Computers and Chemical Engineering 26 (2002) 1399–1413

Computers & Chemical Engineering

www.elsevier.com/locate/compchemeng

Markovian inventory policy with application to the paper industry

K. Karen Yin [a,*], Hu Liu [a,1], Neil E. Johnson [b,2]

[a] *Department of Wood and Paper Science, University of Minnesota, St. Paul, MN 55108, USA*
[b] *IT Minnesota Pulp and Paper Division, Potlatch Corporation, Cloquet, MN 55720, USA*

Received 20 August 2001; received in revised form 18 April 2002; accepted 18 April 2002

Abstract

This paper concerns problem formulation and solution procedure for inventory planning with Markov decision process models. Using data collected from a large paper manufacturer, we develop inventory policies for the finished products. To incorporate both variability and regularity of the system into mathematical formulation, we analyze probabilistic distribution of the demand, explore its connection with the corresponding Markov chains, and integrate these into our decision making. In particular, we formulate the Markov decision model by identifying the chain's state space and the transition probabilities, specify the cost structure and evaluate its individual component; and then use the policy-improvement algorithm to obtain the optimal policy. Application examples are provided for illustration.

Keywords: Inventory; Production planning; Markov chain; Markov decision process; Optimal policy; Paper manufacturing

1. Introduction

Inventory management is one of the crucial links of any supply chain. To manufacturers, it entails managing product stocks, in-process inventories of intermediate products as well as inventories of raw material, equipment and tools, spare parts, supplies used in production, and general maintenance supplies. In a broader sense, it comprises all kinds required to run a business including storage, personnel, cash and transportation facilities, etc. This paper focuses on inventory of finished products. A manufacturing company needs an inventory policy for each of its products to govern when and how much it should be replenished. Good inventory management offers the potential not only to cut costs but also to generate new revenues and higher profits. On the contrary, undersupply causes stockout and leads to lost sales; whereas oversupply hinders free cash flow and may cause forced markdowns. As a result of improper inventory policies, both will diminish earnings and can have enough impact to make a company non-profitable. Due to the everchanging market conditions, the dynamic and random nature of the demands, the close and complicated relationship between resource/production planning and product inventory management, as well as the process uncertainties, matching supply with demand has always been a great challenge. Being able to offer the right product at the right time for the right price remains frustratingly elusive (Fisher, Raman, & McClelland, 2000) to manufacturers and retailers.

Process scheduling and planning have attracted growing attention in many industries. Numerous papers in the area of design, operation and optimization of batch as well as continuous plants have been published (see, Applequist, Samikoglu, Pekny, & Reklaitis, 1997; Bassett, Pekny, & Reklaitis, 1997; Pekny & Miller, 1990; Petkov & Mararas, 1997 and the references therein). The main objective of inventory management is to increase profitability. A frequently used criterion for choosing the optimal policy is to minimize the total costs, which is equivalent to maximizing the net income in many cases. Scientific inventory management requires a sound mathematical model to describe the behavior of the underlying system and, quite often, an

* Corresponding author. Tel.: +1-612-624-1761; fax: +1-612-625-6286

E-mail addresses: kyin@umn.edu (K.K. Yin), liux0295@tc.umn.edu (H. Liu), neil.johnson@potlatchcorp.com (N.E. Johnson).

[1] Fax: +1-612-625-6286.

[2] Fax: +1-218-879-1079.

0098-1354/02/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0098-1354(02)00113-8

optimal policy with respect to the model. A large number of works have been published in the past decades (see, e.g. Arrow, Karlin, & Scarf, 1958; Buchan & Koenigsberg, 1963; Buffa, 1980; Gupta, Maranas, & McDonald, 2000; Johnson & Montgomery, 1974; Starr & Miller, 1962; Veinott, 1965 among many others.). Many models have been developed for various inventory situations. The first inventory model appeared in the literature more than 70 years ago (Wilson, 1934) is frequently referred to as the Wilson formulation. This is a fixed order quantity system that selects the order quantity to minimize the total costs in the inventory management. Several of its variations, such as modified reorder point system with periodic inventory counts, the replenishment system, and multiple reorder systems etc. have been widely used. Many inventory systems possess complications that require models capable of handling specific problems in certain situations. Despite the large number of models developed, however, there is still a wide gap between theory and practice.

Similar to many other dynamic processes in the real world, demand variation encountered by retailers or manufacturers is both random and seasonal in nature. A random/stochastic process may be considered as an ensemble of random variables defined on a common probability space and evolving over time. The observed data are statistical time series, which are single realizations of the underlying process. Contrary to those from the deterministic processes, the outcomes from a stochastic process are not unique. Time series of on-line data collected from repetitions of the same experiment will not be the same; levels of demand for a product change from week to week. It is desirable or sometimes necessary to quantify the dynamic relationships among these random events so as to better understand and effectively handle process uncertainties. Considering that the dynamics of such systems are often governed by Markov chains, we resort to Markovian models for solution.

Markov chain, a well-known subject introduced by Markov in 1906, has been studied by a host of researchers for many years (Chung, 1960; Doob, 1953; Feller, 1971; Kushner & Yin, 1997). Markovian formulations (see Chiang, 1980; Taylor & Karlin, 1998; Yang, Yin, Yin, & Zhang, 2002; Yin, Zhang, Yang, & Yin, 2001; Yin & Zhang, 1997, 1998; Yin, Yin, & Zhang, 1995 and the references therein) are useful in solving a number of real-world problems under uncertainties such as determining the inventory levels for retailers, maintenance scheduling for manufacturers, and scheduling and planning in production management. Markov chain approach has been applied in the design, optimization, and control of queueing systems, manufacturing processes, reliability studies and communication networks, where the underlying system is formulated as stochastic control problem driven by Markovian noise. This paper is concerned with problem formulation and solution procedure for inventory planning using Markov decision process (MDP) models.

We conducted this work using data collected from the Potlatch Corporation. A diversified forest product company, Potlatch's manufacturing facilities convert wood fiber into fine, coated paper. In order to improve productivity and product quality as well as to lower production cost, this company has recently completed a 15-year modernization project. Much effort has also been made to take advantage of the tremendous progress in information technology to capture, store and to analyze production and trade data so as to enhance production planning and supply chain management. This work is concerned with inventory planning of coated fine paper produced by Potlatch Corporation's Minnesota Pulp and Paper Division (MPPD).

In the pulp and paper industry, the search for good inventory control policy started several decades ago. Many models have been developed and used. Nevertheless, inventory management remains to be a challenging problem to paper manufacturers and wholesalers. The dynamic and random nature of the demands makes their forecasting very difficult or sometimes impossible. Despite the existence of the large number of inventory models, very often managers cannot find a single one suitable to their needs. As a result, decision has to be made based on a combination of experience, mathematical models, and even on gut feels of a few individuals. It is desirable to shift such experience-based decision making to an information-based decision making. This will require a systematic use of historical data and a theoretically sound mathematical model applicable to the real situation. This work is intended in this direction.

Incorporating both variability and regularity of the system into mathematical formulation, we will analyze the probabilistic structure of the system of interest, explore its connection with the corresponding Markov chains, and integrate these into our decision making. In this work, we consider the demand to be a random variable. We assume periodical review, or the inventory level is checked at fixed intervals and ordering decisions are only made at these times. In addition to the MDP model, a replenishment system model will also be used for comparison.

This paper is organized as follows. Several important concepts and frequently used models in inventory management are outlined first. Followed by a review of Markov chain and the MDP models in Section 3. Section 4 discusses the optimal policy and the policy-improvement algorithm. Application examples are included in Section 5 for illustration and comparison. Summary and further discussion are given in Section 6.

2. Several inventory control models

Operations in inventory systems as well as several important quantities involved are illustrated schematically in Fig. 1. After receiving an order of amount Q, the inventory level will diminish continuously due to sales until the receipt of a new order. The average inventory level I during a period has a direct impact on the carrying costs. Since inventory stock represents a considerable investment, it is desirable to maintain it at the lowest possible level provided that the customer service can be guaranteed. One class of models belong to the fixed order quantity systems, in which the reorder point P is the level of inventory at which a reorder should be placed. There is usually a time lag, also referred to as the lead time, L, between placing a reorder and receiving it. Therefore the average expected demands during this lead time should be included in the reorder amount. To avoid situations such as temporary out-of-stock, back orders, and/or possible lost sales resulted from the demand variation, it is also necessary to include a buffer/safety stock, B, in determining the reorder point. Another class of models are the so-called replenishment system models, where the reorder time is fixed and the reorder quantity varies according to the stock on hand (not shown in the figure). A fixed amount M, called the replenishment level, is used as the upper limit in the determination of the reorder quantity.

The original Wilson formulation is also known as the economic order quantity model or the economic lot-size model. A fixed quantity system, it assumes that the costs in inventory management consist of two parts, ordering cost and carrying cost. Its essence is to choose an economic order quantity to minimize the total costs. A reorder of a quantity Q is placed whenever the inventory falls below the reorder point P. Such system is based on the assumption that perpetual inventory records are kept so that whenever the inventory level falls below the reorder point, a new order can be placed immediately.

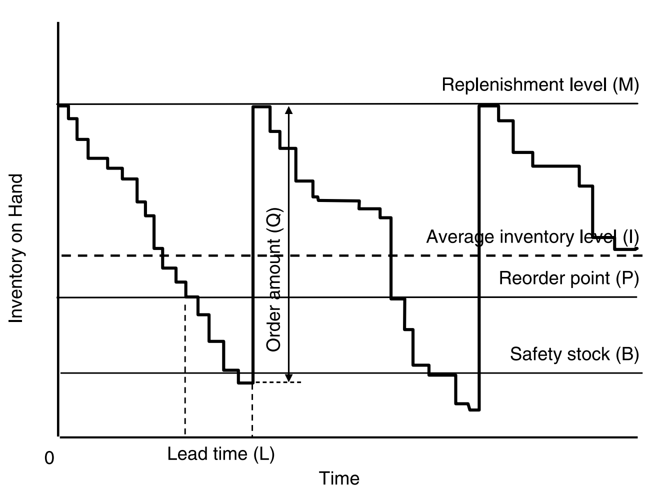

Fig. 1. Operation of an inventory system.

There are several modifications of the simple Wilson formulation and are useful for those cases in which perpetual inventory records are not available. Modified reorder point system with periodic inventory reviews was developed for systems in which reviews are made regularly. The review time is needed in using this model.

The original Wilson model is a single reorder point system. To handle situations where demand during the lead time frequently exceeds the order quantity, multiple reorder systems have been developed. One of the two popular approaches uses the prescribed customer service level (the probability of being out-of-stock) to seek the optimal quantity policy by minimizing the costs. An alternative approach determines the cost of stockouts first before adding it to the cost function to be minimized.

The inventory cost is not explicitly considered in a replenishment system; and neither is the fixed reorder quantity included. However, periodic reviews are required therein. Both the original replenishment model and many of its variations have been widely used. In the original model, the replenishment level M is determined by

$$M = B + \bar{S}_w(L + r), \tag{1}$$

where B is the safety stock; $\bar{S}_w$ denotes the average sales per unit time; L and r are the lead time and the time interval between reviews, respectively. The reorder quantity Q is computed from

$$\begin{cases} Q = M - I & \text{for } L < r \\ Q = M - I - q_0 & \text{for } L \geqslant r \end{cases} \tag{2}$$

where I is the current inventory level and q_0 is the quantity already in order. A widely used optional replenishment system, also known as (S, s) policy, places a lower limit s on the size of the reorder. Since the optional replenishment model possesses the advantages (Buchan & Koenigsberg, 1963) of having a quick response to increased demand and the ability of preventing high inventory during periods of lower demand, it is applicable to a wide range of operating conditions and can usually result in lower costs. For more information of the various inventory models, the readers are referred to Buchan and Koenigsberg, Hillier and Lieberman (1999) and Taylor and Karlin (1998) and the references therein.

3. Markov decision processes

Many processes, such as inventory systems in the real world, have uncertainty associated with them. In the meantime, they exhibit some degree of regularity. It is desirable to incorporate both variability and regularity into mathematical models and treat them quantitatively from a probability point of view. Advances in statistics and stochastic processes have allowed us to do so.

A stochastic process (see, Chiang, 1980; Chung, 1960; Doob, 1953; Feller, 1971) may be considered as a collection of random variables $\{\alpha_t(\omega)\}$ (or $\{\alpha_t\}$) defined on a common probability space and indexed by the parameter $t \in T$, where T is a suitable set and the index t often represents time. For a fixed t, $\alpha_t(\omega)$ is a random variable. For each ω, $\alpha_t(\omega)$ is called a sample path or a realization of the process. Note that the parameter ω is often omitted for brevity. The state space of $\alpha(t)$ is the collection of all values it may take. Stochastic processes can be classified by their index, their state space, and other properties such as stationary vs. non-stationary and jump vs. smooth sample path, etc.

3.1. Markov chain and Markov property

Markov chain is concerned with a particular kind of dependence of random variables involved: When the random variables are observed in sequence, the distribution of a random variable depends only on the immediate preceding observed random variable and not on those before it. In other words, given the current state, the probability of the chain's future behavior is not altered by any additional knowledge of its past behavior. This is the so-called Markovian property. For a discrete-time process, $T = (0, 1, 2, \ldots)$ and

$$P\{\alpha_{t+1} = j | \alpha_0 = i_0, \ldots, \alpha_{t-1} = i_{t-1}, \alpha_t = i\} = P\{\alpha_{t+1} = j | \alpha_t = i\} \quad (3)$$

If the index t of α_t is continuous, i.e. $T = [0, \infty)$, the Markov property means that the values of α_s $(s > t)$ are not influenced by the value of α_u for $u < t$. We consider discrete-time processes in this work.

A stochastic process is a Markov chain if it possesses the Markovian properties and its state space is finite or countable. Throughout the paper, we consider time homogeneous Markov chains. Namely,

$$P_{ij}^{t,t+1} = P\{\alpha_{t+1} = j \,|\alpha_t = i\} \quad \text{for all } t. \quad (4)$$

In other words, the transition probabilities are independent of t. In this case, we say that the Markov chain is stationary and we may drop the time variable t for simplicity.

3.2. Problem statement

We are interested in inventory policies capable of handling situation where the demand during a period is random, and where stock replenishments take place periodically, e.g. every week, every 2 weeks or once every month, etc. Assume that the total aggregate demand for a specific product during any given period n is a random variable d_n. The frequency distribution of the demands encountered in a real-world situation often approximates one of these three probability distributions: Poisson, normal, and exponential (see, e.g. Buchan & Koenigsberg, 1963; Hillier & Lieberman, 1999). The goodness of fit can be examined by using historical data. For cases where none of these three is suitable, the sample mean, variance, and sample frequency distribution of the demands are still easily obtainable from the past records. Making decisions on whether and how much to stock requires a stochastic model. One possible approach is to use a multi-period model that incorporates the probability aspect into dynamic programming. An alternative approach resorts to the MDP model for solution.

Observing the randomness and regularity in the inventory process, we choose to describe it with discrete-time finite-state Markov chains. To establish the mathematical model requires connections linking the elements of the physical process to the logical system, i.e. the Markov chain. More specifically, we need to specify the key elements of the discrete-time Markov chain, which entails designating its state space and prescribing the dependence relations among the random variables based on the real process data.

3.3. State space and transition probabilities of a Markov chain

Let $d_1, d_2, \ldots$ represent the demands (in lbs) for a particular product during the first week, the second week, ... Assume that d_n are independent, identically distributed random variables whose future values are unknown. Let $\tilde{X}_n$ denote the stock of certain product on hand at the end of the nth week. The states of the stochastic process, $\{\tilde{X}_n\}$, consist of the possible values of its stock size. The stock levels at two consecutive periods are related by the current demand d_n and the inventory policy chosen. For example, under the simple (S, s) policy, which requires that replenish to S if the stock level is lower than s; otherwise do not replenish, the current and the next stocks $\tilde{X}_n$ and $\tilde{X}_{n+1}$ satisfy

$$\tilde{X}_{n+1} = \begin{cases} \tilde{X}_n - d_{n+1} & \text{if } s < \tilde{X}_n \leqslant S, \\ S - d_{n+1} & \text{if } \tilde{X}_n \leqslant s. \end{cases} \quad (5)$$

Since the successive demands $d_1, d_2, \ldots$ are independent random variables, the amounts in stock, $\tilde{X}_0, \tilde{X}_1, \ldots$

Table 1
States of the Markov chain

State	Amount of stock at hand
0	$0<\tilde{X}_i\leqslant s$
1	$s<\tilde{X}_i\leqslant s+u$
2	$s+u<\tilde{X}_i\leqslant s+2u$
3	$s+2u<\tilde{X}_i\leqslant s+3u$
$\vdots$	$\vdots$
$m-1$	$s+(m-2)u<\tilde{X}_i\leqslant s+(m-1)u$
m	$s+(m-1)u<\tilde{X}_i\leqslant S$

Table 2
Decisions and actions

Decision	Action
0	Do not replenish
1	Replenish u lbs
2	Replenish $2u$ lbs
3	Replenish $3u$ lbs
$\vdots$	$\vdots$
K	Replenish Ku lbs

constitute a Markov chain whose transition probability matrix can be calculated according to Eq. (5). The weight in stock at time n, $\tilde{X}_n$, is a continuous random variable. To simplify the solution procedure, we discretize it via the following transformation.

Let u denote the minimum order/production amount. For easy presentation, assume the amounts of replenishment will be u or its multiples. We should note that more general cases using any replenishment amount can be treated similarly. Let $s\geqslant 0$ and $S>s$ be a low and the highest possible levels of the inventory. Let

$$X_n=\left\lfloor\frac{\tilde{X}_n-s}{u}\right\rfloor \tag{6}$$

where $\lfloor Z \rfloor$ denotes the integer part of Z. Observe that X_n is a discrete random variable, which indicates the level of the stocks and takes values in $\mathscr{M}=\{0,\ldots,m\}$, where

$$m=\left\lfloor\frac{S-s}{u}\right\rfloor. \tag{7}$$

Such discretization allows us to model this inventory system by an $(m+1)$-state Markov chain, whose state space $X_n\in\mathscr{M}$ is shown in Table 1.

Due to the periodic replenishment and the random demand, at the end of each period the stock level may undergo $m+1$ possible events: it may jump from the current level, i, to a higher one j,$(i<j\leqslant m)$; it may fall into a lower level k, $(0\leqslant k<i)$; or it may stay in the same state.

To establish a stochastic model requires the transition probabilities, that are the bases for dynamic modeling and optimization. The transition probability matrix $P=\|P_{ij}\|$ of the stationary Markov chain $\{X_n\}$ is of the form

$$P=\begin{pmatrix} P_{00} & P_{01} & P_{02} & \cdots & P_{0m} \\ P_{10} & P_{11} & P_{12} & \cdots & P_{1m} \\ P_{20} & P_{21} & P_{22} & \cdots & P_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{m0} & P_{m1} & P_{m2} & \cdots & P_{mm} \end{pmatrix}. \tag{8}$$

To completely define a Markov process requires specifying its initial state (or the initial probability distribution, in general) and its transition probability matrix. For the inventory management problem, the former is usually available, whereas the latter is affected by the random demand as well as the replenishment activities.

3.4. Decisions, actions, and policies

As mentioned earlier, the inventory system evolves over time according to the joint effect of the probability laws and the sequence of decisions and actions. It fits to the general finite-state discrete-time MDPs. The stock on hand at the end of each period is recorded. Subsequently, a decision is made and an action is taken. Table 2 lists the possible decisions, labeled 0, 1, 2,..., K, and a verbal description of their corresponding actions. The question needs to be answered is which decision should be chosen at any given time and state. In other words, an inventory policy is needed.

A policy is a rule that prescribes decisions to be made for each state of the system during the entire time period of interest. Conceivably, there are a number of possible policies for each problem. Characterized by the values $\{\delta_0(R), \delta_1(R),\ldots, \delta_m(R)\}$, any policy R specifies decisions $\delta_i(R)=k$, $(k=0, 1,\ldots, K)$ for all states i, $(i=1,\ldots, m)$ at every time instant. We consider stationary policy only. Namely, the decision is determined by the current state of the system, regardless of time. Very often the problem is to choose a policy that minimizes the long-run expected (average) cost. Table 3 presents two of the many potential policies, R_a and R_b, applicable to this inventory problem.

A policy R requires that the decision $\delta_i(R)$ be made whenever the system is in state i. Effected by this policy as well as the random demand, the system will move to a new state j according to the corresponding probabilities P_{ij}. For countable items, the determination of the transition probability matrix is relatively straightforward. Since we are dealing with products measured by weights, we have 'discretized' them into different levels illustrated in Eq. (6) and Table 1. Eq. (9) gives the transition probabilities P_{ij}, $(i=0, 1,\ldots, 4, j=0, 1,\ldots, 4)$

of a system with five states ($\mathscr{M}=\{0, 1,\dots, 4\}$) if the policy prescribes decision 0 for all states therefore no replenishment at any time.

$$P=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ P\left\{d>\frac{u}{2}\right\} & P\left\{d\leqslant\frac{u}{2}\right\} & 0 & 0 & 0 \\ P\left\{d>\frac{3u}{2}\right\} & P\left\{\frac{u}{2}<d\leqslant\frac{3u}{2}\right\} & P\left\{d\leqslant\frac{u}{2}\right\} & 0 & 0 \\ P\left\{d>\frac{5u}{2}\right\} & P\left\{\frac{3u}{2}<d\leqslant\frac{5u}{2}\right\} & P\left\{\frac{u}{2}<d\leqslant\frac{3u}{2}\right\} & P\left\{d\leqslant\frac{u}{2}\right\} & 0 \\ P\left\{d>\frac{7u}{2}\right\} & P\left\{\frac{5u}{2}<d\leqslant\frac{7u}{2}\right\} & P\left\{\frac{3u}{2}<d\leqslant\frac{5u}{2}\right\} & P\left\{\frac{u}{2}<d\leqslant\frac{3u}{2}\right\} & P\left\{d\leqslant\frac{u}{2}\right\} \end{bmatrix}. \tag{9}$$

3.5. Markov decision process

Owing to their applicability to a wide range of problems in engineering, management science, and biological and social science, MDP models have attracted growing attention in recent years. Many problems in operations research such as stock options, resource allocation, queueing and machine maintenance fit well in the framework of MDP models (Taylor & Karlin, 1998; Yin & Zhang, 1997). Very often the problem is to choose an optimal policy or the best rule for making decisions at each time instant. For most problems it is sufficient to consider only those policies (Hillier & Lieberman, 1999) depending on the state of the system at the present time, and the possible decisions available. The abovementioned inventory management is one of such examples.

As has been noted, the evolution of the system is affected by the random demands as well as the replenishment activities, which are governed by the inventory policy. Let X_n be the state of the system at time n; and let Δ_n be the decision/action chosen. Then under any fixed policy R the pair $Y_n=(X_n,\ \Delta_n)$ forms a two-dimensional Markov chain with transition probabilities

$$P[X_{n+1}=j,\Delta_{n+1}=k'|X_n=i,\Delta_n=k]=p(j\,|i,k)p(k'|\,j\,), \tag{10}$$

where $p(j\,|i,\ k)$ is the conditional probability of the chain's moving to state j at time $n+1$ provided that the current state is $X_n=i$ and a decision $\Delta_n=k$ is taken; and $p(k'|\,j\,)$ is the probability of a specific decision $\Delta_{n+1}=k'$ being chosen at a particular state $X_{n+1}=j$. For a given feedback policy, the decision $\delta_I(R)=k$ is prescribed for every state $i=0, 1,\dots, m$, thus $p(k|i\,)=1$. Consequently, when the system is in state i and the policy R is used and an action based on the decision $\delta_i(R)=k$ is excised, the probability of its moving to state j at the next time period P_{ij} is given by

$$P[X_{n+1}=j\,|X_n=i,\delta_i(R)=k]=p(j\,|i,k) \tag{11}$$

Starting from X_0, the realization of the underlying stochastic process is X_0, X_1,... and the decisions made are Δ_0, Δ_1,... Note that $\Delta_1=\delta_{X_n}(R)\in\{0,\ 1,\ 2,\dots,\ K\}$ if the feedback policy is used. The sequences of observed states and decisions made are called the MDP.

3.6. The long-run expected average cost

Among the many candidate policies, we seek the 'optimal' one in the sense that it will minimize the (long-run) expected average cost per unit time. It should be noted that another consideration in practice is that the policy should be relatively simple and easily implementable.

Suppose a cost $C_{X_n\Delta_n}$ is incurred when the process is in state X_n and a decision Δ_n is made. A function of both $X_n=0, 1,\dots, m$, and $\Delta_n=0, 1,\dots, K$, $C_{X_n\Delta_n}$ is also a random variable. Its long-run expected average cost per unit time over a period of N is

Table 3
Examples of possible policies

Policy	Description	$\delta_0(R)$	$\delta_1(R)$	$\delta_2(R)$	...	$\delta_m(R)$
R_a	Replenish $(m-i\,)u$ lbs for state i	m	$m-1$	$m-2$	...	0
R_b	Replenish $(m-i\,)u$ lbs for state $i<2$	m	$m-1$	0	...	0

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}E[C_{X_n\Delta_n}]=\sum_{i=0}^{m}\sum_{k=0}^{K}\pi_{ik}C_{ik} \tag{12}$$

where π_{ik} is the stationary (limiting) probability distribution associated with the transition probabilities in Eq. (10). Note that for a regular Markov chain (Taylor & Karlin, 1998), the limit probability distribution satisfies

$$\pi_{ik}\geqslant 0 \quad \text{for all } i,k; \quad \text{and}$$

$$\sum_{i=0}^{m}\sum_{k=0}^{K}\pi_{ik}=1. \tag{13}$$

Also,

$$\pi_{jk'}=\sum_{i=0}^{m}\sum_{k=0}^{K}\pi_{ik}p(j\,|i,k)p(k'|\,j)$$

$$\text{for } j=0,1,\ldots,m \quad \text{and } k'=0,1,\ldots,K. \tag{14}$$

Our objective is to find a policy that minimizes the long-run expected average cost given in Eq. (12) where π_{ik} is related to the policy through Eqs. (13) and (14).

4. Optimal policy and the policy-improvement algorithm

A policy R can also be written in a matrix form

$$D=\begin{pmatrix} D_{00} & D_{01} & D_{02} & \cdots & D_{0K}\\ D_{10} & D_{11} & D_{12} & \cdots & D_{1K}\\ D_{20} & D_{21} & D_{22} & \cdots & D_{2K}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ D_{m0} & D_{m1} & D_{m2} & \cdots & D_{mK}\end{pmatrix} \tag{15}$$

The first subscript i of any element D_{ik} in the matrix $\|D\|$ represents the state; and the second subscript k stands for the decision. Observe that D_{ik} take values of 0 and 1 only. $D_{ik}=1$ means $\delta_i(R)=k$ that calls for decision k and its corresponding action when the system is in state i; whereas $D_{jk}=0$ means $\delta_j(R)=0$, i.e. no action will be taken when the system is in state j.

Since for any given policy, the decision to be made and the action to be taken in any state i has been specified, Eqs. (12)–(14) can be further simplified by replacing π_{ik} with π_i, or

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}E[C_{X_n\Delta_n}]=\sum_{i=0}^{m}\sum_{k=0}^{K}\pi_i C_{ik} \tag{16}$$

and

$$\pi_i\geqslant 0 \quad \text{for } i=0,1,\ldots,m; \quad \text{and } \sum_{i=0}^{m}\pi_i=1 \tag{17}$$

where π_i is the stationary distribution and

$$\pi_i=\lim_{n\to\infty}P_{ii}^{(n)}=\sum_{j=0}^{m}\pi_j P_{ji}. \tag{18}$$

Although the expected cost of a policy can usually be expressed as a linear function of D_{ik}, the linear programming (LP) method is not directly applicable here due to its requirement of continuous variables whereas D_{ik} being discrete. This difficulty is surmountable (Hillier & Lieberman, 1999) by modifying the interpretation of the policy as displayed in matrix (Eq. (15)). In the modification, the D_{ik} are considered to be probability distributions for the decision k to be made when the system is in state i, i.e.

$$D_{ik}=P\{\text{decision}=k|\text{state}=i\} \quad \text{for } i=0,1,\ldots,m;$$

$$k=0,1,\ldots,K, \tag{19}$$

thus

$$0\leqslant D_{ik}\leqslant 1 \quad \text{and} \quad \sum_{k=0}^{K}D_{ik}=1. \tag{20}$$

Having changed D_{ik} into continuous variables, such treatment enables us to use the LP approach in finding the optimal solution.

Rather than using the LP method, we resort to the Policy-Improvement Algorithm in this work, where D_{ik} take values of 0 and 1 only. Let $g(R)$ represent the long-run expected average cost per unit time following any given policy R, i.e.

$$g(R)=\sum_{i=0}^{m}\pi_i C_{ik} \tag{21}$$

Denote $v_i^n(R)$ the total expected cost of a system starting in state i and evolving in a period of length n. By definition, it satisfies the following recursive formula

$$v_i^n(R)=C_{ik}+\sum_{j=0}^{m}P_{ij}(k)v_j^{n-1}(R). \tag{22}$$

Eq. (22) means that this total expected cost $v_i^n(R)$ consists of two parts, the cost incurred in the first time period, C_{ik}, and the total expected costs thereafter. Note that C_{ik} is also an expected cost and

$$C_{ik}=\sum_{j=0}^{m}q_{ij}(k)P_{ij}(k) \tag{23}$$

where $q_{ij}(k)$ and $P_{ij}(k)=p(j\,|i,k)$ are the expected cost and the probability when the system moves from state i to state j after decision k is made, respectively. It can be shown that (Hillier & Lieberman, 1999)

$$g(R)=C_{ik}-v_i(R)+\sum_{j=0}^{m}P_{ij}(k)v_j(R) \quad \text{for } i=0,1,\ldots,m. \tag{24}$$

For a system of $m+1$ states, Eq. (24) consists of $m+1$ simultaneous equations but $m+2$ unknowns, $g(R)$ and $v_i(R)$ ($i=0$, 1,..., m). To obtain a unique solution, it is customary to specify $v_m(R)=0$. Solving the system of equations (Eq. (24)) yields the long-run average expected cost per unit time $g(R)$ if the policy R

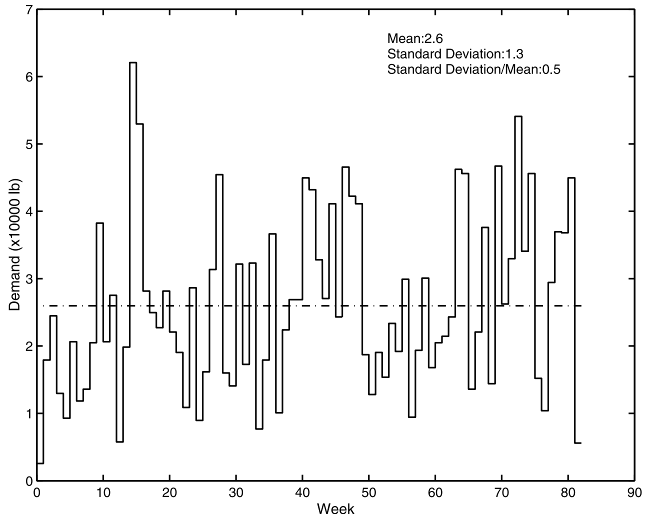

Fig. 2. Weekly demands for $P-1$ (June 7, 1999–December 25, 2000).

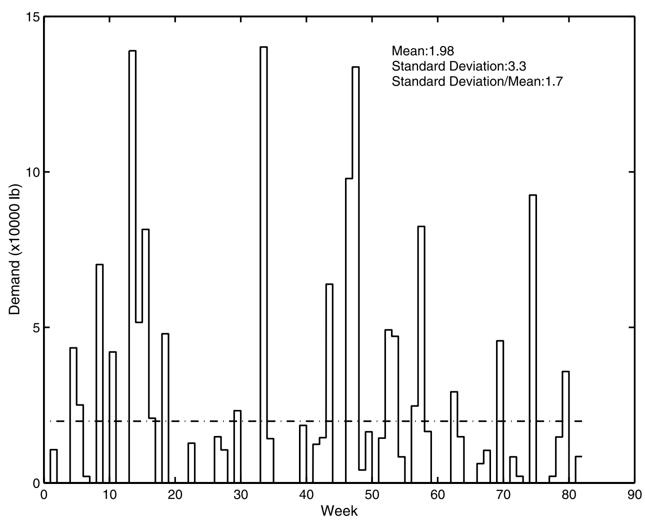

Fig. 3. Weekly demands for $P-2$ (June 7, 1999–December 25, 2000).

is used. An optimal policy is one that results in the lowest cost $g(R^*)$. A policy-improvement algorithm, an iteration procedure consisting of two steps, allows us to obtain the optimal policy. Since there are only a finite number of possible stationary policies when the state space is finite, we will be able to reach the optimal policy in a finite number of iterations (Ross, 1983). The procedure begins by choosing an arbitrary policy R_1. For the given policy R_1, the transition probabilities $P_{ij}(k)$ are available hence the expected costs $C_{ik_{R_1}}$ in Eq. (23) can be computed. Subsequently, the values of $g(R_1)$, $v_0(R_1)$, $v_1(R_1)$,..., $v_{m-1}(R_1)$ can be obtained from Eq. (24). In the second step, the current values of $v_i(R_1)$ are used to find an improved policy R_2. Specifically, For each state i, choose such decision $\delta_i(R_2)$ that makes the right-hand-side of Eq. (24) a minimum, i.e.

$$\delta_i(R_2) = \underset{k}{\operatorname{argmin}}\left\{C_{ik_{R_2}} - v_i(R_1) + \sum_{j=0}^{\infty} P_{ij}(k_{R_2})v_j(R_1)\right\} \quad \text{for all } i, \tag{25}$$

where arg $\min_k f(k)$ is the value of $k \in \{0, 1,..., K\}$ that minimizes $f(k)$. The set of the best decisions for all states $(i=0, 1,..., m)$ constitute the second, or the improved policy R_2, whose long-run expected average cost per unit time is given by Eq. (21). Repeating this iteration procedure until the two successive R's are the same. For more information of the policy-improvement algorithm, readers are referred to the book of Hillier and Lieberman (1999).

5. Application examples

Those outlined in the previous sections are powerful tools of formulating models and obtaining the optimal policies for the control of such systems that belong to the MDPs. Their application to a real inventory management problem is illustrated in this section. We will examine the demand data, study its probability distribution, designate the possible decisions and the corresponding actions taken, identify the Markov decision model related to the underlying system by defining its state space and the transition probabilities, specify the cost structure and evaluate its individual component; and then use the policy-improvement algorithm to obtain the optimal policy. The sale data shows that the demands can be subdivided into groups based on their distribution and/or their variability. Therefore we will use two products having different probability distributions and exhibiting different levels of variation for comparison.

5.1. The random demands

The database includes customer demand for the MP-PD's more than 1100 products during the period of June 1999–December 2000. To protect private commercial information, names of the products used in the paper have been changed; and the original data have also been twisted with the main features reserved. Demands for some of the products show trends and/or seasonality. However, many of them exhibit random, unpredictable and rapid changes, which renders the attempt of forecasting infeasible. Fig. 2 and Fig. 3 display the weekly demands for products $P-1$ and $P-2$. Variation of the former is relatively small—the ratio of its standard deviation to the mean is 0.5. The changes in demand for $P-2$ appear more erratic. Its standard deviation is 1.7 times of its mean value. Of the over 1100 products we studied, this ratio ranges from 0.4 to 9.1. Only one third of the products have the ratio less than one.

Examining actual frequency distributions of the weekly demand data suggests that many of them approximate either normal or exponential distribution. $P-2$, for example, fits in approximately with an exponential distribution as displayed in Fig. 4. Product $P-1$ represents another class of demands having a close-to-normal distribution (Fig. 5). For those not fitting well to either of these two, the statistics such as means and standard deviations needed in our procedure are computed from the available data.

Since the final-product inventory under consideration fits well into the framework of the MDP formulation, we have formulated it as a MDP and defined the cost function. Our objective is to seek inventory policies that enable us to maintain a high level of customer service at a minimum cost. To the best of our knowledge, this is

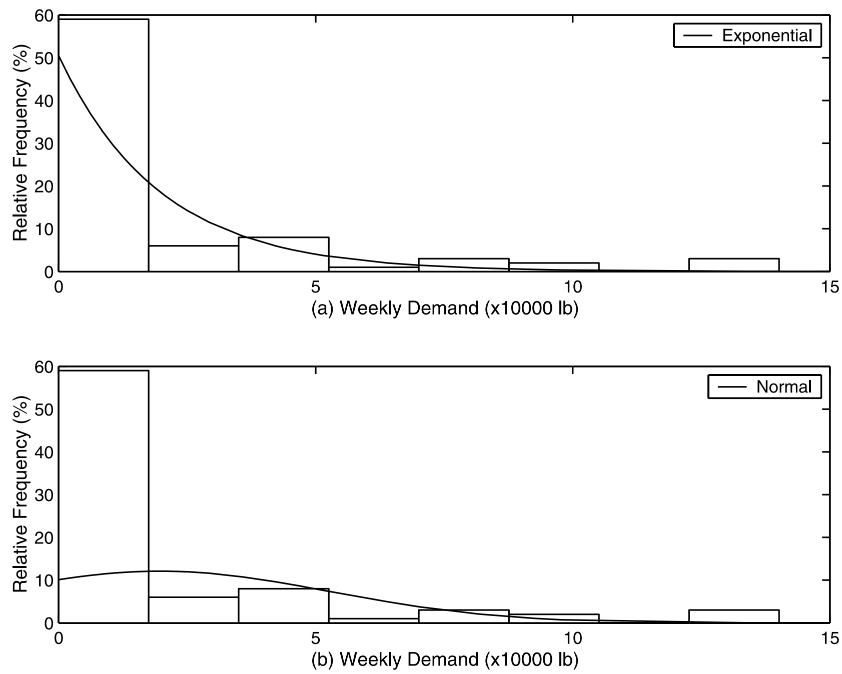

Fig. 4. Comparisons of the actual demand distribution of $P-2$ with (a) exponential; (b) normal distributions.

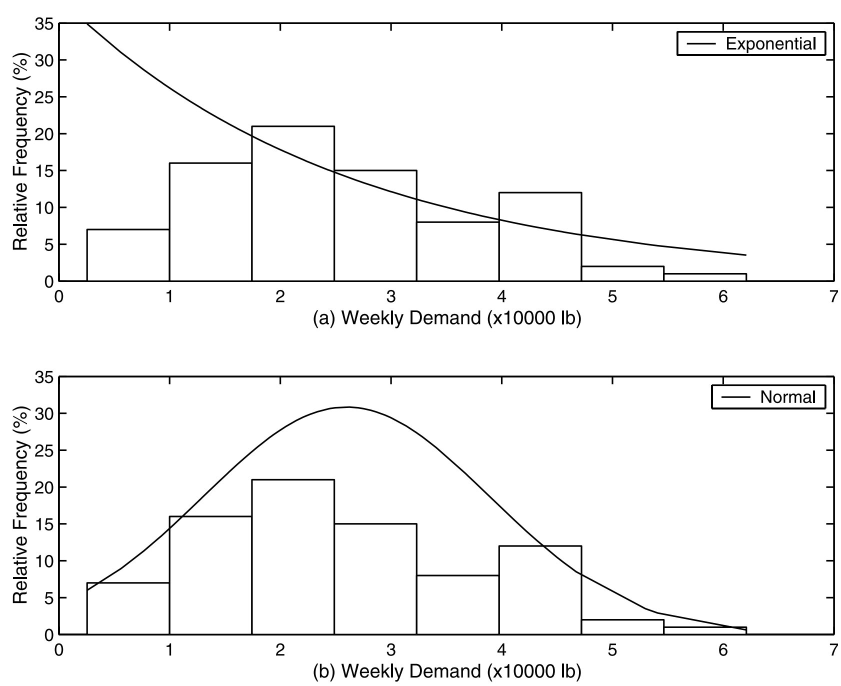

Fig. 5. Comparisons of the actual demand distribution of $P-1$ with (a) exponential; (b) normal distributions.

Table 4
States of the Markov chain for $P-2$

State	Inventory amount ($\times 10\,000$ lbs)
0	$0<\tilde{X}_i \leqslant 14.78$
1	$14.78<\tilde{X}_i \leqslant 20.72$
2	$20.72<\tilde{X}_i \leqslant 26.67$
3	$26.67<\tilde{X}_i \leqslant 32.61$
4	$32.61<\tilde{X}_i \leqslant 38.55$

Table 5
Decisions and actions

Decision	Action
0	Do not replenish
1	Replenish 59 400 lbs
2	Replenish 118 900 lbs
3	Replenish 178 300 lbs
4	Replenish 237 700 lbs

the first attempt of using MDP models in the paper industry.

5.2. Lead time, buffer stock and a replenishment model

Inventory of each product is checked weekly to determine whether and how much this item should be produced. An order is placed upon determination. It usually takes 3 weeks to fill an order. Therefore we consider the lead time to be a constant of 3 weeks herein. In calculating the stock level, it is also assumed that backlogged demand is allowed and the unsatisfied portion is transferred to the next period. In practice, it is necessary to have enough stock on hand to cover the expected sales during lead time, i.e. the expected demand during the upcoming 3 weeks. In addition, since actual demand during lead time often exceeds this expected value because of demand fluctuation, a buffer/safety stock is needed. Presumably, the size of the buffer should be determined by both the degree of variation and the required customer service level. Let D_1 and $\overline{D}_1$ be the random weekly demand and its expected value; α be the designated service level, e.g. 95%, 97%,... etc. Let c_1 be the critical value at which

$$\Pr\{D_1 \geqslant c_1\} = 1-\alpha. \tag{26}$$

In other words, the probability of the demand's being greater than c_1 is $1-\alpha$. Then the buffer stock is determined by

$$B=(c_1-\overline{D}_1)L \tag{27}$$

where L is the lead time. Note that the 'service level' is used in the determination of the buffer stock, which requires a prescribed confidence level. The higher the confidence level, the higher the buffer needed and so the higher the inventory level.

Having established the demand distribution, the lead time as well as the size of the buffer stock, we can proceed to formulate the MDP model from the available demand data and, subsequently, to develop the optimal policy by using the policy-improvement algorithm. The policy so obtained will then be examined in terms of its adequacy and performance, evaluated by the required inventory level, the occurrence of stockout, and the number of orders to be placed, as detailed in the following subsections.

For comparison, an (S, s) replenishment policy is also applied to the same system. This policy requires that if the end-of-period stock is less than s, then an amount sufficient to increase the quantity of stock on hand to the level S is ordered; otherwise, no replenishment is undertaken. The replenishment level $S=M$ is determined by

$$M=B+c_3. \tag{28}$$

Corresponding to the random tri-weekly demand D_3, c_3 in Eq. (28) is the critical value at which

$$\Pr\{D_3 \geqslant c_3\} = 1-\alpha.$$

The reorder point P is the same as the lower level s and is determined by

$$P=B-\overline{D}_1 \tag{29}$$

5.3. The MDP model and the optimal policy

In practice, the order amount is not totally arbitrary. There is usually a minimum order/production amount, u, to be used. The choice of u will affect the state space of the Markov chain and hence the final policy according to our formulation (Eq. (6) and Table 1). We choose u to be the mean value of the 3-week demand. The state space of the Markov chain for $P-2$ is shown in Table 4. The upper bound of the state 0 is the level of the buffer stock. We choose the upper bound of state 4 to be M, i.e. the replenishment level in Eq. (28).

The decisions and their corresponding actions are listed in Table 5. Orders are the minimum amount $u = 59\,400$ lbs or its multiples as discussed in the Section 3.4. The inventory level at the end of each 3-week period is calculated by

$$\text{Inventory Level} = \text{Beginning Inventory} + \text{Amount Received} - \text{Demand}.$$

As mentioned in the previous sections, under a given policy the pair of random variables $Y_n=(X_n, \Delta_n)$ forms a two-dimensional Markov chain. To evaluate a policy requires $p(j\,|i,\ k)$, the conditional probability of the chain's moving to state j at the $(n+1)$th time period given the current state i and the kth decision specified by R. In particular, for a system having 5 states and 5 different decisions, there are five transition probability matrices to be evaluated. The procedure for determining the transition probabilities is the same as that for determining the transition probabilities in Eq. (9). Eq. (30) presents the five possible transition probability matrices of this MDP of $P-2$ under the five decisions tabulated in Table 5. The elements in the matrix $P(k)$ are the transition probabilities $P_{ij}(k)$ under decision k. An entry of '–' means that the decision is not valid due to an infeasible action, e.g. the suggested replenishment amount exceeds the maximum value M. Such inadmissible decision/action will not be used.

$$P(0)=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.63 & 0.37 & 0 & 0 & 0 \\ 0.26 & 0.37 & 0.37 & 0 & 0 \\ 0.11 & 0.15 & 0.37 & 0.37 & 0 \\ 0.04 & 0.07 & 0.15 & 0.37 & 0.37 \end{bmatrix}$$

$$P(1)=\begin{bmatrix} 0.63 & 0.37 & 0 & 0 & 0 \\ 0.26 & 0.37 & 0.37 & 0 & 0 \\ 0.11 & 0.15 & 0.37 & 0.37 & 0 \\ 0.04 & 0.07 & 0.15 & 0.37 & 0.37 \\ - & - & - & - & - \end{bmatrix}$$

$$P(2)=\begin{bmatrix} 0.26 & 0.37 & 0.37 & 0 & 0 \\ 0.11 & 0.15 & 0.37 & 0.37 & 0 \\ 0.04 & 0.07 & 0.15 & 0.37 & 0.37 \\ - & - & - & - & - \\ - & - & - & - & \end{bmatrix}$$

$$P(3)=\begin{bmatrix} 0.11 & 0.15 & 0.37 & 0.37 & 0 \\ 0.04 & 0.07 & 0.15 & 0.37 & 0.37 \\ - & - & - & - & - \\ - & - & - & - & - \\ - & - & - & - & - \end{bmatrix}$$

$$P(4)=\begin{bmatrix} 0.04 & 0.07 & 0.15 & 0.37 & 0.37 \\ - & - & - & - & - \\ - & - & - & - & - \\ - & - & - & - & - \\ - & - & - & - & - \end{bmatrix} \quad (30)$$

Table 6
Iteration results

Iteration no.	Policy				
1	1	0	0	0	0
2	4	3	2	1	0
3	3	2	1	0	0
4	3	2	1	0	0

Table 7
The optimal policy

State	Inventory amount	Action
0	$0<\tilde{X}_i\leqslant 14.78$	Replenish 178 300
1	$14.78<\tilde{X}_i\leqslant 20.72$	Replenish 118 900 lbs
2	$20.72<\tilde{X}_i\leqslant 26.67$	Replenish 59 400 lbs
3	$26.67<\tilde{X}_i\leqslant 32.61$	Do not replenish
4	$32.61<\tilde{X}_i\leqslant 38.55$	Do not replenish

In an inventory system, the main costs that affect profit may include manufacturing cost, holding cost, shortage cost, salvage cost, and discount rates, etc. In this work we consider two components, an average manufacturing cost of $\$C_{\mathrm{m}}$/lb and an average shortage cost of $\$C_{\mathrm{s}}$/lb. The manufacturing costs include all those incurred during the production. The shortage costs are resulted from lost of sales due to insufficient inventories. Let d denote the demand. Let Q_k be the ordering amount under decision k; and $\tilde{X}_i$ be the average inventory amount when the system is in state i. Then the expected cost

$$C_{ik}=C_m Q_k + C_s \sum_{j=0}^{m} \max\{(d-\tilde{X}_i),0\} P_{ij}(k)$$

$$\text{for } i=0,\ldots 4,\ k=0,\ldots 4 \quad (31)$$

where $P_{ij}(k)$ is given in Eq. (30).

The choice of the initial policy R_1 is arbitrary. Using the policy-improvement algorithm as outlined in Section 4 usually can yield the optimal policy in only a few iterations. Table 6 presents the intermediate and final results when the first policy calls for ordering 59 400 lbs if the system is in state 0 and no ordering placed otherwise.

By using the Markov decision model and the policy-improvement algorithm, we obtain the optimal policy shown in Eq. (32), in which an entry $D_{ik}=1$ means that the policy calls for the kth decision if the system is in state i. For example, $D_{21}=1$ means that decision 1 and its corresponding action (to replenish 59 400 lbs) will be exercised if the system is in state 2 (stock at hand is between 207 200 and 266 700 lbs).

$$D_{P-2}=\begin{bmatrix} D_{00} & D_{01} & D_{02} & D_{03} & D_{04} \\ D_{10} & D_{11} & D_{12} & D_{13} & D_{14} \\ D_{20} & D_{21} & D_{22} & D_{23} & D_{24} \\ D_{30} & D_{31} & D_{32} & D_{33} & D_{34} \\ D_{40} & D_{41} & D_{42} & D_{43} & D_{44} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (32)$$

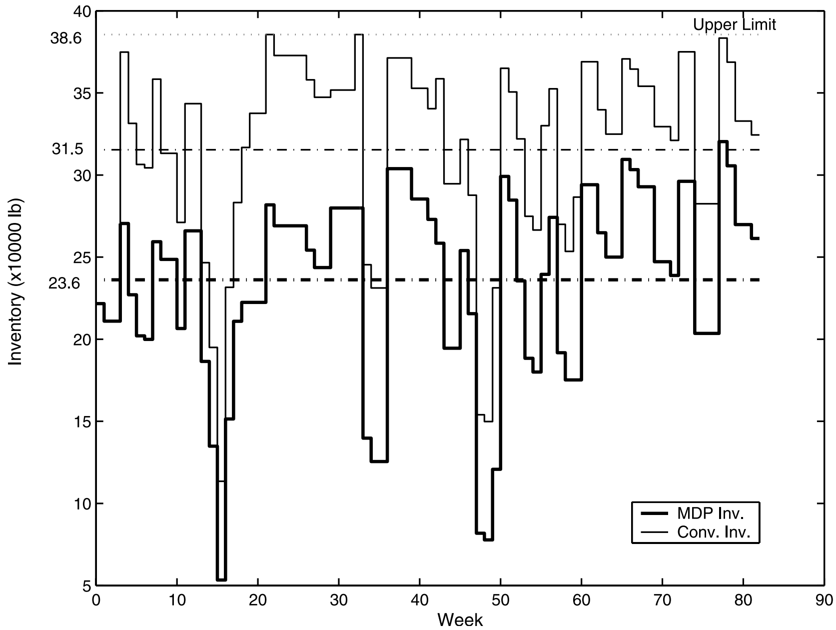

Fig. 6. A comparison of the MDP policy and the conventional replenishment policy—Product: $P-2$; Service level: 97.6%.

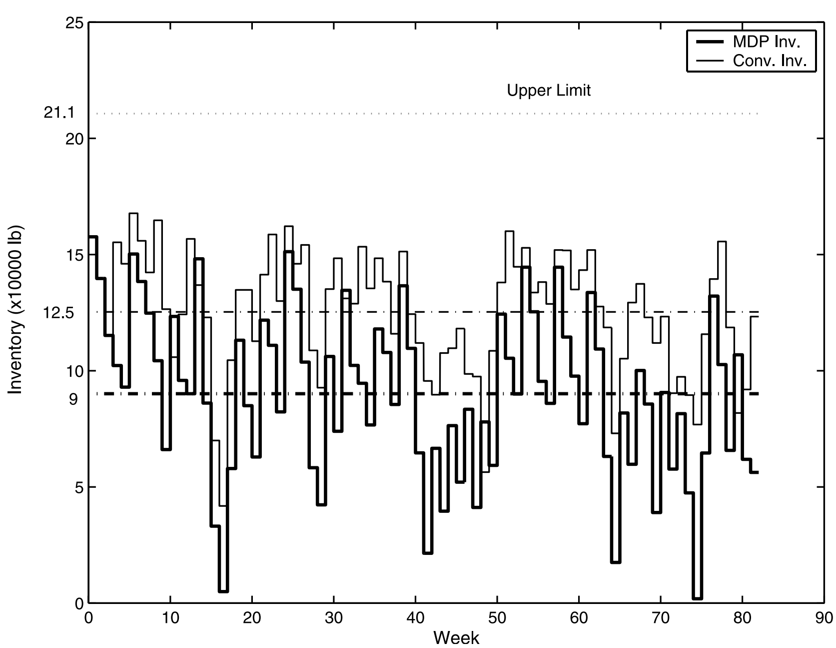

Fig. 7. A comparison of the MDP policy and the conventional replenishment policy—Product: $P-1$; Service level: 98%.

Table 7 provides a verbal description of this optimal policy, which prescribes the decision/action to be made under all possible conditions.

It can be seen that the concepts and development of the MDP model are more involved than most frequently used inventory models. Once a policy as depicted in Table 7 becomes available, however, its implementation is rather straightforward. It can usually provide us with better results as illustrated in Fig. 6 and Fig. 7.

Fig. 6 compares the MDP policy with the conventional (S, s) (Eq. (28) and Eq. (29)) inventory policy based on the real demand data for product $P-2$, of which the weekly mean was 19 800 lbs. Its actual weekly demand ranges from the lowest value of zero to the highest 140 000 lbs. Such high variability makes the prediction of future demand very difficult. Using conventional policies often leads to stockout or requires a higher inventory level. Since the randomness has been included in the MDP model, better results can be expected. It indeed is the case as shown in our comparison.

Fig. 6 shows that no stockout occurs under either policy. The same upper bound $M = 386\,000$ lbs was applied to both methods in their policy development. However, the MDP policy results in an average inventory level of 236 000 lbs that is much lower than the level of 315 000 lbs required by the conventional replenishment policy. The service level in the figure is a prescribed number. A service level of $\alpha = 97\%$ means that the allowed probability of stockout is $1 - 97\% = 3\%$. The calculation shows that using MDP method, a total of 21 reorders have been made during the 83-week period, which is also lower than the 28 times required in the conventional method.

For the low-variability class, the MDP policy also performs better however not as drastic as the high-variability one. Fig. 7 shows that no stockout in either case, the average inventory level of 90 000 lbs from MDP policy is lower than the 125 000 lbs required by the conventional replenishment policy. Its reorder time of 27 is also lower than the 52 times required for the conventional method.

It is conceivable that the inventory level will be reduced if a small amount of stockout is allowed. As displayed in Fig. 8, a change of the service level from 97.6% to 96.5% results in a lowered inventory level from 236 000 lbs (Fig. 6) to 189 000 lbs, corresponding to a 25% inventory savings. Its trade-off is a $3/82 = 3.65\%$ stockout. Fig. 9 shows that when a $2/82 = 2.4\%$ stockout is allowed to happen, a 10% savings in inventory are achieved for $P-1$.

Same procedure as above was applied to several other products. Similar conclusions were obtained. Of the two different methods, the MDP system consistently yielded better policies than the traditional replenishment ones. In general, it results in lower average inventory level and/or lower stockout and requires fewer reorders to be placed. These advantages are more pronounced for products with highly variable demands, for which most of other methods do not perform well. The computation time needed to develop an MDP policy and to obtain results shown in, e.g. Fig. 7, was 19 s using a Dell 700 MHz PC.

6. Summary and discussion

This paper concerns Markovian inventory policy and its application in the paper industry. The Markov decision models have proved to be useful in many systems and are powerful tools for inventory planning. After discussing the model formulation and the solution procedure in general, we apply it to a problem of inventory control of finished paper products. Emphases have been put on several key steps in the model development, such as obtaining the state space of the Markov chain, designating the possible decisions and actions, calculating the transition probabilities, defining the cost structure and evaluating the cost function, and determining the optimal policy. Such practical issues as lead time, inventory level, stockouts are discussed in detail. We have simplified the decision-making procedure by discretizing the continuous demands into discrete levels.

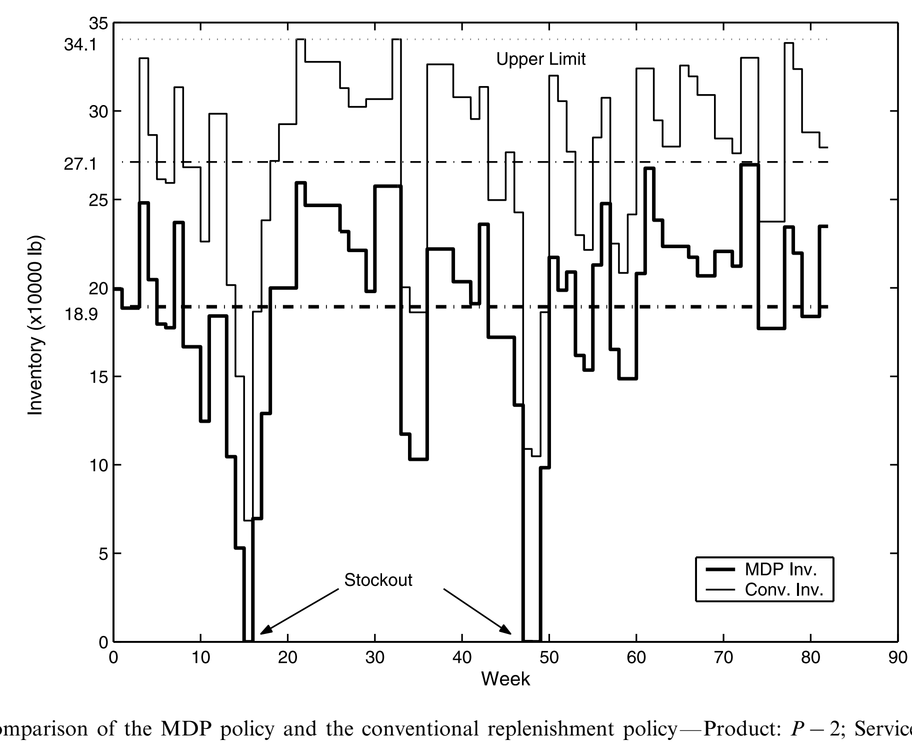

Fig. 8. A comparison of the MDP policy and the conventional replenishment policy—Product: $P-2$; Service level: 96.5%.

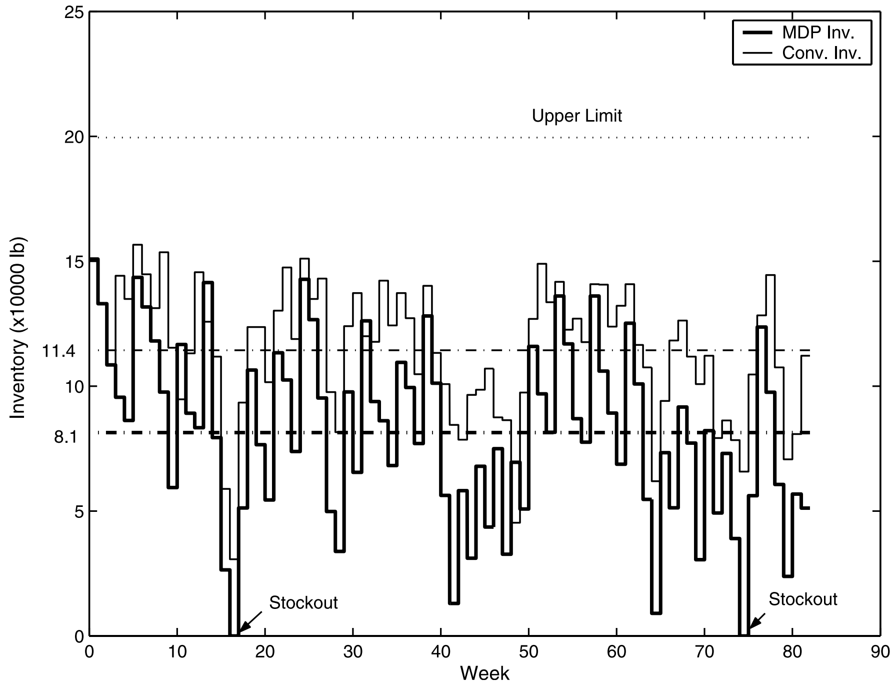

Fig. 9. A comparison of the MDP policy and the conventional replenishment policy—Product: $P-1$; Service level: 97%.

Such treatment is applicable to many cases in chemical industry where the products are measured by their weights. Our calculation results have shown that the MDP models consistently provide better results than conventional models, especially for systems exhibiting high variabilities. When there are trends in the system, the mean is not a constant, and the covariance is a function depending not only on the time lag but rather in a much more complex fashion. Such non-homogeneous MDP can be treated using methods discussed in Yin and Zhang (1998).

It is well known that many time series, particularly sale data, often show seasonality. Building seasonal models and using them for forecasting have become routine in inventory planning. This issue was not addressed herein because the data we used are not seasonal. If the series shows a marked seasonal pattern, the methods developed above need to be modified.

Forecasting product demand with mathematical models derived from historical data has been a common practice in industries. This can lead to reasonably good prediction provided that the demand variation is relatively small, or that certain trends such as seasonality are easily identifiable. The basic assumption of forecasting is that markets and demands are predictable within certain accuracy. Given the rapid and often unpredictable changes in today's global economy, numerous uncertainty involved in the dynamic process, as well as the complicated relationships among the elements and links of any given supply chain, in many cases this assumption is not true, which renders the forecasting unreliable or totally mistaken. As a result, production and inventory decisions can no longer be made solely based on forecasting alone. Other tools are needed to adequately address the variability issue. MDP models offer a good alternative for such purpose.

Stochastic modeling and simulation have become frequently used and powerful tools in quantifying dynamic relations of sequences of random events and uncertainties. Considering the complex and random natures of many chemical processes, yet their still limited use of stochastic modeling and simulation, it is conceivable that Markov chain and stochastic modeling will find more applications in the chemical engineering field.

References

Applequist, G., Samikoglu, O., Pekny, J., & Reklaitis, G. (1997). Issues in the use, design and evolution of process scheduling and planning systems. *ISA Transactions*, *36*(2), 81–121.

Arrow, K. J., Karlin, S., & Scarf, H. (1958). *Studies in the mathematical theory of inventory and production*. Stanford, CA: Stanford University Press.

Bassett, M. H., Pekny, J. F., & Reklaitis, G. V. (1997). Using detailed scheduling to obtain realistic operating policies for a batch processing facility. *Industrial & Engineering Chemical Research*, *36*(5), 1717–1726.

Buchan, J., & Koenigsberg, E. (1963). *Scientific inventory management*. Englewood Cliffs, NJ: Prentice-Hall.

Buffa, E. S. (1980). *Modern production/operations management* (6th ed.). New York: Wiley.

Chiang, C. L. (1980). *An introduction to stochastic processes and their applications*. New York: Robert E. Krieger Publishing Co.

Chung, K. L. (1960). *Markov chain with stationary transition probabilities*. Berlin: Springer.

Doob, J. L. (1953). *Stochastic processes*. New York: Wiley.

Feller, W. (1971). *An introduction to probability theory and its applications*. New York: Wiley.

Fisher, M. L., Raman, A., & McClelland, A. S. (2000). Rocket science retailing is almost here—are you ready? *Harvard Business Review*, *July–August*, 115–125.

Gupta, A., Maranas, C. D., & McDonald, C. M. (2000). Mid-term supply chain planning under demand uncertainty: customer demand satisfaction and inventory management. *Computers & Chemical Engineering*, *24*(12), 2613–2621.

Hillier, F. S., & Lieberman, G. J. (1999). *Introduction to operations research* (6th ed.). Oakland CA: Holden-Day, Inc.

Johnson, L. A., & Montgomery, D. C. (1974). *Operations research in production planning and inventory control*. New York: Wiley.

Kushner, H. J., & Yin, G. G. (1997). *Stochastic approximation algorithms and applications*. New York: Springer.

Pekny, J., & Miller, D. (1990). Exact solution of the no-wait flowshop scheduling problem with a comparison to heuristic methods. *Computers & Chemical Engineering*, *14*(9), 1009–1023.

Petkov, S. B., & Mararas, C. D. (1997). Multiperiod planning and scheduling of multiproduct batch plants under demand uncertainty. *Industrial & Engineering Chemistry Research*, *36*(11), 4864–4881.

Ross, S. (1983). *Introduction to stochastic dynamic programming*. New York: Academic Press.

Starr, M., & Miller, D. (1962). *Inventory control*: *theory and practice*. Englewood Cliffs, NJ: Prentice-Hall.

Taylor, H. M., & Karlin, S. (1998). *An introduction to stochastic modeling*. Boston: Academic Press.

Veinott, A. F. (1965). The optimal inventory policy for batch orderings. *Operations Research*, *13*(3), 424–432.

Wilson, R.H. (1934). A scientific routine for stock control. *Harvard Business Review*, 13.

Yang, H., Yin, G., Yin, K., & Zhang, Q. (2002). Control of singularly perturbed Markov chains: a numerical study, to appear in *The ANZIAM Journal (formerly known as Journal of the Australian Mathematical Society*, *Series B: Applied Mathematics).*

Yin, G., Zhang, Q., Yang, H., & Yin, K. (2001). Discrete-time dynamic systems arising from singularly perturbed Markov chains. *Nonlinear analysis: theory*, *methods & applications*, *47*(7), 4763–4774.

Yin, G. G., & Zhang, Q. (1998). *Continuous Markov chains and applications*, *a singular perturbation approach*. New York: Springer.

Yin, G. G., & Zhang, Q. (1997). *Mathematics of stochastic manufacturing systems*. Providence: American Mathematical Society.

Yin, K., Yin, G. & Zhang, Q. (1995). Approximating the optimal threshold levels under robustness cost criteria for stochastic manufacturing systems, *Proceeding IFAC Conference of Youth Automation YAC '95*, 450–454.