ELSEVIER European Journal of Operational Research 92 (1996) 281–301

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

Theory and Methodology

Production strategies for a stochastic lot-sizing problem with constant capacity

N.P. Dellaert *, M.T. Melo [1]

Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

Received November 1993; revised June 1995

Abstract

This paper presents a single item capacitated stochastic lot-sizing problem motivated by a Dutch company operating in a Make-To-Order environment. Due to a highly fluctuating and unpredictable demand, it is not possible to keep any finished goods inventory. In response to a customer's order, a fixed delivery date is quoted by the company. The objective is to determine in each period of the planning horizon the optimal size of production lots so that delivery dates are met as closely as possible at the expense of minimal average costs. These include set-up costs, holding costs for orders that are finished before their promised delivery date and penalty costs for orders that are not satisfied on time and are therefore backordered. Given that the optimal production policy is likely to be too complex in this situation, attention is focused on the development of heuristic procedures. In this paper two heuristics are proposed. The first one is an extension of a simple production strategy derived by Dellaert [5] for the uncapacitated version of the problem. The second heuristic is based on the well-known Silver–Meal algorithm for the case of deterministic time-varying demand. Experimental results suggest that the first heuristic gives low average costs especially when the demand variability is low and there are large differences in the cost parameters. The Silver–Meal approach is usually outperformed by the first heuristic in situations where the available production capacity is tight and the demand variability is low.

Keywords: Production planning; Lot-sizing; Heuristics; Markov decision process

1. Introduction

In this paper we consider a single item capacitated lot-sizing problem motivated by a Dutch manufacturer of steel pipes operating in a Make-To-Order (MTO) environment. Companies working in the MTO sector of industry manufacture products designed specifically to meet the needs of each individual customer. In general, since demand cannot be predicted with a high degree of confidence, no finished goods inventory is stocked. As a result, a delivery date is quoted in response to a customer order and the production planning problem involves determining lot sizes so that delivery promises are satisfied.

The above elements contrast sharply with those of the Make-To-Stock (MTS) sector. In this case, demand can be forecasted with some degree of confi-

* Corresponding author.

[1] On leave from D.E.I.O. (Universidade de Lisboa, Portugal). This research was supported by J.N.I.C.T. (Portugal) under contract BD/2264/92-IA.

0377-2217/96/$15.00
SSDI 0377-2217(95)00166-2

dence. The end item is usually available and can be dispatched directly to the customer. The production planning problem involves determining lot sizes to replenish inventory. Lot sizes are typically based on some function of set-up, inventory and production costs.

Hendry and Kingsman [9] identify the main differences between MTO and MTS companies. In recent years, MTO companies have become a sizable sector of the manufacturing industry as reported by Tobin et al. [26] and Mattsson et al. [16]. Although an increasing number of companies has been shifting its production from the MTS sector to the MTO sector, most of the literature on production planning concentrates on MTS systems. As a consequence, an extensive collection of results is available. Both Silver and Peterson [24] and Graves et al. [8] provide excellent surveys of models and methods for the deterministic and the stochastic demand cases. The MTO area has not received the same degree of attention. Hendry and Kingsman [9] discuss the applicability of some popular production planning procedures like MRP, MRP II, OPT and JIT to MTO companies. According to the authors, such procedures do not address the special characteristics and needs of MTO systems. Despite this belief, several case studies are reported with a claim of success. These include, among others, Hoey et al. [11], McAreavey et al. [17], McClelland [18] and McClelland and Wagner [19].

A key element of a MTO system concerns the ability to meet agreed delivery dates for customer orders. Since MTO companies work in a competitive environment, it is important to maintain those delivery dates to promote customer satisfaction and future business. Kingsman et al. [13] propose a methodology for the quotation of orders while Hendry and Kingsman [10] discuss the control of manufacturing lead times to meet promised delivery dates.

A common tool used in many of the aforementioned papers to evaluate the performance of different production strategies, is simulation. In contrast, we propose in this paper a model for an MTO system for which mathematical analysis is possible. Important elements like demand uncertainty, limited production capacity and complete backordering of items not ready by their delivery date are incorporated in our model. Special attention is given to the derivation of two lot-sizing rules which extend previous work of Dellaert [5] for a similar problem without capacity constraints. To the best of our knowledge, only methods based on queueing theory have been subject to a mathematical analysis of MTO problems. Graves et al. [8] provide a comprehensive review on queueing models for production systems while Dellaert [4] reports on an application of queueing theory to a multi-product MTO problem.

This paper is organized as follows. In the next section we give a detailed description of our problem. In order to cope with the uncertainty regarding the amount as well as the time of receipt of the orders, we model the problem as a Markov Decison Process with discrete state and time space. This is presented in Section 3. Section 4 is devoted to a dynamic programming algorithm for the calculation of the optimal production policy. Since in most practical situations finding an optimal solution is computationally too complex, we concentrate in Section 5 on the design of two heuristic procedures which extend two lot-sizing strategies proposed by Dellaert [5] for the uncapacitated version of our problem. Section 6 reports on computational experience for some test problems. Finally, in Section 7 a summary and some concluding remarks are presented.

2. Problem description

A Dutch company producing a large variety of steel pipes provides the setting for the study in this paper. The company is an example of an actual MTO system and was first observed by Dellaert and Wessels [7]. The production process consists of several stages where the pipes receive a number of treatments. Except for the welding process, there is sufficient capacity in the remaining stages so that most of the time the orders have no substantial waiting times. The welding process which represents the bottleneck of the system, consists of several welding machines capable of dealing with pipes of different sizes. Each time a new type of pipe is produced on a welding machine, parts of the machine are rebuilt. The demand for the pipes is highly uncertain and can not be predicted in advance. As a result, no finished goods inventory is kept and when the company receives an order from a customer, it quotes a promised delivery date. The company's management feels that the irregular and uncertain nature of the demand along with the customer-specific nature of the pipes dictates a policy of production only to order.

Due to the complex nature of the problem, a full description of reality would certainly lead to a mathematical model far beyond our solving capabilities. For this reason some elements will not be included in order to obtain a simple model for which quantitative analysis of lot-sizing rules will be possible. We will consider a single type of product and one machine. The main objective will be to identify all aspects that affect such a simple situation and to assess the performance of different lot-sizing strategies. We expect to gain from such an analysis much insight into the problem and use this information to include more realistic elements in a future study.

We assume that the planning horizon is infinite and that periodic review controls the moments at which decisions concerning the production of known orders are taken. Each review period has a fixed length and the decisions are made only at predetermined times corresponding to the end of the periods. Following the wishes of the management, orders are divided into different categories varying from very urgent orders to orders with a large requested delivery date. In addition, the company is engaged with a regular number of customers. Due to these elements it seems reasonable to assume that the customers are classified into N different groups. A long-term agreement established between the company and the customers controls the assignment of the delivery dates on receipt of the orders. This consists of offering each customer a fixed lead time so that all customers belonging to a group i $(1 \le i \le N)$ obtain a lead time equal to i periods when they place an order. In other words, in an arbitrary period t the company promises to customers of group i $(1 \le i \le N)$ to have their orders ready by the end of period $t+i$.

A further assumption in our model is that the demands for the next N periods are stationary independent stochastic variables with known probability distribution. The reason for this supposition is that it will enable us to give a mathematical steady-state analysis of the problem. Every order placed by a customer is manufactured on one single machine. Customer orders that are not available at the promised delivery date are backlogged. Every period, customers of group i $(1 \le i \le N)$ can place together at most L_i orders. The production level in each period of the horizon is restricted to at most C units. This limitation may be related to conditions such as the number of hours worked per shift, the number of shifts scheduled during a production period, or the capabilities of the production equipment.

Every time the production of a lot is started, a fixed set-up cost is incurred. This cost has a significant value and accounts for the preparation costs of the machine, administrative costs, the wages of skilled workers who have to adjust the machine, etc. Orders that are finished before their delivery date have to be temporarily stocked and for them a holding cost is paid per advanced period. This cost is charged because many customers find it inconvenient to receive untimely deliveries. On the other hand, orders that are not ready by their delivery date lead to penalty costs. This provides an incentive to have orders manufactured on time.

Given the above description of the main features of the problem and the associated cost structure, our objective is to determine in each period of the planning horizon the size of production lots so that delivery dates are met as closely as possible at the expense of minimal average costs.

Cruickshanks et al. [3] present a model for an MTO company which shares most elements of our problem. However, backordering of demand is not allowed since customers do not tolerate late delivery due to their own tight production schedules. In order to smooth production in such a setting, Cruickshanks et al. [3] introduce the concept of a planning window which represents the amount by which the promised delivery date exceeds the production lead time for a given product. Based on simulation, the authors find that, as expected, the variation in production decreases as the planning window increases. Furthermore, a small increase in the planning window can lead to a large reduction in the production variation. By dropping the constraint of no backordering, the authors also derive an approximate model which is tested in a simulation study with normally i.i.d. and stationary demand. The results obtained show some predictive value with regard to the behaviour of the original model. Miltenburg and Sparling [20] address the MTO environment in a more general way by discussing cycle time management (CTM). The purpose of CTM is to reduce the total cycle time of activities that occur during order processing, design of the products, supply management of components, production and distribution of the orders to the customers. The authors develop a simple stochastic model, a Markov chain model and a queueing model to identify the more time-consuming activities of the cycle time. Another problem related to ours is presented by Ten Kate [12] who considers product families, set-up times and processing times among the different families. Contrary to our case, the same lead time is assigned to all families. The modelling approaches include simulation and queueing theory for the derivation of two strategies for the problem of order acceptance. Markland et al. [15] discuss a multi-type problem with time-dependent capacity bounds, arising in an MTO manufacturer of tapes. For this more general problem the authors combine lot-sizing with production scheduling and propose three approaches for generating solutions: a zero-one, goal program and two heuristics. Other problems with a similar cost structure to ours correspond to the generalization of the deterministic inventory model with backordering of Zangwill [28] to the case of stochastic time-varying demand. Silver [22] and later Bookbinder and Tan [2], propose different heuristic procedures for the minimization of the expected total relevant costs over a finite number of periods.

In Dellaert [5] a similar problem without capacity constraints is considered. For that situation, Dellaert proposes a simple decision rule, the so-called (x, T)-rule, where the known demand for the first T periods is produced if the demand for the current period is at least x units. The results obtained for several test problems show that this rule performs reasonably well by offering low average costs and being easy to implement. Due to these good results we propose in this paper an extension of the (x, T)-rule to more complex situations caused by the presence of production capacities. A second lot-sizing strategy is also derived based on the adaptation by Dellaert [5] of the well-known Silver–Meal heuristic [23]. Before giving the details of the two new lot-sizing rules, we present in the next section a Markov decision model for our problem.

3. The model

As in the uncapacitated lot-sizing problem considered by Dellaert [5], the problem is modelled as a Markov Decision Process with discrete state and time space. The state space is denoted by R and consists of order state vectors $r = (r_1, r_2, \dots, r_N)$, where, at the end of an arbitrary period t, r_1 is the number of orders of unit size to be delivered during period $t+1$, including backorders from earlier periods; r_i $(2 \le i \le N)$ is the number of orders of unit size to be delivered during period $t+i$. The set of all possible values for the kth component of the order state vector r is given by R_k, $1 \le k \le N$.

If the available capacity C is greater than the maximum amount that can be ordered by all customers for a certain period, i.e. if $C \ge \sum_{i=1}^{N} L_i$, then the state space is finite. However, if $C < \sum_{i=1}^{N} L_i$ the level of congestion rises due to the accumulation of orders in r_1 and as a result the state space becomes infinite. In this case, the problem can be transformed in order to obtain a finite state space, at least if the average amount of ordered products during each period is less than C. In Section 5 this situation will be discussed briefly for the extension of the (x, T)-rule.

Associated with each state $r \in R$ there is a finite, non-empty set of actions $A(r)$. In the uncapacitated situation, we can restrict our attention to lot sizes that cover exactly an integer number of periods of requirements [5]. This is certainly not the case when production capacities are taken into account. In this case, the nature of the solution becomes considerably more complicated. As observed by Maes and Van Wassenhove [14] for the static version of the problem, where the demand is known completely for a number of periods and backordering is not permitted, a solution can often be improved by allowing the demand for a certain period to be split over different production lots. Therefore, we define an action a as the amount of orders that is produced during one period, that is the size of a lot. In particular, the action $a = 0$ specifies that production is delayed.

Let q_r^a denote the one stage costs of taking action a on observing state r. If $a = 0$ then penalty costs of p units per order have to be paid for all the orders in r_1. If less than r_1 items are produced (i.e. $a < r_1$), it is clear that in addition to set-up costs s, penalty costs for $r_1 - a$ unsatisfied orders are also incurred. In case $a = r_1$ we only have set-up costs. Finally, if $a > r_1$, set-up costs together with holding costs h per order and per period have to be considered. Since $a - r_1$ items will be finished before their delivery date, they will account for the holding costs. In order to describe these costs we are interested in the number of periods for which the total demand can be produced, that is we want to find the value of

$$k_r^a = \max\left\{k \in \{0, 1, \ldots, N\} : \sum_{i=1}^{k} r_i \le a\right\}. \tag{1}$$

Moreover, let us define w_r^a as

$$w_r^a = \begin{cases} a - \sum_{i=1}^{k_r^a} r_i & \text{if } 1 \le k_r^a \le N-1, \\ 0 & \text{if } k_r^a = 0 \text{ or } k_r^a = N. \end{cases} \tag{2}$$

Therefore, given $a > r_1$, holding costs are incurred for the orders $(r_2, \ldots, r_{k_r^a})$ and also for that part of $r_{k_r^a+1}$ that is produced, i.e. w_r^a.

Using the above notation it follows that for a given state $r \in R$ and an action $a \in A(r)$, the corresponding one stage costs q_r^a have the form:

$$q_r^a = \begin{cases} pr_1 & \text{if } a = 0, \\ s + p(r_1 - a) & \text{if } 0 < a \le r_1, \\ s + h \sum_{i=1}^{k_r^a - 1} i r_{i+1} + h k_r^a w_r^a & \text{if } r_1 < a \le \sum_{i=1}^{N} r_i, \end{cases} \tag{3}$$

Furthermore, we can now give a description of the action space $A(r)$. Taking $C' := \min\{\sum_{i=1}^{N} r_i, C\}$ it is clear that

$$A(r) = \begin{cases} \{0, r_1, r_1 + 1, \ldots, C'\} \\ \quad \text{if } 0 \le r_1 \le C \text{ and } r_1 p \le s, \\ \{r_1, r_1 + 1, \ldots, C'\} \\ \quad \text{if } 0 < r_1 \le C \text{ and } r_1 p > s, \\ \{C\} \\ \quad \text{if } r_1 > C. \end{cases}$$

Observe that any action d such that $0 < d < r_1 \le C$ is never optimal since the corresponding one stage costs q_r^{d} given by $s + p(r_1 - d)$ are always higher than the one stage costs of action $a = r_1$. For this reason, action d is not included in the action space.

In order to obtain the transition probability from a state vector r to a state vector z, we introduce d_{ij} as the probability that customers of group i order together j units for a certain period. Moreover, if we take action a in state r, let $Q_a(r)$ denote the state at the end of the next period, just before the new orders are added to the order state. To illustrate the meaning of $Q_a(r)$, suppose that $N > 3$ and on observing the state $r = (r_1, \ldots, r_N)$ with $r_3 > 1$, action $a = r_1 + r_2 + 1$ is taken. This implies that $Q_a(r) = (0, r_3 - 1, r_4, \ldots, r_N, 0)$. If $a \le r_1$ then $Q_a(r) = (r_1 - a + r_2, r_3, \ldots, r_N, 0)$. Hence, using the above definition of k_r^a, $Q_a(r)$ is given by

$$Q_a(r) = \begin{cases} (r_1 - a + r_2, r_3, \ldots, r_N, 0) \\ \quad \text{if } 0 \le a \le r_1, \\ (0, \ldots, 0, r_{k_r^a+1} - w_r^a, r_{k_r^a+2}, \ldots, r_N, 0) \\ \quad \text{if } r_1 < a < \sum_{i=1}^{N} r_i, \\ (0, \ldots, 0) \\ \quad \text{if } a = \sum_{i=1}^{N} r_i. \end{cases} \tag{4}$$

Note that in case $r_1 < a < \sum_{i=1}^{N} r_i$, the first $k_r^a - 1$ components of $Q_a(r)$ are equal to zero.

Let $J \subseteq (0, \ldots, L_1) \times (0, \ldots, L_2) \times \cdots \times (0, \ldots, L_N)$ be the set of all possible one-period demands $(j_1, j_2, \ldots, j_N)$. If on observing the state r we choose action $a \in A(r)$, we enter a state z given by $Q_a(r) + (j_1, j_2, \ldots, j_N)$ and the corresponding transition probability P_{rz}^a is defined by

$$P_{rz}^a = \begin{cases} \prod_{i=1}^{N} d_{ij_i}, & \text{if } (j_1, j_2, \ldots, j_N) \in J, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Given the Markov Decision Process just described, our objective is to find a production policy for which the long-run average costs per period are minimal. The next section will be dedicated to this point.

4. Optimal production policy

If the number of possible states is limited, i.e. if the total expected demand in each period does not exceed the available capacity C, then the state space R is finite and the optimal production policy can be determined by using the same method of successive approximations as described by Odoni [21] and applied to the uncapacitated lot-sizing problem by Dellaert [5]. Our goal is to find the policy that minimizes the total expected costs per transition which are denoted by g. Defining $v_n(r)$ as the total expected costs from the next n transitions when the current state is r and an optimal policy is followed, the iteration scheme takes the form described in the optimality principle by Bellman [1]:

$$v_{n+1}(r) = \min_{a \in A(r)} \left[q_r^a + \sum_{z \in R} P_{rz}^a v_n(z) \right], \quad r \in R,\ n = 0, 1, \dots \tag{6}$$

with q_r^a and P_{rz}^a given by (3) and (5), respectively.

Applying the results of Odoni [21] in a similar way as done by Dellaert [5], it can be proved that any production rule achieving the minima in (6) also has minimal costs per transition. Therefore, starting with $v_0(r) = 0$ for all $r \in R$, the above iteration scheme can be used until a satisfactory degree of convergence is achieved. It follows that g may be estimated by $\frac{1}{2}[\min_{r \in R}\{v_{n+1}(r) - v_n(r)\} + \max_{r \in R}\{v_{n+1}(r) - v_n(r)\}]$. This estimate becomes nearly exact for large n.

Solving a Markov Decision Process by the above algorithm is quite common (see for e.g. Tijms [25]) but from a practical point of view the applicability of this technique is limited due to the extremely fast growth of storage (and time) requirements as the dimension of the state space increases. Computation times become prohibitively large even for small problems. Even if the number of actions we would have to consider would be limited, like in the deterministic case, computation times would soon become too large. For this reason, it is desirable to derive good heuristics, that is fast lot-sizing rules with relatively low average costs per period. In the next section we will extend two lot-sizing strategies developed by Dellaert [5] for the uncapacitated problem.

5. Heuristic procedures

Intuitively appealing heuristic methods are from a practitioner's point of view very attractive. Moreover, they can provide insights into the problem which were not obvious before. The well-known Silver–Meal heuristic [23] is an example of a procedure with intuitive appeal and which gives near-optimal costs in a variety of tests [24]. In contrast, the Wagner–Whitin algorithm [27] is not used extensively in practice, not only because of the relatively complex nature of the algorithm that makes its understanding difficult for practitioners, but also because of the computational effort it requires. The non-realistic nature of the demand also accounts for the low acceptance of the algorithm. In this section we extend two simple lot-sizing strategies derived by Dellaert [5] for the uncapacitated version of our problem. Before describing these heuristic methods in detail, we introduce some of the notation that will be used in Subsections 5.1 and 5.2.

Let $X_{t,t+i}$ denote the number of orders arriving during period t for period $t+i$ with $i = 1, 2, \dots, N$. In other words, $X_{t,t+i}$ contains the demand placed by customers of group i during period t. Fig. 1 gives an illustration of these variables. All demand for a certain period is found in the same column.

The expectation of the random variable $X_{t,t+i}$ is denoted by $u_{t,t+i}$ and due to the assumption of stationary demand is determined by

$$u_{t,t+i} = \mathbb{E}(X_{t,t+i}) = \sum_{j>0} j d_{ij} = u_i. \tag{7}$$

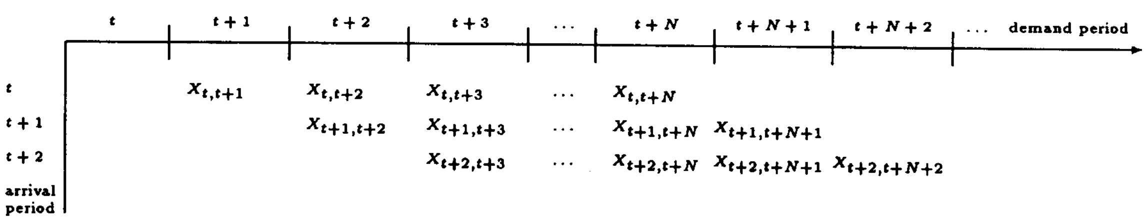

Fig. 1. Demand for various periods.

If we define b_{ilt} as the probability that during the last i periods before an arbitrary period t, customers order a total number of l orders for period t, it follows that

$$b_{ilt} = \mathbb{P}\left\{ \sum_{n=t-i}^{t-1} X_{nt} = l \right\} = \sum_{J_{il}} \prod_{k=1}^{i} d_{kj_k} = b_{il}, \qquad (8)$$

where J_{il} is the set of all possible one-period demands $(j_1, j_2, \ldots, j_N)$ for which the sum of the first i components equals l.

5.1. The (x, T, δ)-rule

For the uncapacitated stochastic lot-sizing problem, the so-called (x, T)-rule proved to be a simple decision strategy by offering low average costs and being easy to implement [5]. Given a pair (x, T), the rule consists of producing the known demand for the first T periods if the required deliveries for the current period together with the backlogged orders (i.e. the value of r_1) are at least x units. Moreover, production can take place less than T periods after the previous one if in the meantime enough orders have arrived.

In this subsection we will extend the above rule to situations with limited available capacity. Although the decision to produce during a certain period will also be based exclusively on the value of r_1, the presence of capacity constraints may not make it possible to manufacture the known demand for T periods completely in certain cases. In situations of capacity shortages, at most C units of product can be manufactured and the analysis of the rule is simplified by assuming that production does not take place during periods for which all the required deliveries have been produced previously. This assumption leads to higher costs than necessary and could be dropped in practice. In case $r_1 \geq x$ and the known orders for the first T periods do not exceed C units, we proceed like in the unconstrained situation by considering the possibility of producing again in the next period or later. Furthermore, for situations of capacity shortages a parameter δ is introduced with the following meaning:

$$\delta = \begin{cases} 1 & \text{if the maximum capacity is used whenever } r_1 \geq x, \\ & r_1 < C \text{ and the total demand for } T \text{ periods exceeds } C, \\ 0 & \text{otherwise.} \end{cases} \qquad (9)$$

This means that in case $r_1 \geq x$ and $r_1 + \cdots + r_T > C$, we manufacture C units of product if $\delta = 1$. Otherwise, the amount produced covers an integer number of periods and therefore, the available capacity may not be used completely in certain cases. Finally, if $r_1 \geq x \geq C$, it is clear that the only possible decision to take in this situation is to fill up the capacity, irrespective of the value of δ.

We will call this new strategy the (x, T, δ)-rule. A relation can be established between the new rule and the well-known (R, s, S) strategy for MTS systems with periodic review [24]. The length of each review period corresponds in our case to $R = 1$. While the (R, s, S) strategy bases its decisions on the inventory position, in an MTO situation the decisions can only be based on the known demand, since no finished goods inventory is kept. Our decision variable x plays a similar role as the parameter s. In the MTS case, the inventory position is raised to S if in the current review instant is not above the reorder point s. This maximum level S is related to the decision variable T in the unconstrained problem. However, under a limited production capacity it is not always possible to follow the strategy induced by T and so the (x, T, δ)-rule deviates from the (R, s, S) strategy in those cases. Observe that the

Table 1
Illustration of the effect of δ

r	δ	Production in $t+1$	Holding costs	Next possible production	Penalty costs
(2, 1, 3, 2, 1)	1	$2+1+(3-1)=5$	$h+2h\times 2$	$t+3$	pu_1
(2, 1, 3, 2, 1)	0	$2+1=3$	h	$t+3$	pu_1
(2, 1, 1, 1, 2)	1	$2+1+1+1=5$	$h+2h+3h$	$t+2$	–
(2, 1, 1, 1, 2)	0	$2+1+1+1=5$	$h+2h+3h$	$t+2$	–

parameter δ only becomes operative when $r_1 \geq x$, $r_1 < C$ and $r_1 + \cdots + r_T > C$.

To give an illustration of the performance of the (x, T, δ)-rule, suppose that $N = 5$, $C = 5$, $x = 2$ and $T = 4$. In Table 1 we illustrate the effect of the parameter δ on the total amount produced, upon the observation of a certain order state vector at the end of a given period t.

Observe that for the first order vector, the known demand for periods $t+1$ and $t+2$ is produced completely and as a result we make a "jump" from period $t+1$ to period $t+3$. By not allowing production during period $t+2$, the demand that arrives during period $t+1$ with a delivery date equal to $t+2$ will be at least one period too late. Recall that u_1 represents the expected number of orders with a lead time of one period (see (7)). Therefore, penalty costs of pu_1 units will have to be paid. Moreover, in period $t+3$ production may take place again only if there are at least 2 required deliveries for that period. In case $\delta = 1$, the requirements for period $t+3$ consist of the unsatisfied demand for $t+3$ together with the number of orders that arrive during periods $t+1$ and $t+2$ for periods $t+2$ and $t+3$ respectively, i.e. a total of $1 + X_{t+1,t+2} + X_{t+1,t+3} + X_{t+2,t+3}$ orders. In the second order state (i.e. $r = (2, 1, 1, 1, 2)$), any choice of δ leads to the same result since $r_1 = 2$ and $r_1 + \cdots + r_4 = 5$. Hence, if the number of orders for period $t+2$, which is given by the stochastic variable $X_{t+1,t+2}$, is equal to or greater than 2, production will take place again in that period.

As already mentioned, given a triplet (x, T, δ), the decision to produce depends exclusively on the value of r_1. Since demand is stationary and we are using a fixed production strategy, there exist time points at which the stochastic process that represents the order state, probabilistically restarts itself. From the coming analysis it will become clear that one of these regenerative points takes place at the end of a period in which the required deliveries for T periods have been produced entirely. We say that a cycle is completed every time a renewal occurs and use the term "absorption" to denote such event. Since the probabilities for each possible order state vector are the same in every regenerative point, the average costs per period associated with the triplet (x, T, δ) are determined by the expected costs during a cycle divided by the expected length of a cycle. Although the vector $r = (r_1, r_2, \ldots, r_N)$ provides all necessary information about the order state, it is not suitable for the analysis of situations with a large number of order state vectors. On the other hand, the decision whether to produce or not in a certain period only depends on the value of r_1. Therefore, like in the unconstrained situation, we will only use in our state description the first component of the demand vector.

Concerning the values of $(r_2, \ldots, r_T)$, they are influenced by the time elapsed since the last absorption. In order to show this, consider $N = 5$, $T = 4$ and assume that during period t the known demand for T periods is manufactured. As illustrated in Fig. 2, at the end of period t, the required deliveries for each one of the periods $t+k$ with $k = 1, 2, 3$ correspond to the orders that arrive during period t with those delivery dates. These orders are given by the stochastic variables $X_{t,t+k}$ for $k = 1, 2, 3$. With re-

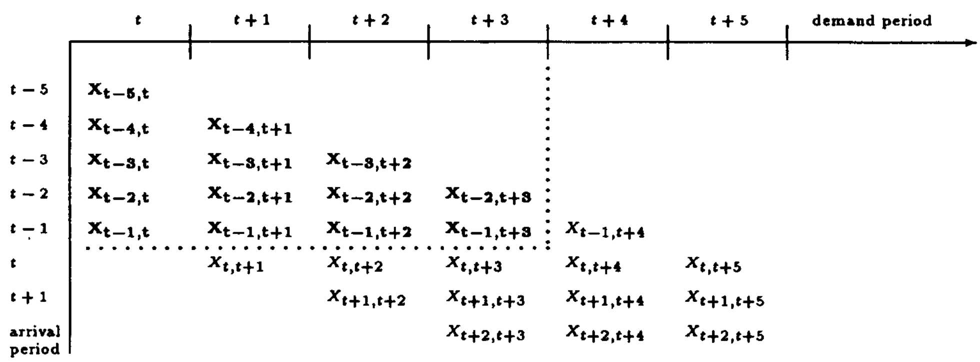

Fig. 2. Example with $N = 5$ and $T = 4$. The demand in bold represents part of the orders that are produced during period t.

Table 2
Order state description at the end of period $t+1$

Component	Value	Delivery date of demand
r_1	$X_{t,t+1}+X_{t,t+2}+X_{t+1,t+2}$	$t+2$
r_2	$X_{t,t+3}+X_{t+1,t+3}$	$t+3$
r_3	$X_{t-1,t+4}+X_{t,t+4}+X_{t+1,t+4}$	$t+4$
r_4	$X_{t,t+5}+X_{t+1,t+5}$	$t+5$

gard to the demand for period $t+4$, known at the end of period t, it can also be seen by Fig. 2 that it is related to the stochastic variables $X_{t-1,t+4}$ and $X_{t,t+4}$.

Suppose now that the required deliveries for period $t+1$ are smaller than x. This implies that production does not occur during that period. Consequently, the amount given by $X_{t,t+1}$ is backordered and due to the rolling horizon effect we obtain by the end of period $t+1$ the order state components as shown in Table 2.

If at the end of period $t+1$, the r_1-value is still smaller than x, then at the end of the following period the value of the three stochastic variables $X_{t,t+3}$, $X_{t+1,t+3}$ and $X_{t+2,t+3}$ is added to r_1. If at the end of period $t+2$ or one of the later periods the value of r_1 is still smaller than x, it is not difficult to see that the value of N stochastic variables will have to be added to r_1. Concerning the values of r_k for $k=2, 3, 4$, observe that at the end of period $t+2$ or one of the later periods, $6-k$ stochastic variables will give each r_k. This means that from period $t+3$ on, each r_k-value is always determined by the same number of variables independently of the time already elapsed since the last absorption.

From this example we can see that for given N and $T \le N$, the value of r_k for $k=2,\dots,T$ is fixed by $N-k+1$ parts of demand in any period after $T-k$. However, during the first $T-k$ periods after absorption the demand for r_k contains less components. Observe also that whenever $r_1 \ge x$ and it is not possible to satisfy the whole demand for the next T periods, we do not allow production during a certain number of periods. The first period after which production can take place again corresponds to a situation where each r_k is given by exactly $N-k+1$ stochastic variables and $2 \le k \le T$. If in situations of capacity shortages we did not proceed in this way, the changes in the r_k-values would be much more difficult to describe. Although the analysis of the (x, T, δ)-rule is considerably simplified in these cases, the updating of the value of r_1 still requires more calculations than in the uncapacitated situation. Later on this will be discussed in detail.

In order to describe the above behaviour of $(r_2,\dots,r_T)$, let us denote by $P_{t,t+i}(r_k=l_k)$ the probability that in period $t+i$, the kth component of the order state vector r has the value $l_k \in R_k$ for $2 \le k \le T$, given that during period t absorption occurred. From the above discussion it follows that

$$P_{t,t+i}(r_k=l_k) = \begin{cases} \mathbb{P}\left(\sum_{n=t}^{t+i-1} X_{n,t+k+i-1}=l_k\right) & \text{if } 1 \le i \le T-k, \\ \mathbb{P}\left(\sum_{n=t+k+i-N-1}^{t+i-1} X_{n,t+k+i-1}=l_k\right) & \text{if } i \ge T-k+1. \end{cases}$$

Let J_{i,l_k} denote the set of all possible one-period demands $(j_1, j_2,\dots, j_N)$ for which the sum of the components from k until $k+i-1$ equals l_k, i.e. $\sum_{n=k}^{k+i-1} j_n = l_k$. Similarly, let $\bar{J}_{k,l_k}$ be the set for which $\sum_{n=k}^{N} j_n = l_k$. Since demand is stationary we can write

$$P_{t,t+i}(r_k=l_k) = \begin{cases} \sum_{J_{i,l_k}} \prod_{n=k}^{k+i-1} d_{n,j_n} & \text{if } 1 \le i \le T-k \\ \sum_{\bar{J}_{k,l_k}} \prod_{n=k}^{N} d_{n,j_n} & \text{if } i \ge T-k+1 \end{cases} \tag{10}$$

$$= P_i(r_k=l_k).$$

Given a triplet (x, T, δ), the corresponding average costs per period are determined by using a similar Markov chain to that of the unconstrained case. Each state is described by a pair (i, j) where i represents the number of periods passed since the last absorption and j contains the value of r_1. The state space in the time-direction will not be limited, i.e. $i \ge 1$ but a limit will be set upon the r_1-component. In the unconstrained situation this limit equals x but this value can certainly not be used in the presence of capacity constraints. Therefore, we will take a sufficiently large value M such that the average costs per period do not change when we increase M by at least one unit. This strategy also applies to the situations mentioned in Section 3 for which $C < \Sigma_{i=1}^{N} L_i$. In those cases, if the average total amount of ordered products during a period is less than C it is possible to transform the problem in order to obtain a finite state space. However, such a transformation is rather complex and causes the computation of the average costs per period to be more difficult to obtain. An alternative way yielding approximately the same results consists of choosing a large M. Fig. 3 gives an illustration of the states that form the Markov chain for a given triplet (x, T, δ).

In order to describe the possible transitions and their expected costs in the above chain, we consider three different cases. If i periods have elapsed since the last absorption we may enter a state (i, j_1) with $0 \le j_1 \le x-1$, a state (i, j_2) with $x \le j_2 \le C$ or a state $(i, C+j_3)$ with $1 \le j_3 \le M-C$. In the first case, since production does not take place, we can move to any state $(i+1, j_1+k)$ with $k \ge 0$ and only penalty costs of pj_1 units are incurred. A different situation occurs when we are in state $(i, C+j_3)$ with $1 \le j_3 \le M-C$. Since the amount produced is exactly C units, we enter a state $(i+1, j_3+k)$ with $k \ge 0$. The costs incurred in this case consist of set-up costs and also penalty costs for j_3 orders that will only be finished after their promised delivery date.

Denoting by c_{ij} the expected costs in state (i, j) with $i \ge 1$ and $0 \le j \le M$, it follows from the above analysis that

$$\begin{aligned} c_{i,j_1} &= pj_1, && i \ge 1,\ 0 \le j_1 \le x-1, \\ c_{i,C+j_3} &= s + pj_3, && i \ge 1,\ 1 \le j_3 \le M-C. \end{aligned} \tag{11}$$

In case we enter a state (i, j_2) with $x \le j_2 \le C$, absorption occurs when the known demand for the next T periods does not exceed the available capacity, i.e. $j_2 + \Sigma_{j=2}^{T} r_j \le C$. Otherwise, we only manufacture the known orders for the next y periods completely with $y = k_r^C < T$ and k_r^C defined in a similar way as in (1). We may also produce part of the demand in r_{y+1} which is given by δw_r^C with w_r^C introduced in (2). From period i we then "jump" to period $i+y$ and do not allow production to take place in any of the periods $i+1, \dots, i+y-1$ which leads to penalty costs for all the required deliveries for those periods. The effect and the costs associated with both the absorption case (A) and the non-absorption case (NA) are listed next. For a detailed analysis we refer the reader to Dellaert and Melo [6].

(A): $j_2 + \Sigma_{k=2}^{T} r_k \le C$
Effect: Produce $(j_2, r_2, \dots, r_T)$ and move from state (i, j_2) to the absorption state $(0', 0)$.
Costs: $s + h\Sigma_{k=2}^{T}(k-1)r_k$.

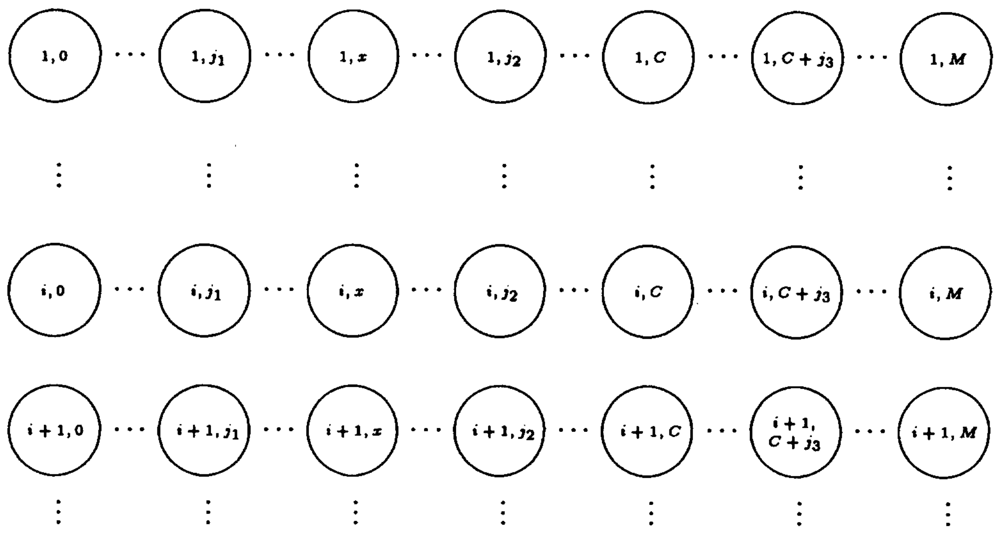

Fig. 3. States of the Markov chain in the (x, T, δ)-rule.

(NA): There exists some $y(y = k_r^C < T)$ such that $j_2 + \sum_{k=2}^{y} r_k \leq C$ and $j_2 + \sum_{k=2}^{y+1} r_k > C$.

Effect: Produce $(j_2, r_2, \ldots, r_y, \delta w_r^C)$ and move from state (i, j_2) to the intermediate state $((i+y)', m_y)$ with $m_y = r_{y+1} - \delta w_r^C$. From this state, enter a state $(i+y, m_y + n)$ with n the number of orders that arrived during periods $i, i+1, \ldots, i+y-1$ for periods $i+1, \ldots, i+y$.

Costs: $s + h\sum_{k=2}^{y}(k-1)r_k + h\delta w_r^C + p\sum_{k=2}^{y}(y+1-k)\sum_{m=1}^{k-1} u_m$ with u_m defined by (7).

In order to obtain a general formula for the above expected holding and penalty costs we introduce some notation. In what follows we replace the index j_2 by j for notational convenience.

For fixed j and $x \leq j \leq C$, let l' represent an N-dimensional vector with components l'_k given by

$$l'_k = \begin{cases} j & \text{if } k = 1 \\ l_k & \text{if } 2 \leq k \leq T, \\ C & \text{if } k = T+1 \text{ and } T < N, \\ 0 & \text{if } T+2 \leq k \leq N \text{ and } T < N-1, \end{cases}$$

where l_k is an element of the set R_k of possible values for r_k and $2 \leq k \leq T$.

Using $k_{l'}^C$, we also define the sets S_y^j such that

$$S_y^j := \{l = (l_2, \ldots, l_T) : k_{l'}^C = y\}$$

for every $y = 1, \ldots, T$, and $j = x, \ldots, C$. Observe that for fixed y and j, the set S_y^j contains all possible demand vectors $(l_2, \ldots, l_T) \in R_2 \times \cdots \times R_T$ for which the first y components and possibly also part of component $y+1$ can be produced.

With the notation introduced, we can now derive a general expression for the expected holding costs $h_{ij}(T)$ in every state (i, j) with $i \geq 1$ and $x \leq j \leq C$.

$$h_{ij}(T) = h \sum_{y=1}^{T} \sum_{l \in S_y^j} (y-1) l_y \prod_{k=2}^{y} P_i(r_k = l_k) 1_{\{y \geq 2\}} + y\delta w_{l'}^C \prod_{k=2}^{y+1} P_i(r_k = l_k) 1_{\{y \leq T-1\}}. \quad (12)$$

In the first part of the above formula, for fixed y and every $l \in S_y^j$ only the expected holding costs for the orders for period $i+y$ are calculated. In fact, the additional costs for the demand $l_2, \ldots, l_{y-1}$ were already included while examining the sets $S_2^j, \ldots, S_{y-1}^j$ since $\sum_{k=2}^{y} l_k \leq C - j$ implies that $l_2 \leq C - j, \ldots, \sum_{k=2}^{y-1} l_k \leq C - j$. Moreover, the formula in (12) also contains the expected holding costs associated with absorption by assuming that in this case a jump to "period $i+T$" is made.

With regard to the penalty costs we need to consider the situations covered by (NA). If for instance $T = 3$ we will jump from period i to period $i+2$ if $r_2 \leq C - j$ and $r_2 + r_3 > C - j$. Hence, the demand that arrives during period i to be delivered by the end of period $i+1$ will certainly not be finished by its delivery date. Therefore, using the expected demand u_m as defined in (7), the expected penalty costs in state (i, j) with $i \geq 1$ and $x \leq j \leq C$ are described as follows:

$$\bar{p}_{ij}(T) = p \sum_{y=2}^{T-1} \left(\sum_{k=2}^{y} (y+1-k) \sum_{m=1}^{k-1} u_m \right) \sum_{l \in S_y^j} \prod_{k=2}^{y+1} P_i(r_k = l_k) \quad \text{if } T > 2. \quad (13)$$

For values of T smaller than 3 there are no penalty costs. Therefore,

$$\bar{p}_{ij}(T) = 0 \quad \text{if } T \leq 2 \quad (14)$$

Finally, the total expected costs in state (i, j) with $i \geq 1$ and $x \leq j \leq C$ are determined by

$$c_{ij} = s + h_{ij}(T) + \bar{p}_{ij}(T) \quad \text{if } T \geq 2, \quad (15)$$

with $h_{ij}(T)$ and $\bar{p}_{ij}(T)$ given by (12), (13) and (14). Clearly, if $T = 1$ we simply have $c_{ij} = s$.

Following the analysis of the expected costs in every possible state (i, j) of the Markov chain presented in Figure 3, we concentrate next on the determination of the transition probabilities. It is not difficult to verify that (see also Dellaert and Melo [6])

$$\mathbb{P}\{(i, j) \to (i+1, k)\} = b_{i+1,k-j},$$
$$i = 1, \ldots, T-2,\ j = 0, \ldots, x-1,\ k = j, \ldots, M,$$
$$\mathbb{P}\{(i, j) \to (i+1, k)\} = b_{N,k-j},$$
$$i \geq T-1,\ j = 0, \ldots, x-1,\ k = j, \ldots, M,$$
$$\mathbb{P}\{(i, C+j) \to (i+1, k)\} = b_{i+1,k-j},$$

$i=1,\dots,T-2,\ j=1,\dots,M-C,\ k=j,\dots,M,$

$\mathbb{P}\{(i,C+j)\to(i+1,k)\}=b_{N,k-j},$

$i\geq T-1,\ j=1,\dots,M-C,\ k=j,\dots,M.$

Moreover, the probability of occurring absorption in state (i, j) with $i \geq 1$ and $x \leq j \leq C$ is denoted by $D_{ij}(T)$ and given by

$$\mathbb{P}\{(i,j)\to(0',0)\}=\sum_{l\in S_T^j}\prod_{m=2}^{T}P_i(r_m=l_m)=D_{ij}(T). \qquad (16)$$

The transition probability from state (i, j) with $i \geq 1$ and $x \leq j \leq C$ to one of the states $(i+y, k)$ with $1 \leq y \leq T-1$ and $1 \leq k \leq M$ requires some elaboration. Defining this probability by $P_{i+y,k}^{ij}$, note that it is associated with the situation in which the orders for periods $i, i+1,\dots,i+y-1$ are produced completely but together with the demand (or part of it) for period $i+y$ they exceed the available capacity. As a result, we jump to an intermediate state $((i+y)', m_y)$ with m_y denoting the number of items of r_{y+1} that are not manufactured, i.e. $r_{y+1}-\delta w_r^C$. Hence, if we define p_{i+y,m_y}^{ij} as the probability that m_y items for period $i+y$ are not produced during period i given that $x \leq j \leq C$, we obtain

$$p_{i+y,m_y}^{ij}=\sum_{l\in S_y^j}\prod_{k=2}^{y+1}P_i(r_k=l_k),$$

with $1 \leq y \leq T-1$ and $m_y=\min(l_{y+1}-\delta w_{l'}^C, M)$.

Thus, the transition probability $P_{i+y,k}^{ij}$ takes into account the above probability together with the probabilities for the arrival of orders during periods $i,\dots,i+y-1$ for periods $i+1,\dots,i+y$. By assuming that during the last n periods before period $i+n$ with $1 \leq n \leq y$, customers place a total number of l_n orders for period $i+n$, we obtain

$$P_{i+y,k}^{ij}=\sum_{L}\prod_{n=1}^{y}b_{n,l_n}p_{i+y,m_y}^{ij}, \quad 1\leq y\leq T-1,$$

$1 \leq k \leq M,$

where L denotes the set of all possible values $l_1 \geq 0,\dots,l_y \geq 0$ and m_y is determined by $k-\sum_{q=1}^{y}l_q$.

Given the above transition probabilities we can now determine easily the probability of visiting state (i, j) during one cycle, i.e. between two absorptions. Let $p_{ij}(T)$ denote this probability for a triplet (x, T, δ). In the first period after the last absorption we have for $T \geq 2$

$$p_{1j}(T)=\begin{cases}b_{1j} & \text{if } 0\leq j\leq M-1,\\ 1-\sum_{k=0}^{M-1}b_{1j} & \text{if } j=M.\end{cases} \qquad (17)$$

In case $T \geq 3$ and $2 \leq i \leq T-1$ we are either in state $(i, 0)$ or in one of the states (i, j) with $1 \leq j \leq M$. Clearly,

$$p_{i0}(T)=p_{i-1,0}(T)b_{i0}. \qquad (18)$$

State (i, j) with $1 \leq j \leq M$ can be reached from different positions in the Markov chain. We can move directly from the previous period $i-1$ where the value of r_1 was smaller than x or greater than C. A third possibility of entering state (i, j) consists of moving from one of the previous periods $i-m$ during which production of the orders in $r_1,\dots,r_m$ took place. Hence, for every $2 \leq i \leq T-1$ and $1 \leq j \leq M$ we obtain

$$p_{ij}(T)=\sum_{k=0}^{\min\{j,x-1\}}p_{i-1,k}(T)b_{i,j-k}+\sum_{m=1}^{\min\{i-1,T-1\}}\sum_{k=x}^{C}P_{i,j}^{i-m,k}p_{i-m,k}(T)+\sum_{k=1}^{\min\{j,M-C\}}p_{i-1,k+C}(T)b_{i,j-k}. \qquad (19)$$

Since after T periods the transition probabilities no longer depend on the i-value, replacing b_{i0} and $b_{i,j-k}$ in (18) and (19) by b_{N0} and $b_{N,j-k}$ respectively, yields the probabilities $p_{ij}(T)$ for $i \geq T \geq 2$ and $0 \leq j \leq M$.

With respect to $T=1$ the expression of the probabilities $p_{ij}(T)$ is as follows:

$$p_{ij}=\begin{cases}b_{Nj} & \text{if } i=1 \text{ and } 0\leq j\leq M,\\ p_{i-1,0}(T)b_{N0} & \text{if } i\geq 2 \text{ and } j=0,\\ \sum_{k=0}^{\min\{j,x-1\}}p_{i-1,k}(T)b_{N,j-k}+\sum_{k=1}^{\min\{j,M-C\}}p_{i-1,k+C}(T)b_{N,j-k} & \\ \text{if } i\geq 2 \text{ and } 1\leq j\leq M. & \end{cases} \qquad (20)$$

By the formulae (11) and (15) the expected costs c_{ij} in every state (i, j) are determined. Moreover, using (17)–(20), the values of $p_{ij}(T)$ are obtained. Hence, the total expected costs during the first t periods after the last absorption are given by

$$C_t(x, T, \delta) = \sum_{i=1}^{t} \sum_{j=0}^{M} c_{ij} p_{ij}(T), \quad t \geq 1, \qquad (21)$$

With the knowledge of the probability $D_{ij}(T)$ defined in (16) which accounts for the occurrence of absorption in a state (i, j) with $i \geq 1$ and $x \leq j \leq C$, we can estimate the time between two absorptions, provided that t periods elapsed since the last absorption period, by

$$\Gamma_t(x, T, \delta) = \begin{cases} \sum_{i=1}^{t} \left(i \sum_{j=x}^{C} D_{ij}(T) p_{ij}(T) \right) \\ \quad \text{if } t \geq 1 \text{ and } T \geq 2, \\ \sum_{i=1}^{t} \left(i \sum_{j=x}^{C} p_{ij}(T) \right) \\ \quad \text{if } t \geq 1 \text{ and } T = 1. \end{cases} \qquad (22)$$

By dividing (21) by (22) we obtain for the triplet (x, T, δ) the average costs per period between two absorptions provided that t periods have passed since the last absorption:

$$g_t(x, T, \delta) = \frac{C_t(x, T, \delta)}{\Gamma_t(x, T, \delta)}, \quad t \geq 1. \qquad (23)$$

In order to determine the average costs per period associated with the triplet (x, T, δ), we can apply the following algorithm with a prespecified tolerance ϵ.

Step 1. Let $g_0(x, T, \delta) := 0$ and $t := 1$.
Step 2. Compute $g_t(x, T, \delta)$ by (23).
Step 3. **If** $| g_{t-1}(x, T, \delta) - g_t(x, T, \delta) | < \epsilon$
Then stop
Else $t := t + 1$; Return to Step 2.

It is possible to speed up the algorithm by replacing $g_t(x, T, \delta)$ in Step 2 by a forecast incorporating an estimation of the number of periods after period t still required for absorption to occur.

To conclude this section we observe that in order to find the optimal triplet (x, T, δ) it is usually not necessary to determine the average costs of every possible combination of x, T and δ. An easily obtainable lower bound on the costs is given by the costs of the best pair $(\bar{x}, \bar{T})$ for the uncapacitated version of the problem as described by Dellaert [5]. This pair $(\bar{x}, \bar{T})$ can be taken as a starting point. If the available capacity is relatively large then $\bar{T}$ is already the correct T-value. For rather small capacity levels it may be necessary to evaluate the costs for $T = \bar{T} + 1$ as the numerical results in Section 6 will show. Due to the jumps that are made in those situations where only part of the demand for the next T periods is manufactured, it is also convenient to determine the costs for $T = \bar{T} - 1$. Regarding the best choice for x, the $\bar{x}$-value is an upper bound and trials with smaller values must be carried out. If the demand distribution involves a large quantity of orders then it may be too time consuming to look for the optimal x. In that case the range of tests may be reduced and from our numerical experience x-values in the neighbourhood of the optimal do not present large deviations in costs. Finally, although the parameter δ can take either the value 0 or 1, only for large capacities and relatively low holding costs it is advantageous to have $\delta = 0$. Usually, it is better to make full use of the available capacity which means that $\delta = 1$ is the proper choice.

5.2. The Silver–Meal approach

For an inventory management problem with a deterministic time-varying demand rate, Silver and Meal [23] proposed a simple lot-sizing rule which consists of selecting the action that produces the (first local) minimum of the total relevant costs per unit of time. These costs are obtained by dividing the expected costs of an action by the number of periods involved in that action. For the uncapacitated version of our problem, Dellaert [5] used the same criterion and derived an heuristic procedure. In this subsection we will further extend the approach of Dellaert [5] to the situation of constant capacity.

As mentioned in Section 3, choosing action a on observing a state $r \in R$ leads to the one stage costs q_r^a given by (3). However, also indirect costs will be involved during the periods covered by action a $(a \neq 0)$ if we presume that during those periods production will not take place again. Assuming that during some period t, a orders are manufactured corresponding to the known demand for say k periods ($k \geq 2$), then by not allowing production to occur in periods $t+1, \ldots, t+k-1$, penalty costs are incurred to all the orders that arrive in the meantime with these delivery dates. The exact penalty costs are difficult to calculate since the future demands are unknown. However, we can replace them by their expected value during the first k periods as done by Dellaert [5] for the uncapacitated problem. This yields the following penalty function $p_a(k)$:

$$p_a(k) = \begin{cases} 0 & \text{if } k = 0, 1 \\ p \sum_{i=2}^{k} (k+1-i) \sum_{j=1}^{i-1} u_j & \text{if } 2 \leq k \leq N. \end{cases} \tag{24}$$

In the unconstrained situation, the total relevant costs per period associated with action a by following the Silver–Meal criterion, are determined by adding the above penalty costs to the costs q_r^a and dividing them by the number of periods k [5]. In the constrained case, it is not enough to simply consider the indirect costs measured by $p_a(k)$. Observe that since our production level is restricted to at most C orders in every period, a decision made in the current period may lead to an increasing level of congestion in future periods. This means that the required deliveries for the periods ahead may go beyond the limit C and as a result, the number of late orders will grow as well as the penalty costs for not satisfying that demand on time.

In order to estimate the impact on the future penalty costs of an action a chosen in some period t, we define z_i as the projected order state vector in period $t+i$ with $i \geq 0$. Clearly, in period t we have $z_0 = r$. In the subsequent periods $t+i$ we define $z_i = Q_{a_{i-1}}(z_{i-1}) + (u_1, \ldots, u_N)$ with a_{i-1} the action chosen in period $t+i-1$ ($i \geq 1$), $Q_{a_{i-1}}(\cdot)$ the vector given by (4) and $u_k (1 \leq k \leq N)$ the expected demand introduced in (7). We set $a_0 = 0$ and $a_i = \min(C, \sum_{j=1}^{N} z_{ij})$ with z_{ij} the jth component of the vector z_i and $i \geq 1$.

Given $a \in A(r)$, the influence of this action is studied during H periods with $H := \{i: z_{i1} - a > C$ and $z_{l1} - a \leq C$ for every $l > i\}$. In each period $i \leq H$ with $z_{i1} - a > C$, the backordered demand equals $z_{i1} - a - C$ orders for which penalty costs are incurred. Therefore, the expected future penalty costs induced by action a are given by

$$\bar{p}(a) = \begin{cases} 0 & \text{if } H = 0, \\ \sum_{i=1}^{H} \max\{0, p(z_{i1} - a - C)\} & \text{otherwise.} \end{cases} \tag{25}$$

Although in the original Silver–Meal heuristic we would divide the above costs by the total number H of periods, in all the computational tests performed we obtained better results by simply using the expression (25).

To illustrate the calculation of (25), consider $N = 4$, $u = (u_1, \ldots, u_4) = (1.2, \ldots, 1.2)$, $r = (1, 6, 3, 1)$, $a = 3$ and $C = 5$. Starting with $z_0 = r$ and $a_0 = 0$, it is clear that $z_1 = (1+6, 3, 1, 0) + u = (8.2, 4.2, 2.2, 1.2)$. Since $z_{11} - a = 8.2 - 3 > 5$ we proceed by choosing action $a_1 = 5$. This yields $z_2 = (8.2 - 5 + 4.2, 2.2, 1.2, 0) + u = (8.6, 3.4, 2.4, 1.2)$. Once again we observe that $z_{21} - a = 8.6 - 3 > 5$. Taking $a_2 = 5$ we obtain $z_3 = (8.6 - 5 + 3.4, 2.4, 1.2, 0) + u = (8.2, 3.6, 2.4, 1.2)$. It is not difficult to see that $z_{l1} - a \leq 5$ for $l \leq 4$ which means that the effect of action $a = 3$ lasts $H = 3$ periods. Moreover, the expected penalty costs equal $p(8.2 - 3 - 5) + p(8.6 - 3 - 5) + p(8.2 - 3 - 5) = p$.

Given $r \in R$, the expected costs of an action $a \in A(r)$ are determined by q_r^a and the penalty functions (24) and (25). If $w_r^a \neq 0$ it means that the demand in the $(k_r^a + 1)$th component of r is not produced completely. This raises the question of how to deal with this situation in the calculation of the expected costs per period. Among several ways, we consider next two different possibilities. We denote by $g_r^i(a)$ the expected costs per period associated with a state $r \in R$ and an action $a \in A(r)$ in option i ($i = 1, 2$).

(1) Whenever part of the demand for a certain period is manufactured, that is $w_r^a \neq 0$, consider that only a fraction $w_r^a / r_{k_r^a + 1}$ of period $k_r^a + 1$ is in-

volved in the costs. This strategy leads to the cost function:

$$g_r^1(a) = \begin{cases} q_r^a + \bar{p}(a) \\ \quad \text{if } a = 0, \\ q_r^a\left(\dfrac{r_1}{C}\right) + \bar{p}(a) \\ \quad \text{if } A(r) = \{C\}, \\ \dfrac{q_r^a + p_a(k_r^a)}{k_r^a} + \bar{p}(a) \\ \quad \text{if } w_r^a = 0 \text{ and } A(r) \neq \{C\}, \\ \dfrac{q_r^a + f_a p_a(k_r^a + 1)}{k_r^a + f_a} + \bar{p}(a) \\ \quad \text{if } w_r^a \neq 0, \end{cases} \tag{26}$$

with $f_a = w_r^a / r_{k_r^a+1}$. Observe that $A(r) = \{C\}$ corresponds to the case $r_1 \geq C$ which is not covered by the definition of w_r^a in (2). Since C orders are produced in this case, the fraction of the first period that is used is given by C/r_1 and so we divide the one stage costs by this number.

Instead of mixing the penalty function $p_a(\cdot)$ as in (26), we can follow the next option.

(2) Whenever part of the demand for a certain period is manufactured, that is $w_r^a \neq 0$, consider that the remaining demand for that period will be produced one period after its delivery date. This strategy leads to the cost function:

$$g_r^2(a) = \begin{cases} q_r^a + \bar{p}(a) \\ \quad \text{if } a = 0 \text{ or } A(r) = \{C\}, \\ \dfrac{q_r^a + p_a(k_r^a)}{k_r^a} + \bar{p}(a) \\ \quad \text{if } w_r^a = 0 \text{ and } A(r) \neq \{C\}, \\ \dfrac{q_r^a + p_a(k_r^a + 1) + p\left(r_{k_r^a+1} - w_r^a\right)}{k_r^a + 1} + \bar{p}(a) \\ \quad \text{if } w_r^a \neq 0. \end{cases} \tag{27}$$

We can also consider a third possibility for the calculation of the expected costs per period by restricting the state space to only those actions that cover exactly an integer number of periods of requirements. Clearly, this only applies to the state vectors r such that $r_1 < C$. In case $r_1 \geq C$ we must keep $A(r) = \{C\}$. Under this strategy we obtain

$$g_r^3(a) = \begin{cases} q_r^a + \bar{p}(a) \\ \quad \text{if } a = 0 \text{ or } A(r) = \{C\}, \\ \dfrac{q_r^a + p_a(k)}{k} + \bar{p}(a) \\ \quad \text{otherwise}, \end{cases} \tag{28}$$

with k the number of periods covered by action a.

In the original Silver–Meal heuristic only a local minimum in the expected costs per period is guaranteed since the procedure is stopped at the action that gives the first increase in costs. In our case, if we observe a state $r \in R$, we take the action that minimizes the expected costs per period over the whole set $A(r)$. This implies that if we follow one of the approaches i described above ($i = 1, 2, 3$) then we compute $\min_{a \in A(r)} g_r^i(a)$ and choose the corresponding action. Denoting that action by a^i we can use a similar dynamic programming algorithm to the one presented in Section 4 to obtain the average costs per period associated with approach i. The iteration scheme in this case takes the form $v_{n+1}(r) = q_r^{a^i} + \sum_{z \in R} P_{rz}^{a^i} v_n(z)$ for every $r \in R$. In a practical setting, in each period of the planning horizon a certain order vector is observed and the action to be implemented is chosen according to one of the approaches described above. The order vector is updated in view of the action taken and the process is repeated in the next period.

6. Numerical results

In this section we evaluate the effect of the lot-sizing rules presented in the previous section, through a set of test examples. The sample problems are intended to cover variations of different important parameters. These variations include changes in the set-up costs, the holding cost rate, the unit penalty cost, the mean demand, and the variability of the demand for the product.

Regarding the choice of the cost structure, the penalty costs are always larger than the holding costs

Table 3
Gaps in percentages of the optimal average costs for binary demand, $N = 4$, $d = 0.5$ and $s = 50$

C	$(h, p) = (5, 15)$					$(h, p) = (10, 15)$					$(h, p) = (5, 10)$				
	(x, T, δ)	XT	SM1	SM2	SM3	(x, T, δ)	XT	SM1	SM2	SM3	(x, T, δ)	XT	SM1	SM2	SM3
3	(2, 3, 1)	3.83	2.72	2.67	6.47	(2, 2, 1)	1.20	0.45	0.79	1.48	(2, 3, 1)	1.17	5.11	3.66	9.56
4	(2, 3, 1)	0.08	2.13	2.13	3.73	(2, 2, 0)	0.27	4.18	4.18	3.09	(2, 3, 1)	0.14	4.25	4.25	10.51
5	(2, 2, 1)	0.37	2.30	2.30	2.67	(2, 2, 0)	0.81	1.76	1.76	0.72	(3, 2, 1)	0.71	0.62	0.62	3.27
6	(2, 2, 0)	0.33	1.13	1.13	1.13	(3, 2, 0)	0.92	0.92	0.92	0.34	(3, 2, 1)	0.09	0.78	0.78	1.47
7	(2, 2, 0)	0.82	0.62	0.62	0.62	(3, 2, 0)	0.89	0.83	0.83	0.65	(3, 2, 1)	0.11	0.74	0.74	1.13
8	(3, 2, 0)	0.83	0.57	0.57	0.57	(3, 2, 0)	1.00	1.02	1.02	1.00	(3, 2, *)	0.52	0.63	0.63	0.73
9	(3, 2, *)	0.83	0.59	0.59	0.59	(3, 2, *)	1.07	1.15	1.15	1.15	(3, 2, *)	0.80	0.68	0.68	0.65
10	(3, 2, *)	0.84	0.59	0.59	0.59	(3, 2, *)	1.07	1.17	1.17	1.17	(3, 2, *)	0.85	0.75	0.75	0.72

* Same results with 0 and 1

but smaller than the set-up costs. These latter costs are set at least three times the value of the penalty costs and in some cases the proportion increases considerably. As for the holding costs, they can be at most 2/3 of the penalty costs. The selection of such a cost structure reflects the differences in the three types of costs that are commonly observed in many practical situations including those encountered in the Dutch company motivating this study.

We will start with a very simple set of examples for which demand follows a binary distribution. Although such a demand pattern is not likely to be observed in practice, it has the advantage that for relatively small values of N, the dimension of the state space is not too large and so we can apply the dynamic programming algorithm of Section 4 to obtain the optimal average costs. Consequently, each heuristic can be compared with the optimal production policy. Moreover, the average costs associated with the Silver–Meal-like strategy can also be computed directly.

Tables 3 and 4 summarize the results obtained by assuming that the demand of each group i of customers follows the same binary distribution with parameter d, that is $d_{i0} = 1 - d$ and $d_{i1} = d$ for $i = 1, \ldots, N$ and $0 < d < 1$. Among the many tests we performed with binary demand, we selected some of the cases that illustrate features of the heuristics frequently encountered. We consider a period as being one week and take $N = 4$ that is, there are 4 groups of customers to whom delivery promises vary from 1 to 4 weeks. The demand parameter d is fixed at 0.5 which leads to an average amount ordered every week of $2(= \Sigma_{i=1}^{4} u_i)$. With this demand pattern we investigate the effects of having different combinations for the cost parameters.

Table 3 refers to experiments with set-up costs $s = 50$ while in Table 4 the results of increasing s to

Table 4
Gaps in percentages of the optimal average costs for binary demand, $N = 4$, $d = 0.5$ and $s = 90$

C	$(h, p) = (5, 15)$					$(h, p) = (10, 15)$					$(h, p) = (5, 10)$				
	(x, T, δ)	XT	SM1	SM2	SM3	(x, T, δ)	XT	SM1	SM2	SM3	(x, T, δ)	XT	SM1	SM2	SM3
3	(2, 3, 1)	2.39	10.94	8.27	13.03	(2, 3, 1)	0.47	7.27	6.80	8.48	(2, 3, 1)	0.86	7.17	6.36	7.36
4	(2, 3, 1)	0.38	6.51	6.51	14.22	(2, 3, 1)	0.91	11.51	7.26	13.57	(2, 3, 1)	1.07	14.03	12.91	14.03
5	(2, 3, 1)	1.52	6.76	6.76	15.72	(3, 2, 1)	0.06	6.09	4.88	13.16	(3, 3, 1)	0.29	18.16	10.83	18.16
6	(3, 3, 1)	0.06	4.03	4.03	7.63	(3, 2, 1)	0.03	2.07	2.07	4.85	(3, 3, 1)	0.31	8.18	8.18	15.70
7	(3, 3, 1)	0.07	1.74	1.74	2.78	(3, 2, 1)	0.96	0.91	1.18	1.59	(4, 3, 1)	0.17	3.16	3.16	10.76
8	(3, 3, 1)	0.18	0.43	0.43	0.67	(4, 2, 1)	0.91	1.31	1.61	1.53	(4, 3, 1)	0.00	1.46	1.46	6.47
9	(3, 3, 0)	0.46	0.11	0.11	0.11	(4, 2, 1)	0.96	1.57	1.85	1.56	(4, 3, 1)	0.12	0.77	0.55	1.72
10	(3, 3, 0)	0.59	0.06	0.06	0.05	(4, 2, *)	1.03	1.71	1.96	1.69	(4, 3, 1)	0.42	0.51	0.51	0.57
11	(3, 3, *)	0.62	0.05	0.05	0.05	(4, 2, *)	1.05	1.75	1.98	1.74	(4, 3, 1)	0.58	0.67	0.67	0.67
12	(3, 3, *)	0.62	0.05	0.05	0.05	(4, 2, *)	1.05	1.76	1.99	1.76	(4, 3, *)	0.62	0.75	0.75	0.75

* Same results with 0 and 1

90 are presented. For both choices of s, we then analyse the influence of variations in the holding and penalty costs in a certain range of capacity levels. We consider the combinations $(h, p) = (5, 15)$, $(h, p) = (10, 15)$ and $(h, p) = (5, 10)$. The first column of Table 3 indicates the available capacity C. This ranges from low values to values beyond which the influence on the production rules is relatively small. The first column is followed by three sets of columns relative to the above pairs (h, p) selected. For each combination of h and p, and for a given capacity C, the triplet (x, T, δ) which yields the lowest costs in the (x, T, δ)-rule is presented. Also, the gap between the average costs of each heuristic and the optimal average costs is determined. We denote by gap the ratio (Heur − OPT)/OPT × 100% where Heur indicates the average costs per period of a heuristic procedure and OPT the average costs per period of the optimal policy. The columns under XT correspond to the gaps obtained with the best (x, T, δ) triplet. Columns under SM1 denote the gaps given by the first variant of the Silver-Meal approach in (26), while columns under SM2 and SM3 refer to the gaps associated with formulae (27) and (28), respectively.

In Table 3 we can observe that when the holding costs are small and the penalty costs are large (i.e. $(h, p) = (5,15)$), the (x, T, δ)-rule only gives the lowest gaps for $4 \leq C \leq 6$. In all the other cases, at least one of the variants of the Silver–Meal approach performs better. Increasing the holding costs from 5 to 10 leads to larger gaps in general and except for $C = 4, 9, 10$, the (x, T, δ)-rule is never the best production strategy. The effect of having relatively small holding and penalty costs (i.e. $(h, p) = (5, 10)$), leads to different results. In this case, the (x, T, δ)-rule gives the smallest gaps in many capacity levels. Furthermore, with a tight capacity of 3, all variants of the Silver–Meal approach lead to large gaps varying between 3% and 10%. The changes in the x, T and δ values can also be analysed in Table 3. Observe that for the pairs $(h, p) = (5, 15)$ and $(h,p) = (5, 10)$, the best T-value increases from 2 to 3 when the available capacity becomes very tight. This is caused by the fact that the holding costs are relatively small and so when the production level is low it pays off to manufacture orders with a larger delivery date. The effects caused by changes in the penalty costs are shown in the x-value of the tests with the combinations $(h, p) = (5, 15)$ and $(h, p) = (5, 10)$. As one would expect, when having late orders becomes more expensive, production starts earlier. The x-value decreases from 3 to 2 when the available capacity is still relatively large ($C = 7$), while with $(h, p) = (5, 10)$ this only occurs for $C = 4$ and $C = 3$. Finally, the δ-value seems to be much influenced by the magnitude of h. High holding costs like in the tests with the pair $(h, p) = (10, 15)$ lead to more cases with $\delta = 0$, since it compensates not to produce orders for the second period whenever $r_1 + r_2 > C$ and $C \neq 3$. With large capacities, the (x, T, δ)-rule becomes insensitive to the value of δ, as expected.

In order to examine the effect of having high set-up costs, we increased s from 50 to 90 and conducted a number of experiments with the same combinations for h and p. The results obtained are presented in Table 4. A first glance to the table shows that the results obtained are considerably different from those in Table 3. Regardless of the choices for h and p, the (x, T, δ)-rule gives almost in every case the lowest gap. Moreover, the three variants of the Silver–Meal approach produce high gaps when the available capacity decreases. This is particularly striking in the tests with the pair $(h, p) = (5, 10)$ and $C \leq 6$. Nevertheless, variant 2 performs slightly better than the other variants. It seems natural that variant 3 is the worst rule when C is very small, since a tight capacity requires its full use in almost every production period. The influence of high set-up costs is also noticeable in the x-value of the (x, T, δ)-rule. This is natural since it becomes more expensive to start the production. The influence of the holding costs is also stronger in the presence of a high s. Observe that production only covers the demand for two periods when $h = 10$. In all the other cases, we manufacture the orders for $T = 3$ periods. Finally, the parameter δ is greatly affected by $s = 90$. In almost every test, it becomes advantageous to fill up the available capacity not only when C is small but also when C is relatively large.

The following examples are constructed by taking a demand pattern that is closer to real situations than the cases discussed above. We assume the demand for the product to be geometrically distributed since in the company motivating our research, it was

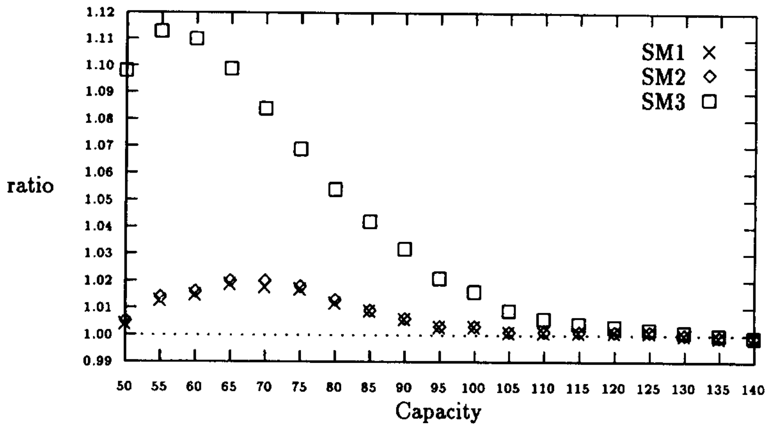

Fig. 4. Cost ratio of the Silver–Meal approach compared to the (x, T, δ)-rule for geometric demand, $\alpha = 5/6$, $k = 5$ and $cv = 0.55$.

observed for most of the product types that by choosing a proper unit size for the steel pipes the demand could be modelled by such a distribution. Regarding the choice of the cost structure, it follows the same criteria as in the binary demand examples, that is $s > p > h$. In order to study the effect of the demand variability upon the competing production rules, two experiments were carried out corresponding to situations with low and high coefficient of variation (cv). The coefficient of variation is the standard deviation of the demand divided by the average demand. Furthermore, due to the large state spaces associated with the new demand pattern, it is no longer possible to determine the optimal policy. Also, the average costs of each variant of the Silver-Meal-like strategy can only be obtained by simulation. In order to compare these costs with those of the (x, T, δ)-rule, the latter is also simulated for different combinations of x, T and δ. In all experiments simulation is carried out through 100000 periods.

Figures 4 and 5 depict the results obtained by taking 4 groups of customers, set-up costs $s = 200$, holding costs $h = 1$ and penalty costs $p = 4$. The choice of such a cost structure is determined not only by the fact that a higher number of orders is placed compared to the binary cases, but also by the information provided by the management of the Dutch company concerning differences in costs. The demand of every group of customers follows a geomet-

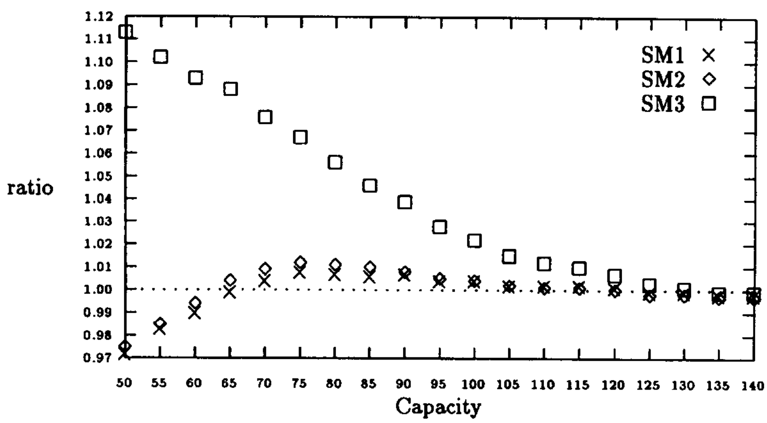

Fig. 5. Cost ratio of the Silver–Meal approach compared to the (x, T, δ)-rule for geometric demand, $\alpha = 10/11$, $k = 0$ and $cv = 1.05$.

ric distribution with the same average of ordered units of product. Since we are interested in the impact of demand variability, we allow the domain of the distribution to be shifted to the right so that different parameter variations leading to the same average but to a different variability, are possible. This means that the orders placed by the ith group of customers ($1 \leq i \leq N$) are generated according to the following expression:

$$d_{ij} = (1-\alpha)\alpha^{j-k}, \quad j = k, k+1, \ldots$$

with $0 < \alpha < 1$ and $k \geq 0$. We fixed the expectation of each group at 10 units and tried the combinations $\alpha = 5/6$, $k = 5$ and $\alpha = 10/11$, $k = 0$. In the first case the coefficient of variation is approximately 0.55 while in the second case we obtain 1.05. In both examples the average demand per period is 40. Fig. 4 presents for $cv = 0.55$ and different capacity levels, the ratios obtained by dividing the average costs of each variant of the Silver–Meal-like strategy by the costs of the best (x, T, δ) triplet.

It can be seen that only for very large capacities ($C \geq 130$) one of the Silver–Meal variants gives lower costs than the (x, T, δ)-rule. Furthermore, SM3 is clearly not suitable for low capacity situations. An increase of the demand variability leads to different results as shown in Fig. 5. For $C < 65$, SM2 and mainly SM1 are better than the (x, T, δ)-rule.

In Table 5 the best choices for x, T and δ are given for both coefficients of variation. The main differences occur in the x-value which increases faster with high demand variability partially due to the shifts produced by the parameter k. As in the binary demand experiments, the best T-value also increases from 2 to 3 when C is very small as a result of having low holding costs in combination with a tight capacity. The only exception occurs for $cv = 1.05$ and $C = 50$. Contrary to the (x, T, δ)-rule, the first two variants of the Silver–Meal strategy seem to handle better situations of low capacity and high demand variability. From the simulations performed we could observe that in those variants the decisions concerning the production of orders are not based exclusively on the value of r_1 like in the (x, T, δ)-rule, but also on r_2 and in some cases even on r_3. The magnitude of the set-up costs together with a small capacity level account for this behaviour. If $r_1 + r_2 < C$ and the required deliveries for the third period are not too large, it proves to be more effective to postpone production. In this way holding costs can be saved.

Table 5
Geometric demand, $N = 4$, $s = 200$, $h = 1$ and $p = 4$

C	$\alpha = 5/6$, $k = 5$, $cv = 0.55$	$\alpha = 10/11$, $k = 0$, $cv = 1.05$
	(x, T, δ)	(x, T, δ)
50	(14, 3, 1)	(8, 2, 1)
55	(17, 3, 1)	(13, 3, 1)
60	(19, 3, 1)	(15, 3, 1)
65	(21, 3, 1)	(19, 3, 1)
70	(22, 2, 1)	(22, 3, 1)
75	(24, 2, 1)	(23, 3, 1)
80	(24, 2, 1)	(24, 2, 1)
85	(26, 2, 1)	(25, 2, 1)
90–95	(28, 2, 1)	(27, 2, 1)
100	(31, 2, 1)	(29, 2, 1)
105	(33, 2, 1)	(29, 2, 1)
110–120	(33, 2, 1)	(30, 2, 1)
125	(33, 2, 1)	(32, 2, 1)
130–135	(34, 2, 1)	(33, 2, 1)
140	(34, 2, 0)	(33, 2, 1)

Finally, we remark that the different production strategies were implemented in Sun Pascal and all experiments conducted on a Sun Sparc Station 5. In the binary demand examples the CPU times in determining both the optimal policy and the average costs of each heuristic took always less than one second. Regarding the simulation experiments, the first two versions of the Silver-Meal approach required on average 3.8 minutes of CPU and the corresponding time for SM3 was 0.25 minutes. The simulation of the (x, T, δ)-rule for a given combination of x, T and δ during 100000 periods took approximately 4.37 seconds and on average 21.37 triplets were evaluated in order to find the lowest costs. The simulation of this rule followed the description provided in Subsection 5.1, that is, whenever the orders for say y periods with $1 \leq y \leq T-1$ are produced during some period i, then a jump to period $i+y$ is made in the sense that production in any of the periods $i+1, \ldots, i+y-1$ is not allowed. Although this jump simplifies considerably the calculation of the average costs, in practice it may be desirable not to take it when large demands occur in the meantime and there is enough capacity. Therefore, we also simulated the (x, T, δ)-rule without considering any jumps. We could observe that in this case the strategy outperforms every variant of the Silver–Meal approach irrespective of the demand variability. Moreover, for $cv = 0.55$ and $C \leq 65$ the best T-value is always 3 while for $cv = 1.05$ this occurs for $C \leq 75$.

7. Summary and conclusions

In this paper we analysed a stochastic lot-sizing problem motivated by a Dutch company operating in an MTO environment. The problem is characterized by highly uncertain demand, fixed delivery dates for customer orders, no possibility for stockkeeping and limited production capacity. We modelled the problem as a Markov Decision Process and used the method of successive approximations to determine the optimal policy and the corresponding long-run average costs. Since from a practical point of view this technique has a limited applicability due to the extremely fast growth of storage (and time) requirements as the dimension of the state space increases, two lot-sizing strategies were proposed.

In the (x, T, δ)-rule, production only takes place during a period for which the required deliveries are at least x units. In that case, the known orders for the next T periods are manufactured if the available capacity is not exceeded. Otherwise, the parameter δ controls the amount to be produced. The second lot-sizing rule is a Silver–Meal-like production strategy where an estimation of the costs per period is obtained for the best action associated with each order state vector. Three different ways of determining these expected costs were proposed. The drawback of having to apply dynamic programming to obtain the average costs of each variant is easily overcome in this case by using simulation.

From the numerical experiments conducted we could learn that both the (x, T, δ)-rule and the Silver–Meal approach are influenced by the level of variability of the demand, the available production capacity and the cost parameters. The (x, T, δ)-rule seems to perform well when the set-up costs are considerably large and the holding cost rate is much smaller than the unit penalty cost. Also, when the demand variability is not too high, the strategy gives low average costs regardless of the capacity level. The Silver–Meal approach is more sensitive to the size of the production capacity and may give high average costs when the set-up costs are very large as displayed by some of the binary demand examples. From the three variants of the rule, the third one is usually not suitable for situations with tight capacities. Even when variants 1 and 2 produce better results than the (x, T, δ)-rule, the differences in costs are small and the solutions may be obtained at the expense of considerable computational effort.

Concerning the company which motivated our research, the different lot-sizing rules were integrated in the operations control level of the hierarchical production planning system used by the company. The detailed decisions produced by the lot-sizing strategies for individual items helped the management to evaluate the impact of planned production capacity and thus the quality of aggregate planning decisions. However, since different types of steel pipes can be manufactured on the same machine, further studies need to be conducted in order to extend the lot-sizing rules to the multi-item situation and devise a practical tool.

Acknowledgements

The authors would like to thank two anonymous referees for their helpful comments on an earlier version of the paper.

References

[1] Bellman, R., *Dynamic Programming*, Princeton University Press, Princeton, NJ, 1957.

[2] Bookbinder, J.H., and Tan, J.-Y., "Strategies for the probabilistic lot-sizing problem with service-level constraints", *Management Science* 34 (1988) 1096–1108.

[3] Cruickshanks, A.B., Drescher, R.D., and Graves, S.C., "A study of production planning in a job shop environment," *Management Science* 30 (1984) 368–380.

[4] Dellaert, N.P., "Multi-item production control for production to order," *Engineering Costs and Production Economics* 17 (1989) 167–173.

[5] Dellaert, N.P., "Production to order: models and rules for production planning," in: *Lecture Notes in Economics and Mathematical Systems* 333, Springer-Verlag, Berlin, 1989.

[6] Dellaert, N.P., and Melo, M.T., "Heuristic procedures for stochastic lot-sizing with capacity constraints," Technical Report TI-93-222, Tinbergen Institute, Rotterdam, The Netherlands, 1993.
[7] Dellaert, N.P., and Wessels, J., "Production scheduling with uncertain demand and no stockkeeping", *Operations Research Proceedings*, 1985, 247–252.
[8] Graves, S.C., Rinnooy Kan, A.H.G., and Zipkin, P.H., (eds.), *Logistics of Production and Inventory*, (Handbooks in Operations Research and Management Science), Elsevier Science, Amsterdam, 1993.
[9] Hendry, L.C. and Kingsman, B.G., "Production planning systems and their applicability to make-to-order companies", *European Journal of Operational Research* 40 (1989) 1–15.
[10] Hendry, L.C., and Kingsman, B.G., "Job release: Part of a hierarchial system to manage lead times in make-to-order companies, *Journal of the Operational Research Society* 42 (1991) 871–883.
[11] Hoey, J., Kilmartin, B.R., and Leonard, R., "Designing a material requirements planning system to meet the needs of low-volume, make-to-order companies (with case study)," *International Journal of Production Research* 24 (1986) 375–386.
[12] Ten Kate, H.A., "Order acceptance strategies and lead times," Paper presented at the Eighth International Working Seminar on *Production Economics*, Igls, Austria, February 21–25, 1994.
[13] Kingsman, B.G., Worden, L., Hendry, L.C., Mercer, A., and Wilson, E., "Integrating marketing and production planning in make-to-order companies," *International Journal of Production Economics* 30–31 (1993) 53–66.
[14] Maes, J., and Van Wassenhove, L.N., "A simple heuristic for the multi item single level capacitated lotsizing problem," *Operations Research Letters*, 4 (1986) 265–273.
[15] Markland, R.E., Darby-Dowman, K.H., and Minor, E.D., "Coordinated production scheduling for make-to-order manufacturing," *European Journal of Operational Research* 45 (1990) 155–176.
[16] Mattsson, L., Olhager, J., Ovrin, P., and Rapp, B., "Computerization of manufacturing planning and control systems: Minicomputer-based systems 1980–1987," *Engineering Costs and Production Economics* 17 (1989) 71–78.
[17] McAreavey, D., Hoey, J., and Leonard, R., "Designing the closed loop elements of a material requirements planning system in a low-volume, make-to-order company (with case study)," *International Journal of Production Research* 26 (1988) 1141–1159.
[18] McClelland, M.K., "Order promising and the master production schedule," *Decision Sciences* 19 (1988) 858–879.
[19] McClelland, M.K. and Wagner, H.M., "Location of inventories in an MRP environment," *Decision Sciences* 19 (1988) 535–553.
[20] Miltenburg, G.J., and Sparling, D., "Managing and reducing total cycle time: Models and analyses," Technical report, Michael DeGroote School of Business, McMaster University, 1993.
[21] Odoni, A., "On finding the maximal gain for markov decision processes," *Operations Research* 17 (1969) 857–860.
[22] Silver, E.A. "Inventory control under a probabilistic time-varying demand pattern," *AIIE Transactions* 10 (1978) 371–379.
[23] Silver, E.A., and Meal, H.C., "A heuristic for selecting lot size requirements for the case of a deterministic time-varying demand rate and discrete opportunities for replenishment," *Production and Inventory Management*, 14 (1973) 64–74.
[24] Silver, E.A., and Peterson, R., *Decision Systems for Inventory Management and Production Planning*, Wiley, 2nd edn., New York, 1985.
[25] Tijms, H.C. *Stochastic Modelling and Analysis – A Computational Approach*, Wiley, New York, 1986.
[26] Tobin, N., Mercer, A., and Kingsman, B.G., "A study of small subcontracting and make-to-order firms in relation to quotation for orders," *International Journal of Operations and Production Management*, 6 (1988) 46–59.
[27] Wagner, H.M., and Whitin, T.M. "Dynamic version of the economic lot size model," *Management Science*, 5 (1958) 89–96.
[28] Zangwill, W.I. "A deterministic multi-period production scheduling model with backlogging," *Management Science* 13 (1966) 105–119.