationship would be more realistic for practical situations. Table 3 illustrates nervousness cost figures when the setup cost-holding cost ratio (S/h) is 40.

3.6. Implementation issues

Because the models are very comprehensive and there are many cases to test, an object-oriented program was constructed with the C++ programming language. Basically, a production schedule was designed to be an object. To specify each object, requirement figures, production and demand amounts, and both cost structures were designed to be private variables. All models were coded as public member functions of the objects. CPU times were obtained by running the models one by one for a single data file, for all seven possible values of the horizon length. Since all arguments of the problem for that particular run were arranged to be the same, it created a robust basis for comparison. A 248 MHz SUN, ultraSPARC-II workstation was utilized for this purpose. The CPU results support the theoretical argument made earlier on the running times. The running times were 1.53, 1.57, 1.79, and 1.04 s for the WW, CJK, MWW, and MSM algorithms, respectively. The CPU time used by the MILP model, 3271.76 s – a little less than an hour – was significantly more than that of any other model.

4. Results

In this section we present a comparison of the various models under a variety of demand distributions. The relative performance of the models was generally insensitive to the choice of demand distributions. None of the models were significantly superior in the Bimodal distribution case. Because the structures of B1 and B2 contain many zero demand periods and are only used in the horizon effect case, a few changes will be sufficient to update previous production schedules. As the probability of having no demand for a period decreases, the robustness of the models decreases, and the difference between the models become more noticeable.

Computational results from different random data files were very close to each other. These results led to the construction of very narrow upper and lower bounds with 95% confidence. In all cases, the resultant cost figures were normalized by dividing them by the corresponding optimal cost that can only be computed by having the perfect information about the future. Naturally, such a figure does not have a nervousness component. This figure enabled us to make better comparisons between all cases.

Fig. 1 shows the horizon effect for various S/h ratios. These results clearly demonstrate that, as the S/h ratio increases, the difference between the models becomes more significant. For example, when the ratio is 20, all models perform similarly. However, when the ratio is 80, it is easy to make a distinction between the models. Here MSM outperforms the others for the high values of the S/h ratio. In all cases of the horizon effect, the total cost approaches the optimal value asymptotically with increasing horizon length. It is important to emphasize that our purpose here is not to identify the best horizon length but to determine which model will outperform the others for the different horizon lengths. In fact, making a comparison between the results from different horizon lengths is unjustifiable. Because more periods are included in the production horizon, the more source of nervousness we would have.

Another significant observation from Fig. 1 is that the conventional cost figures are very stable for all distributions. All of the models were able to identify production schedules very close to optimum even for narrow horizon windows such as 6. Therefore, the nervousness cost component is the one that determines the trends in total cost. For the horizon effect case, nervousness cost figures drop rapidly with the horizon length. Intuitively, that is reasonable because the more we know about the future, the less we pay for unplanned changes in production schedules. This is even valid for the models which are less sensitive to the nervousness concept. For example, the Wagner and Whitin model without nervousness has a similar trend with greater total cost figures. In almost all cases, WW is the model that could provide us the schedules with the smallest conventional cost figures. Since it completely ignores the nervousness costs, it pays more attention

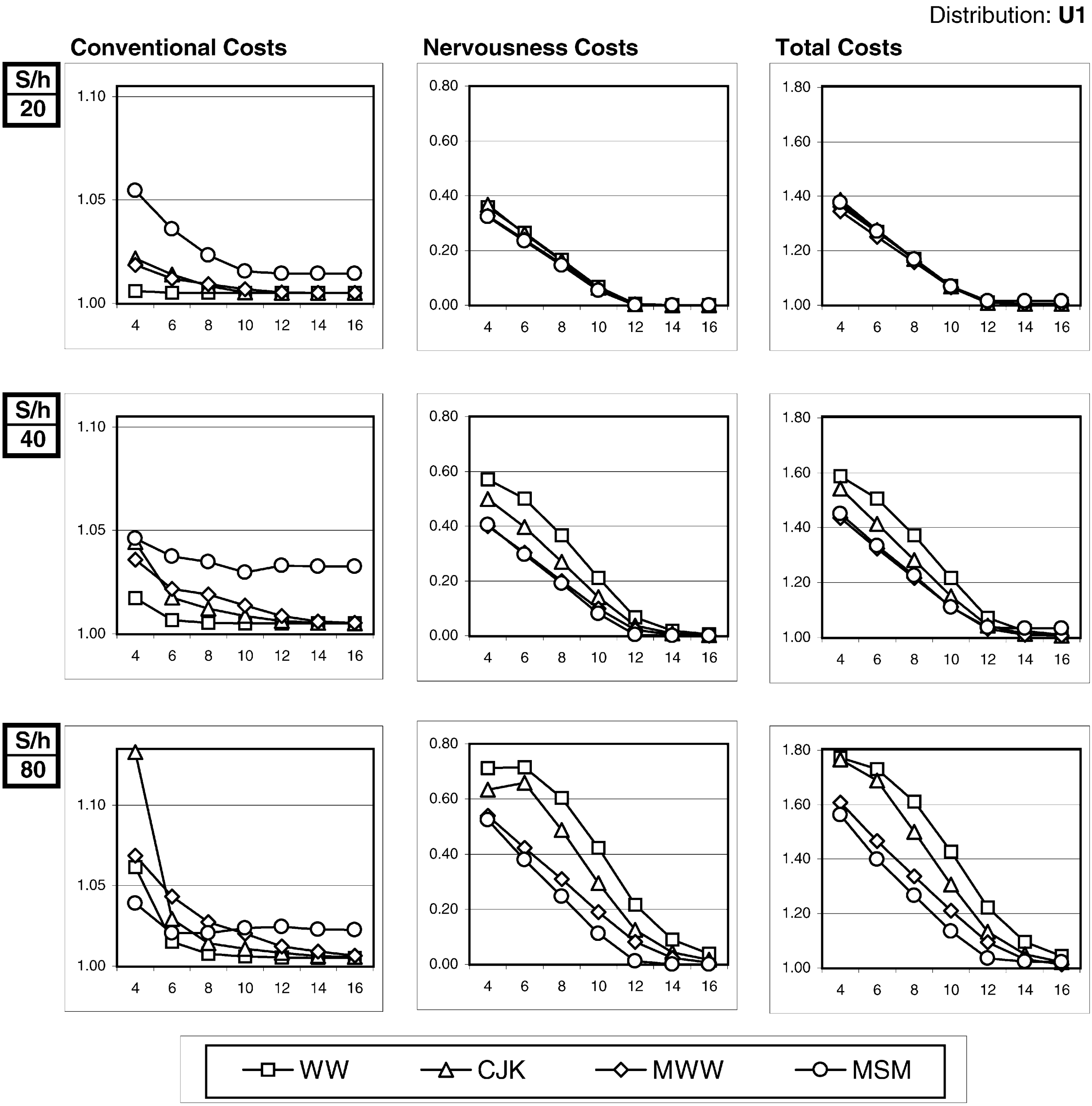

Fig. 1. Comparison of the models in the horizon effect case.

to the conventional cost. However, for total cost it is the worst model in all cases. The CJK model is similar. Its conventional cost figures are reasonable, but since it only penalizes for assigning a brand new production setup, the other two components of the nervousness cost make the total cost higher. In most of the cases, it is MSM that suggests the least nervous schedules, whereas its conventional component is usually higher than the others. The reason behind that might be its

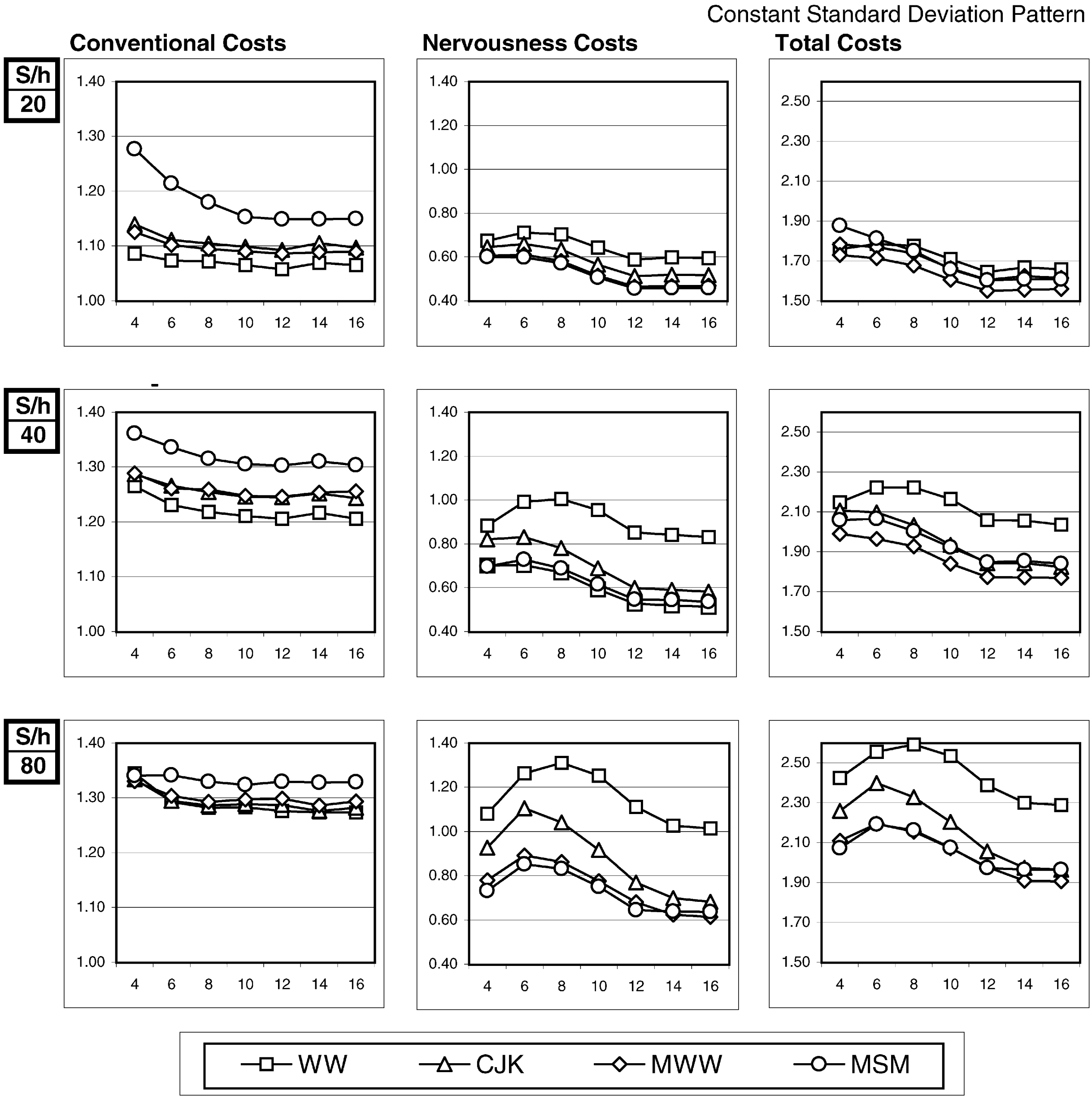

Fig. 2. Comparison of the models in the forecast effect case.

myopic approach to the problem. In contrast to the other models, it tries to identify the best decision for the very first period, then does the same thing for the rest of the window. The MWW model is much better in conventional cost figures and also successful in keeping down nervousness. There is no significant difference between these two models. Both models can be used interchangeably. Although MSM has a better running time, MWW is not computationally prohibitive either.

Fig. 2 summarizes results for the forecast effect. As in the horizon effect case, the performance of the models separates as the S/h ratio increases. On the other hand, because there is more uncertainty involved in horizons, the total cost figures do not tend to approach optimality (perfect information case) but rather some values much higher than the optimal level. Nevertheless, it is still possible to observe a similar asymptotic behavior. When S/h is 40 none of the models could identify a solution within 70% of optimality. For the conventional cost, even WW could not get any closer than 20% of optimality. Evidently, it seems that there is always an unavoidable penalty in the forecast effect case caused by unpredictable changes in the demand figures. This penalty cost increases as the S/h ratio increases. MWW is more successful in most of the cases. As the S/h ratio increases, performance of MSM improves.

Fig. 3 presents the differences between the standard deviation patterns. Uncertainty in the model is reflected in the cost figures. In the increasing order case, the uncertainty is the least and so are the cost figures. On the other hand, for the decreasing order case the cost figures are the highest. In all cases, the myopic MSM is more successful for the small values of the horizon length. For the rest, MWW is slightly better.

The MILP model is not presented in Figs. 1–3, because both of its cost components are nearly identical to those of MWW. The motivation behind the construction of the MILP model was to observe the performance of a model that could guarantee us the optimal solution for a single production window. Because the current cost of altering production volume is not highly emphasized, MWW or even sometimes MSM can easily identify the optimal solution. In order to measure the performance of MILP, some cases in which the cost of altering production volume was significantly higher were established. In the other cases, because of the computational complexity, MWW or MSM should be preferred.

Fig. 4 shows the trends in the cost figures when all three components of the nervousness cost are doubled. The horizon effect case is less sensitive to this change. MWW, MILP, and especially MSM suggest better schedules. However, for the forecast effect case, MILP outperforms all of the other methods. Although deviations are more volatile than the previous cost figures, a similar asymptotic behavior is observed.

Fig. 5 shows the effect of increasing the cost of altering production volume to five times higher than the original cost structure. This corresponds to a production system where decreasing the production volume by 16 units is more expensive than canceling the whole setup. In this case, we thought that it might be quite interesting to visualize the rolling horizon performance of MILP that provides optimality for a single production window. The nervousness cost figures for the models WW and CJK are more than doubled. That is because they ignore that nervousness component. For the horizon effect case, MILP produces slightly better results than the other two procedures (MWW, MSM) for small values of the horizon length. As the horizon gets wider, MSM also becomes preferable. For the forecast effect case, the superiority of MILP is undeniable. It outperforms others significantly for all cases. For this reason, in such extreme cases, MILP is much preferred to the other models.

Fig. 6 summarizes the results of the different cost structures discussed. For all combinations of S/h ratio and horizon length, the model that provides the best result is shown. Therefore, it also identifies which model is the best for a particular case. Again, the results were standardized with respect to the optimal solution with perfect information in order to provide a better basis for comparison and discussion.

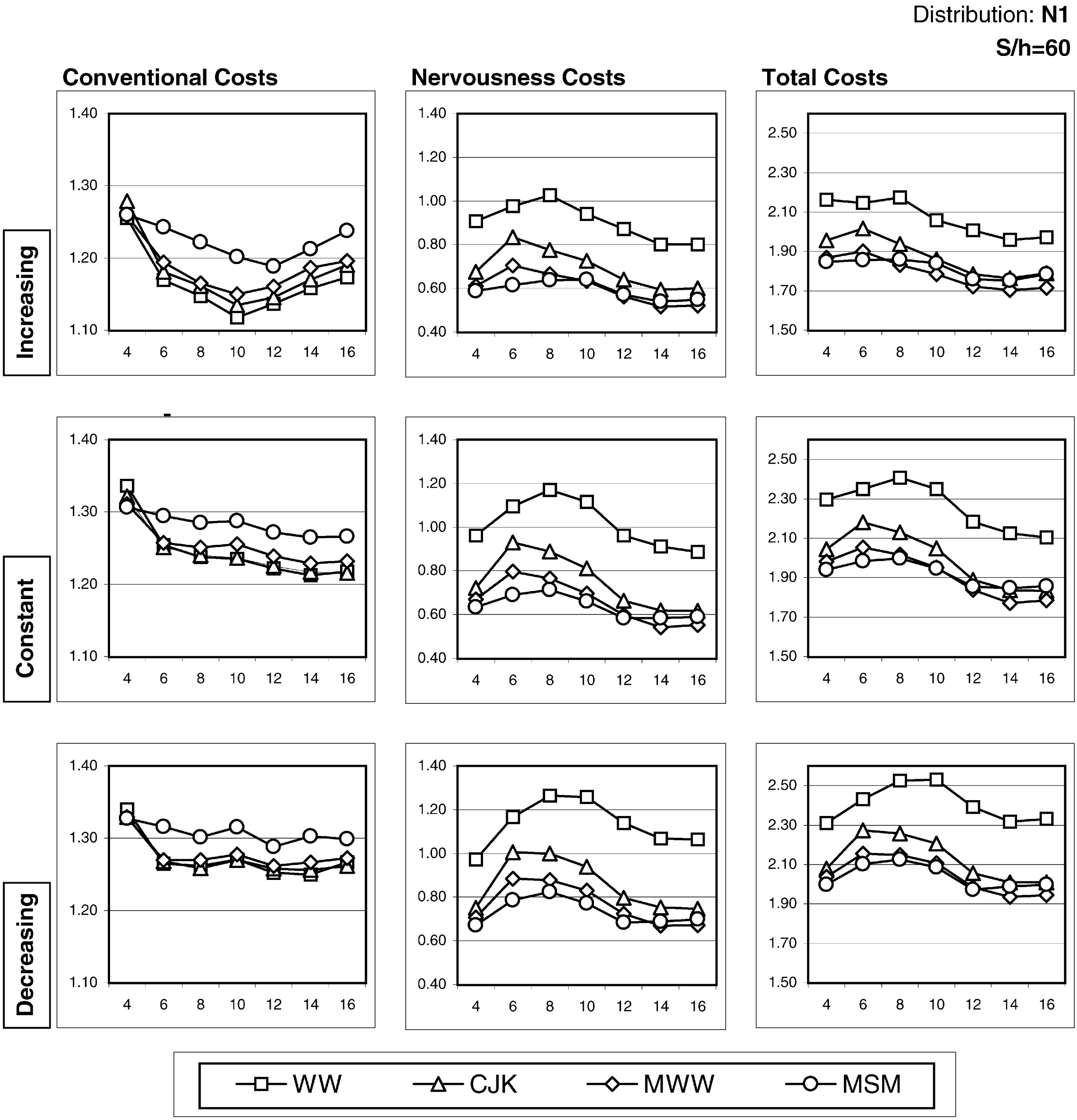

Fig. 3. Comparison of standard deviation patterns for forecast error.

5. Conclusions

Because almost all arrangements in production planning are time dependent, it is an undeniable fact that when new production schedules are being generated, previous ones can never be totally

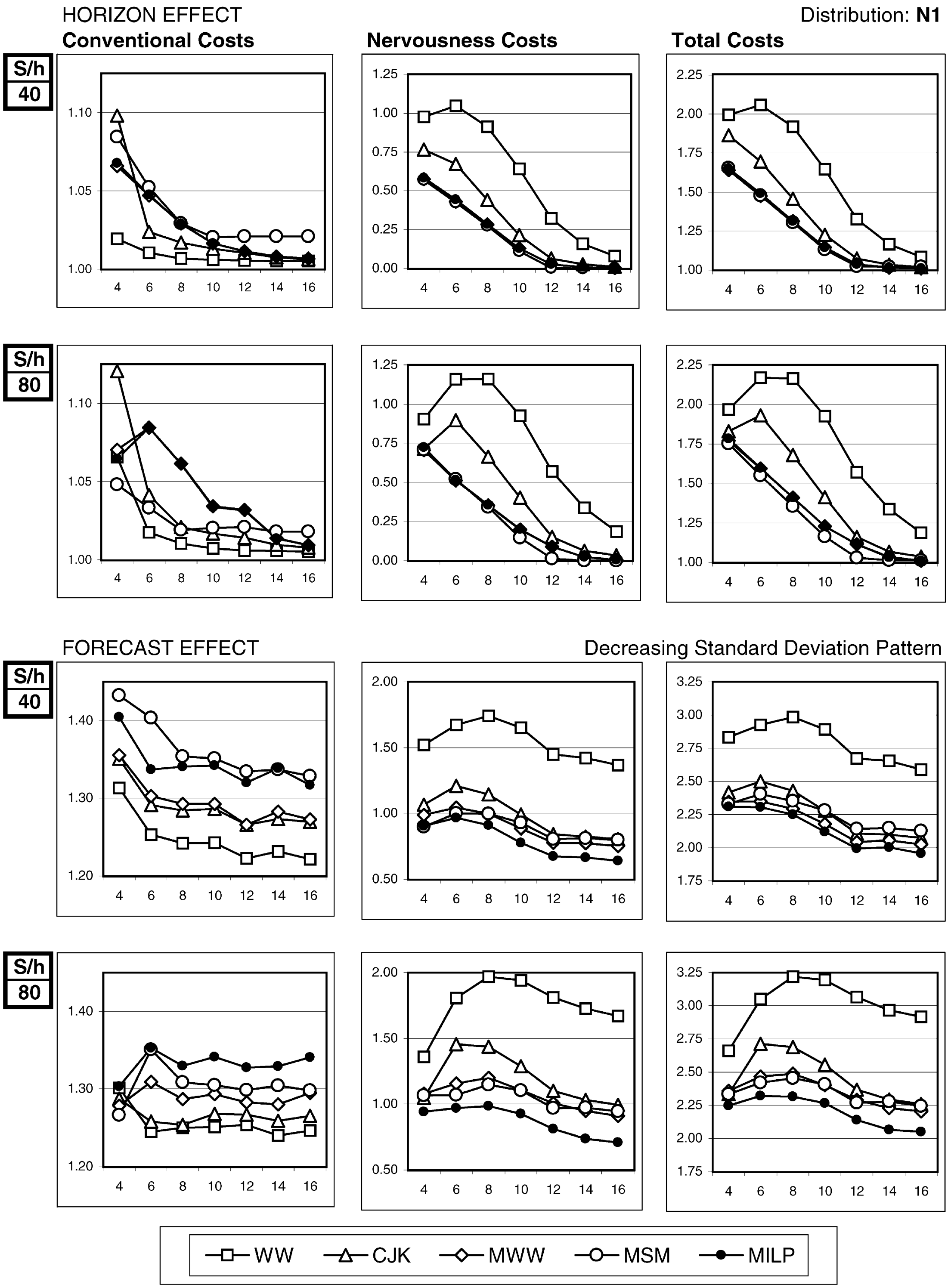

Fig. 4. All nervousness costs are doubled.

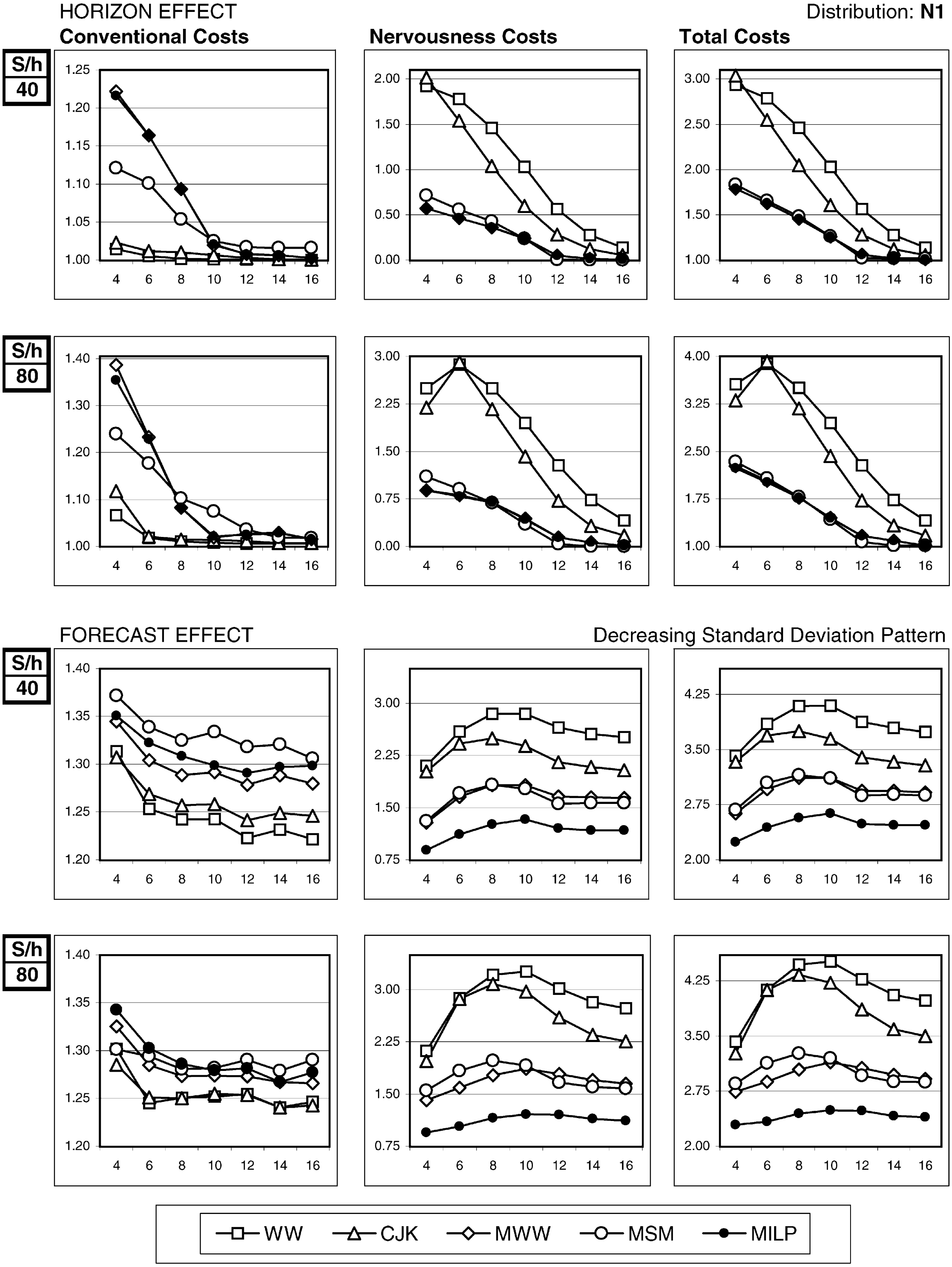

Fig. 5. Altering production volume cost is five times higher.

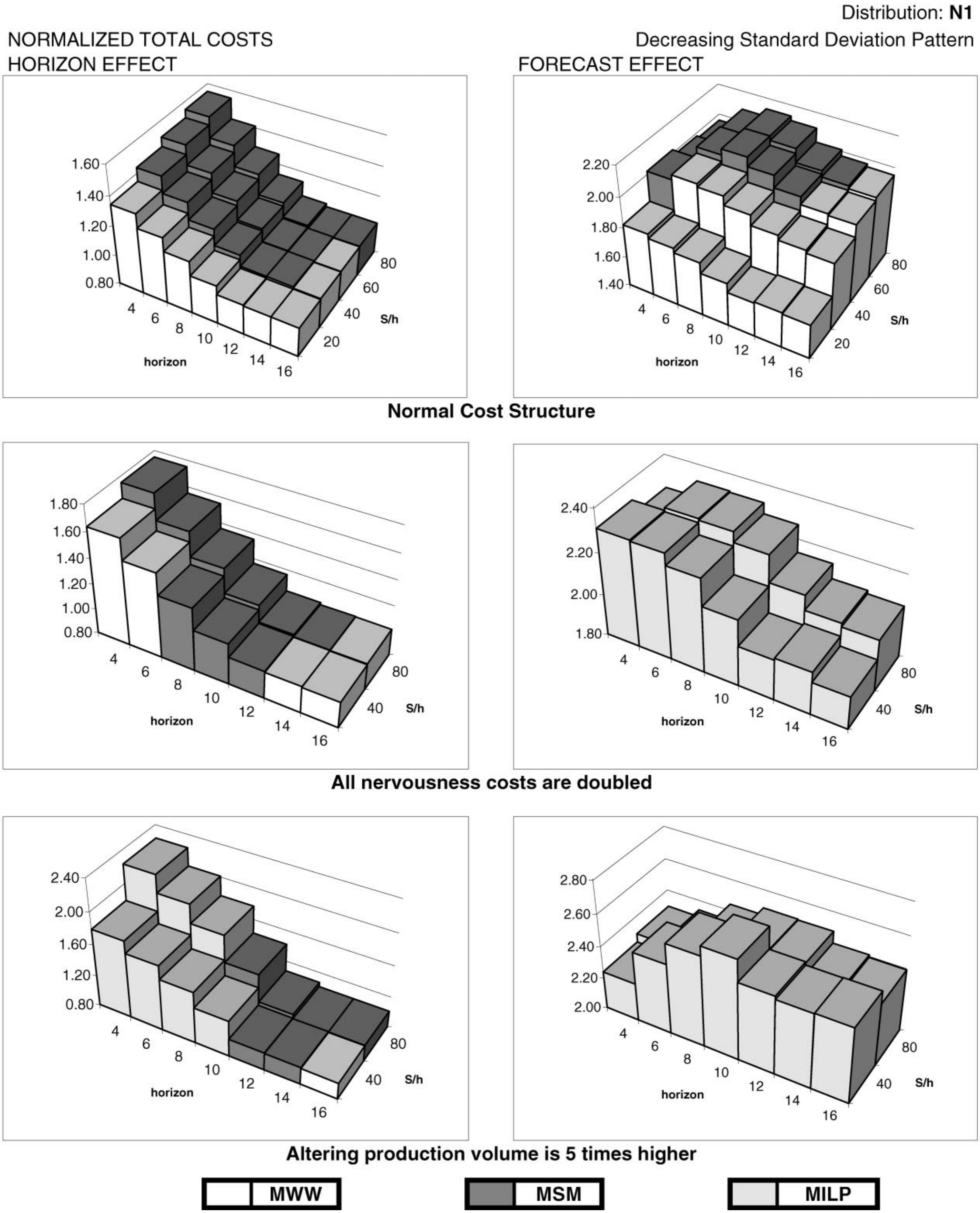

Fig. 6. Best models in parameterized space.

ignored. Therefore, it is perfectly logical to observe some penalties (costs) associated with changes in earlier arrangements. In this work we have suggested a new, broader definition of the nervousness concept that would represent such penalties. Based upon that definition, modified versions of the Wagner–Whitin algorithm and the Silver–Meal heuristic were presented in this paper. Our new comprehensive approach to the nervousness concept changes the optimality conditions of a classical lot-sizing problem for a single horizon. That fact motivated us to construct a Mixed Integer Linear Programming model.

We have compared implementation and performance of the models (WW, CJK, MWW, MSM, and MILP) for the lot-sizing problem. We compared the models in a rolling horizon because a model that may not guarantee optimality for a single period may produce better and more stable production schedules in the long run. This is more likely the case if we have a usual cost structure representing a realistic production system. On the other hand, for some cases a particular method may be more successful because it pays more attention to some components of the problem. Our mixed integer linear programming model (MILP) is the best example of this. When operating a production system that is not flexible to changes in predetermined production volume, the MILP model is the preferable tool to generate new schedules.

In this paper both the horizon and the forecast effects are studied separately. A more realistic approach may suggest combining these effects together. That is, for a single problem we might have a production window in which the first several periods' demand figures are fixed, whereas the remaining periods' demand are subject to change. By making a distinction between the effects, we tried to study the independent cases.

Another important aspect of the methods studied in this paper is that all of them were blind beyond the production horizon. None of the models is interested in the structure of the demand distribution, nor in the probabilities of having some changes in predetermined production volumes. A stochastic model that would take these aspects into account might be more appropriate for the lot-sizing problem on rolling horizon basis. We believe that such extensions would offer interesting directions for future research.

References

[1] Wagner HM, Whitin TM. Dynamic version of the economic lot size model. Management Science 1958;5: 89–96.

[2] Silver EA, Meal HC. A heuristic for selecting lot size quantities for the case of a deterministic time-varying demand rate and discrete opportunities for replenishment. Production and Inventory Management 1973;14:64–74.

[3] Steele C. The nervous MRP system: how to do battle. Production and Inventory Management 1975;16:83–9.

[4] Mather H. Reschedule the reschedules you just rescheduled-way of life for MRP? Production and Inventory Management 1977;18:60–79.

[5] Baker KR. An experimental study of the effectiveness of rolling schedules in production planning. Decision Sciences 1977;8:19–27.

[6] Carlson RC, Jucker JV, Kropp DH. Less nervous MRP systems: a dynamic economic lot-sizing approach. Management Science 1979;25:754–61.

[7] Kropp DH, Carlson RC, Jucker JV. Heuristic lot-sizing approaches for dealing with the MRP system nervousness. Decision Sciences 1983;4:156–69.

[8] De Bolt MA, Van Wassenhove LN. Cost increases due to demand uncertainty in MRP lot sizing. Decision Sciences 1983;14:345–61.

[9] Blackburn JD, Kropp DH, Millen RA. A comparison of strategies to dampen nervousness in MRP systems. Management Science 1986;32:413–29.
[10] Ho CJ, Carter PL. An investigation of alternative dampening procedures to cope with MRP system nervousness. International Journal of Production Research 1996;34:137–56.
[11] Aull RL, LaForge RL. Using the incremental part-period algorithm to compute allowable limits on MRP system nervousness. Production and Inventory Management Journal 1989;30:28–31.
[12] Stidharan SV, Berry WL. Freezing the master production schedule under demand uncertainty. Decision Sciences 1990;21:97–120.
[13] Zhao X, Lee TS. Freezing the master production schedule in multilevel material requirements planning systems under demand uncertainty. Journal of Operations Management 1993;11:185–205.
[14] Gupta SM, Brennan L. MRP systems under supply and process uncertainty in an integrated shop floor control environment. International Journal of Production Research 1995;33:205–20.
[15] Lin NP, Karajewski L. A model for master production scheduling in uncertain environments. Decision Sciences 1995;23:839–61.
[16] Federgruen A, Tzur M. Detection of minimal forecast horizons in dynamic programs with multiple indicators of the future. Naval Research Logistics 1996;43:169–89.

O. Kazan received an M.S. in Industrial Engineering from the State University of New York at Buffalo in 1998. He is now a management consultant with Dopkins & Company, LLP in Buffalo, New York.

R. Nagi is an Associate Professor of Industrial Engineering at the State University of New York at Buffalo. He obtained his Ph.D. from the University of Maryland, College Park in 1991. His research interests are in production systems design, production management, agile and information-based manufacturing.

C. Rump is an Assistant Professor of Industrial Engineering at the State University of New York at Buffalo. He obtained a Ph.D. in Operations Research from The University of North Carolina at Chapel Hill in 1995. His research interests are in queuing and inventory decision models.