Int J Adv Manuf Technol (2005) 26: 1184–1192
DOI 10.1007/s00170-004-2069-8

ORIGINAL ARTICLE

C.O. Kim · J. Jun · J.K. Baek · R.L. Smith · Y.D. Kim

Adaptive inventory control models for supply chain management

Received: 20 October 2003 / Accepted: 11 November 2003 / Published online: 7 July 2004

Abstract Uncertainties inherent in customer demands make it difficult for supply chains to achieve just-in-time inventory replenishment, resulting in loosing sales opportunities or keeping excessive chain-wide inventories. In this paper, we propose two adaptive inventory-control models for a supply chain consisting of one supplier and multiple retailers. The one is a centralized model and the other is a decentralized model. The objective of the two models is to satisfy a target service level predefined for each retailer. The inventory-control parameters of the supplier and retailers are safety lead time and safety stocks, respectively. Unlike most extant inventory-control approaches, modelling the uncertainty of customer demand as a statistical distribution is not a prerequisite in the two models. Instead, using a reinforcement learning technique called action-value method, the control parameters are designed to adaptively change as customer-demand patterns changes. A simulation-based experiment was performed to compare the performance of the two inventory-control models.

Keywords Adaptive inventory control · reinforcement learning · simulation · supply chain

C.O. Kim (✉)
School of Computer and Industrial Engineering,
Yonsei University,
Seoul, 120-749, Republic of Korea
E-mail: kimco@yonsei.ac.kr
Tel: (+822) 2123-2711
Fax: (+822) 364-7807

J. Jun
Research Institute for Information and Communication Technologies,
Korea University,
Seoul, 136-701, Republic of Korea

J.K. Baek
Department of Industrial System Management,
Seoul College,
Seoul, 131-702, Republic of Korea

R.L. Smith
Department of Industrial and Operations Engineering,
University of Michigan,
Ann Arbor, Michigan 48109-2117, USA

Y.D. Kim
Department of Industrial Engineering,
Korea Advanced Institute of Science and Technology,
Daejon, 305-701,
Republic of Korea

1 Introduction

In supply-chain management, the effort of minimizing total costs in terms of reduction in chain-wide inventory has been increasingly addressed and attempted in industry. During the last two decades, however, achieving this objective has been more difficult, as customer demands become more diverse and the life cycles of products are shorter. In most cases, due to unpredictable customer needs and economic situations, customer demands fluctuate with time, showing nonstationary patterns. Uncertainties inherent in customer-demand patterns make it difficult to satisfy customer demands in just-in-time (JIT) mode, resulting in loosing sales opportunities or keeping excessive chain-wide inventories.

In modelling inventory-control problems, it is not practical to assume that customer demands during a period are known a priori in the form of a constant or a statistical distribution. In this respect, adaptive inventory control in supply-chain management should be addressed. By adaptive, we mean that the control parameters of inventory-control models are dynamically adjusted toward satisfying a target service level with the consideration of the nonstationarity of customer demand. The target service level means the percentage of customer demands that have to be satisfied during the time interval between order placement time and inventory replenishment time. This time interval is commonly called lead time.

In this paper, we deal with a two-echelon supply-chain system consisting of one supplier and multiple retailers. The customer demand process is assumed to be nonstationary and unknown. By a nonstationary demand process, we mean that the mean and variance of the demand distribution changes with time. It is assumed that the supplier's orders are always satisfied after a constant lead time from a perfectly reliable single outside source. It is also assumed that, for each retailer, transportation lead time from the supplier to the retailer is given as a constant. However, the retailers' actual lead times are not constants unless the supplier has enough inventory to meet the retailers' orders. Finally, if customer demands are not satisfied at sales points of time, the demands are treated as lost sales.

In this environment, we propose two adaptive inventory-control models for the supply chain: a centralized model and a decentralized model. In both models, the supplier makes use of on-line information about the retailers' inventory status in deciding his order placement time. The goal is to make the average of service levels during lead times as close as possible to a predefined target service level. We assume that the order size of each retailer is predetermined based on the capacity of the delivery system. Therefore, associated decisions are concerned with when the retailers' inventories are replenished. The control parameters of the two models determine the inventory replenishment times.

The centralized inventory-control model is similar to the vendor-managed inventory model [1], in the sense that the retailers no longer control their inventory replenishment times. Instead, the supplier is responsible for maintaining appropriate inventory levels of the retailers. In more detail, at each discrete inspection time, the supplier collects data on each retailer's inventory position (on-hand inventory plus ordered quantity in transition) and sales history. With the data, the supplier makes use of a linear time-series model to predict the time point at which the inventory position of the retailer is anticipated to drop down below zero at first. If the time interval between the inspection time and the predicted time is approximately close to total lead time (supplier's lead time + retailer's transportation lead time + safety lead time), then the supplier places an order for the retailer. In this paper, we call this inventory-control rule a JIT delivery policy. As soon as the supplier receives the ordered quantity from the outside source, he sends the quantity directly to the retailer without keeping it in his warehouse. As a consequence, the inventory level of the supplier becomes completely zero. The safety lead time is a time buffer and is traditionally used for coping with demand uncertainty during lead time. In the centralized model, it is a control parameter to adjust the service level of the retailer in a nonstationary demand situation. The safety lead time exists for each retailer.

In the decentralized inventory-control model, each retailer is allowed to adaptively set safety stock by reflecting the nonstationarity of demand. Just as for the safety lead time, the safety stock is an inventory buffer to cover demand uncertainty during lead time. Once the safety stock is decided, the retailer forecasts demand during its transportation lead time at each inspection time. If, at a certain inspection time, demand during the transportation lead time plus safety stock is very close to its inventory position observed at the inspection time, the retailer places an order to the supplier.

The supplier also makes use of the safety stock set by the retailer. The operational mechanism of the supplier in the decentralized model is almost the same as the JIT delivery policy in the centralized model. That is, at each inspection time, the supplier predicts the time point at which the inventory position of the retailer is anticipated to drop below the safety stock at first. If the time interval between the inspection time and the predicted time is approximately close to total lead time, then the supplier places an order for the retailer. However, unlike the centralized model, after the supplier receives the ordered quantity from the outside source, he keeps the quantity in his warehouse until the retailer actually places an order. In the decentralized model, safety stock and safety lead time are control parameters.

Using a reinforcement learning technique, the control parameters of the two models are designed to adaptively change. The reinforcement learning technique employed in this research is called the action-value method [2], which is suitable to heuristically solve sequential optimization problems in uncertain environments. A representative domain appropriate for applying the action-value method is the stochastic optimization problem, where the value of each action is not known but should be learned through repetitive applications of the action in a real or simulated domain. The main advantage of the reinforcement learning is that it is possible to make good decisions while the learning is progressing. In this respect, reinforcement learning would be appropriate for applying to real-time control problems.

Specifically, at each decision point of time, one of the possible actions is selected based on a probabilistic function of their value estimates. In minimization problems, it is desirable to give more opportunity of being selected to the actions with low value estimates. This idea can be incorporated into the following probabilistic action selection rule:

$$P\{new\ action = a\} = \frac{\mathrm{e}^{-ValueEstimate(a)}}{\sum_{a_i \in AS} \mathrm{e}^{-ValueEstimate(a_i)}} \tag{1}$$

where AS is the set of possible actions. Because the numerator, $\mathrm{e}^{-ValueEstimate(a)}$, in Eq. 1 increases as the value estimate of the action $ValueEstimate(a)$ decreases, the action with the lowest estimated value would be selected with the highest probability. The denominator is a normalization term to make the action selection rule be a probability function.

The result of the selected action (current value) is then used for learning its objective value. The learning formula we employ is called the exponential recency weighted average (pp. 37, Sutton and Barto [2]) and can be defined as

$$\begin{aligned} &NewValueEstimante \leftarrow OldValueEstimate \\ &+ StepSize[CurrentValue - OldValueEstimate] \end{aligned} \tag{2}$$

Each time a specific action is performed, its new value estimate is updated by adding an error (weighted difference of the current value and the old estimate) to the old estimate. The error indicates a desirable direction to which the value estimate moves. *StepSize* is a learning parameter that decides learning speed. It is normally set to a constant, such as 0.1, which has been experimentally verified to be desirable, especially in nonstationary environments (pp. 39, Sutton and Barto [2]). At the next decision point of time, a new action is chosen according to the probabilistic rule with the updated value estimate, and this procedure is repeated until the end of the decision horizon is reached.

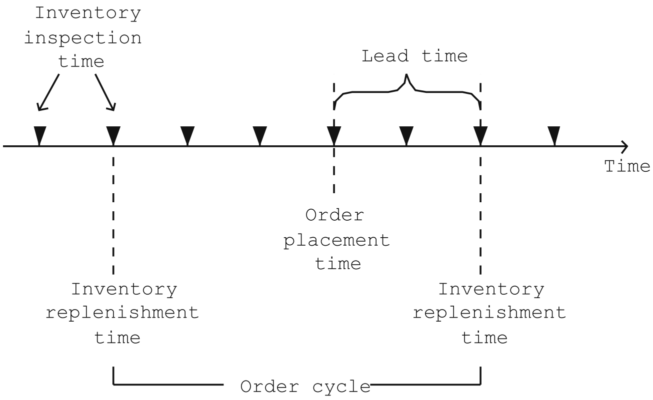

Fig. 1. Inventory replenishment time and order cycle in periodic inspection system

In the context of the problem discussed in this paper, the action corresponds to a control parameter (safety stock or safety lead time). As shown in Fig. 1, for each retailer, the decision point of time implies inventory replenishment time. The time interval between two consecutive inventory replenishment times is called an order cycle. When a new safety stock is selected at an inventory replenishment time, the service level during the lead time is measured at the end of the order cycle. The result of the safety stock is then defined as the absolute deviation of the service level from a target service level. After that, the value estimate of the safety stock is updated according to Eq. 2. The *NewValueEstimate* in Eq. 2 means the weighted average of the absolute deviations of service levels during lead times from the target service level. Therefore, as learning progresses, safety stocks with low service-level deviations will be given high selection probabilities.

The remainder of this paper is organized as follows. In Sect. 2, we review extant inventory-control methods relevant to our models. In Sect. 3, we present the two inventory control models. In Sect. 4, we present the results of a simulation-based performance evaluation. Finally, in Sect. 5, we conclude this research and remark on some future research areas.

2 Literature survey

Related to adaptive inventory control in supply chains, most previous research efforts have been occurred in the mathematical production control area [3–6]. With the objective of minimizing total sum of inventory costs, they formulate the inventory-control problem as a dynamic programming model and adaptively estimate the uncertain parameters of demand distribution using demand history. While the rigorous optimization models show some mathematical convergence results in stationary demand cases, it is unsupported in many applied contexts in which customer demand processes are nonstationary.

The idea of our approach is similar to Packer [7]. He considered the (Q, R) inventory policy in single-site inventory-control problems. He suggested a way of taking advantage of using demand history to decrease inventory-related costs. Specifically, order quantity Q is calculated with the economic order quantity (EOQ) model, for which average demand rate is estimated by the exponential smoothing formula. Then the average demand during lead time and a predetermined safety-stock factor are used for setting reorder point, R.

Moinzadeh [8] proposed a supplier replenishment policy in which an order is placed to the outside source immediately after a retailer's inventory position reaches $R+s$. Thus, s, in a sense, gauges the proactivity of the supplier from information availability. He computationally derived the optimal s under the assumption that customer demand at the retailers is Poisson.

Recently, a few distributed inventory-control models have been proposed. Axsater [9] applied the Stackelberg game model to the decentralized control of a multiechelon inventory system consisting of a central warehouse and multiple retailers. Also, by employing penalty cost concepts, Andersson et al. [10], Lee and Whang [11] and Cachon and Zipkin [12] attempted to distribute decision-making rights to the participants in a supply chain while enforcing each participant to respect others' costs. However, all of them analytically solved the distributed problems under the assumption that the customer demand distribution is known.

Zhao et al. [13] proposed a retailer's early order commitments rule in a decentralized supply chain for enabling suppliers to smooth production, better utilize resources and ultimately reduce costs in the whole supply chain. They investigated the impact of the early-order commitment rule and forecasting models on the supply-chain performance under different scenarios of demand patterns and supplier's capacity tightness. Zhao and Xie [14] also examined the impact of forecasting errors on the value of information sharing between a supplier and retailers. They experimentally showed that, although information sharing gives benefits to the supplier, in most cases, it increases the retailers' costs. This phenomenon becomes more distinctive as the magnitude of the forecasting error increases.

Finally, a reinforcement learning approach was recently applied to a coordination and integration problem of multinational corporations with emphasis on logistics and production management [15]. The problem was formulated as a semi-Markov decision model and solved via a reinforcement learning technique called Q-learning.

3 The inventory-control models

3.1 Notations

We define the following notations to explain the two inventory-control models.

L_0 :	lead time of supplier
L_i :	transportation lead time of retailer i $(i = 1, 2, \dots, N)$
$D_i(t)$:	customer demand of retailer i at inspection time t
Q_i :	order quantity of retailer i
s_{ij} :	jth safety factor of retailer i in the centralized model
st_{ij} :	safety lead time of the supplier when safety factor s_{ij} is applied in the centralized model

$\hat{\sigma}_\varepsilon(t)$: standard deviation of forecast errors estimated at inspection time t

$\overline{SL}(s_{ij})$: average service level of retailer i when safety factor s_{ij} is applied

s^r_{ij} : jth safety factor of retailer i in the decentralized model

ss^r_{ij} : safety stock of retailer i when safety factor s^r_{ij} is applied in the decentralized model

s^s_{ij} : jth safety factor of the supplier in the decentralized model

st^s_{ij} : safety lead time of the supplier when safety factor s_{ij} is applied in the decentralized model

3.2 The centralized model

3.2.1 The JIT delivery policy

The supplier monitors the inventory position of retailer i ($i = 1, 2, \ldots, N$) and sales history at each discrete inspection time. With the sales history data, the supplier updates linear time series model $\hat{D}_i(t') = a_0 + a_1 t'$. This model is used for estimating the amount of customer demand at future time t'. In the model, coefficients a_0 and a_1 are also updated using the exponential smoothing method (see Brown [16] for detailed update formula).

At each inspection time t, the inventory position of retailer i at future time t' is defined as the inventory position observed at the inspection time t minus the sum of the estimated demands during the time interval between the inspection time t and the future time t'. Now suppose that the future time t' is set to the time at which the inventory position of retailer i falls to zero. Then the JIT delivery policy can be briefly stated as follows:

At inspection time t, if $\{$*(the time* t' *that the time series model predicts the inventory position of retailer i reaches its zero)* $- t\}$ $\leq L_0 + L_i + st_{ij}$, *then the supplier issues an order of* Q_i *to the outside source.* (3)

3.2.2 Adaptive control of safety lead time

If the demand process is stationary and its variance is very small, the forecasting model will accurately estimate demand during the total lead time. As a result, the JIT policy can replenish retailers' inventories at the time the inventory positions of retailers are close to zero. However, a problem arises if the retailers encounter sudden genuine changes in the underlying demand processes in terms of the changes of mean and (or) variance, resulting in the overestimation (or underestimation) of the demand. For example, suppose that, for retailer i, demand during the total lead time is underestimated. Then the ordered quantity triggered by the JIT policy will be delivered to the retailer after the retailer's inventory becomes a minus level, resulting in significant loss of sales. Therefore, it is necessary to expedite the order process. On the other hand, if the demand is overestimated, the order process should be delayed in order to avoid holding more inventory than needed. Of course, forecasting errors can be reduced to some extent with more sophisticated time series models. However, the models cannot fundamentally resolve the problem of forecasting errors generated due to the change of demand process.

Safety lead time can adjust order placement time. For example, suppose that demand is underestimated. In this case, as shown in Fig. 2a, adding a positive safety lead time to actual lead time (supplier's lead time + retailer's transportation lead time) enforces the JIT policy to place an order earlier than the policy without the safety lead time. Because the delivery of an ordered quantity takes the actual lead time, the JIT policy considering positive safety lead time brings the effect of expediting the order process. Similarly, forecasting with a negative safety lead time will delay order processes, and this is effective when demand is overestimated (see Fig. 2b). The reinforcement learning formulas in Eq. 1 and Eq. 2 are used for determining appropriate safety lead time.

In general, the safety lead time can be obtained from a multiplication function of lead time and forecast error [17]. Let $S_i = \{s_{i1}, s_{i2}, \ldots, s_{ik}\}$ is the set of safety factors for retailer i, in which some s_{ij}, $j = 1, 2, \ldots, k$, may have negative values. Then as shown in Fig. 3, at inventory replenishment time t, safety lead

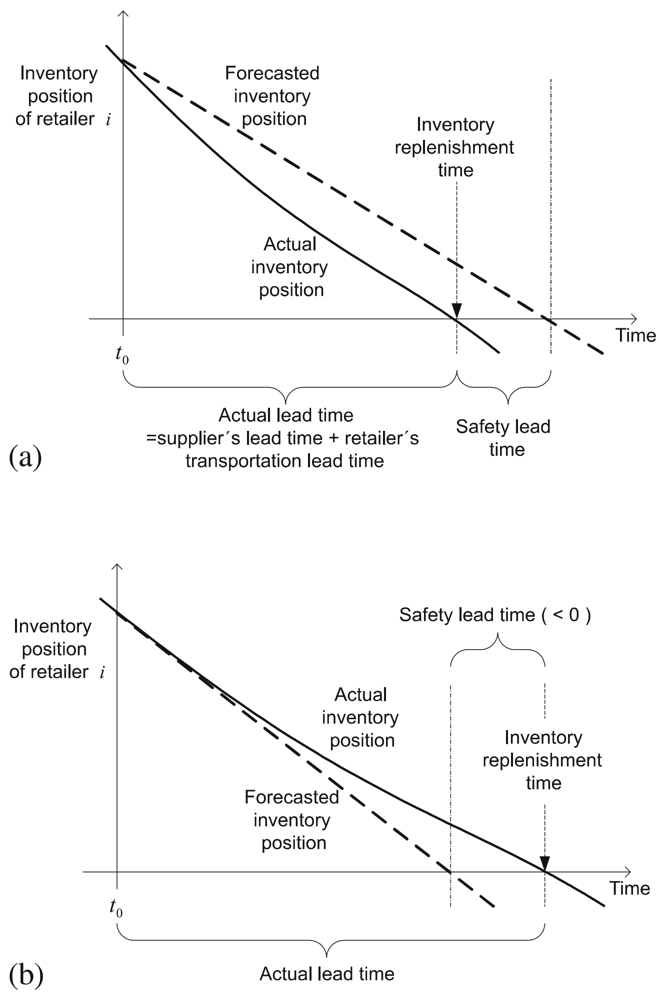

Fig. 2. The role of safety lead time in the JIT delivery policy

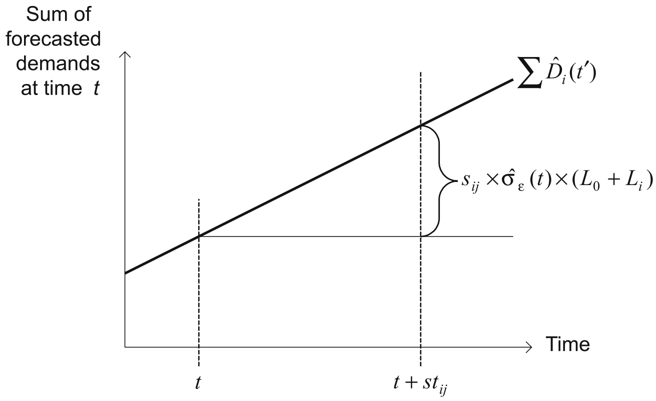

Fig. 3. Derivation of supplier's safety lead time

time st_{ij} corresponding to safety factor s_{ij} can be derived from

$$\text{Find } st_{ij} \text{ such that } \sum_{t'=t}^{t+st_{ij}} \hat{D}_i(t') = s_{ij} \times \hat{\sigma}_\varepsilon(t) \times (L_0 + L_i)$$

where estimated standard deviation of forecast errors $\hat{\sigma}_\varepsilon(t)$ is commonly approximated as 1.25× mean absolute deviation (see Brown [16] for detailed justification about the approximation).

Suppose that some s_{ij} is selected at an inventory replenishment time and an order for retailer i is placed at inspection time t according to the JIT policy specified in Eq. 3. The ordered quantity will be delivered to retailer i at time $t + L_0 + L_i$. Then average service-level estimate of s_{ij}, $\overline{SL}(s_{ij})$, is updated by the following reinforcement learning rule:

$$\begin{aligned}\overline{SL}_{new}(s_{ij}) = {} & \overline{SL}_{old}(s_{ij}) \\ & + StepSize[service\ level\ during\ \ (t, t + L_0 + L_i) \\ & - \overline{SL}_{old}(s_{ij})] \end{aligned} \tag{4}$$

If $\overline{SL}_{new}(s_{ij})$ moves to target service level, then s_{ij} can be regarded as an appropriate safety factor for the current demand pattern. Hence, the selection chance of s_{ij} at the next inventory replenishment time should be increased. To reflect this idea, the next safety factor is determined according to the following rule:

$$P\{new\ safety\ factor = s_{ij}\} = \frac{e^{-(SL_{new}(s_{ij}) - TSL)^2}}{\sum_k e^{-(SL_{new}(s_{ik}) - TSL)^2}} \tag{5}$$

where TSL is a target service level.

The completed inventory-control procedure of the centralized model is explained as follows:

<u>Supplier</u>

Step 0. The supplier selects safety factor s_{ij} initially for each retailer i ($i = 1, 2, \ldots, N$).

Step 1. At inspection time t, if order placement condition 3 is satisfied for retailer i, then the supplier issues an order of size Q_i to the outside source.

Step 2. If the ordered quantity, Q_i, has arrived from the outside source, then the supplier immediately delivers Q_i to retailer i.

Step 3. After retailer i receives the ordered quantity, Q_i (inventory replenishment time), the supplier updates $\overline{SL}_{new}(s_{ij})$ according to the learning formula in Eq. 4. The next safety factor is selected according to the probabilistic rule in Eq. 5. Set $s_{ij} =$ the next safety factor. Go to Step 1.

<u>Retailer</u>

Do nothing.

3.3 The decentralized model

In this model, the supplier and retailers are allowed to control safety lead time and safety stocks, respectively. Detailed control procedures of the safety stock and safety lead time are as follows.

3.3.1 The retailers

Retailer i ($i = 1, 2, \ldots, N$) places an order to the suppler when the inventory position reaches its reorder point. The reorder point is controlled by safety stock. If service levels during successive lead times are much lower than target service levels, retailer i sets its reorder point high by adding large amounts of safety stock, which results in the expedition of the order process. On the other hand, in the case that service levels during successive lead times are higher than the target service level, the reorder point is set to a low one by adding a small amount of safety stock.

In general, safety stock is used for covering the variation of demand during the retailer's lead time. Under the assumption that lead time is fixed, safety stock is commonly given as

$$\begin{aligned}\text{Safety stock} = {} & \text{safety factor} \\ & \times \text{estimated standard deviation of forecast error} \\ & \times \text{retailer's lead time}\end{aligned}$$

In the decentralized model, however, retailer i's lead time is not fixed. Rather, it is equal to its transportation lead time plus waiting time incurred in the case that the supplier has not enough inventory to satisfy retailer i's order. Due to this problem, we approximate safety stock of retailer i as

$$\begin{aligned}\text{Safety stock} \approx {} & \text{safety factor} \\ & \times \text{estimated standard deviation of forecast error} \\ & \times (\text{the latest actual lead time of retailer})^\beta\end{aligned}$$

where unknown actual lead time is replaced by a function of the latest actual lead time and β measures the rate at which the safety stock increases as the latest actual lead time increases. (See Bernard [17] for the detailed role of β in determining safety stock.)

Suppose that the jth safety factor of retailer i, s^r_{ij}, is selected at inventory replenishment time t. The corresponding safety stock ss^r_{ij} is given as $s^r_{ij} \times \hat{\sigma}_\varepsilon(t) \times$ (the latest actual lead time of retailer)$^{\beta}$. Then retailer i places an order if the following condition is satisfied:

At inspection time t, if
$\{$*estimated demand during*$[t, t+$ *transportation lead time of retailer i*$] +$ *safety stock* $\geq$ *inventory position of retailer i*$\}$*,*
then retailer i places an order of Q_i *to the supplier.* (6)

After retailer i receives the ordered quantity at inventory replenishment time t, $\overline{SL}(s^r_{ij})$ is updated as

$$\overline{SL}_{new}(s^r_{ij}) = \overline{SL}_{old}(s^r_{ij}) + StepSize\big[service\ level\ during(t, t+L^a_i) - \overline{SL}_{old}(s^r_{ij})\big] \quad (7)$$

where L^a_i is the time interval between order placement time and the time at which the inventory of retailer i is replenished (actual lead time). After that, the next reorder point is determined according to the probabilistic rule in Eq. 5.

The operational steps of retailer are as follows:

Retailer

Step 0. Retailer i selects a safety factor s^r_{ij} initially.
Step 1. At inspection time t, set reorder point according to Eq. 6, and retailer i places an order if its forecasted inventory position reaches the reorder point.
Step 2. After retailer i receives the ordered quantity, Q_i, update service level according to Eq. 7. Select the next safety factor according to the probabilistic rule in Eq. 5. Set $s^r_{ij} =$ the selected safety factor. Go to Step 1.

3.3.2 The supplier

In the decentralized model, the supplier's inventory-control policy is almost the same as the JIT policy in the centralized model. However, the objective functions of the supplier in both models are different. While the supplier's objective in the centralized model is to meet target customer service levels set by the retailers, in the decentralized model, it is to satisfy the retailers' orders without delay. In other words, the supplier's role in the decentralized model is to make the retailers' actual lead time as close as possible to their transportation lead times, for which safety lead times are needed. The retailers control customer service levels with safety stocks.

For retailer i, the supplier forecasts the time point t' at which the inventory position of retailer i is anticipated to drop down below its safety stock at first. If the time interval between the current inspection time t and the forecasted time t' is approximately close to the total lead time (supplier's lead time + retailer's transportation lead time + safety lead time), then the supplier places an order for retailer i. This JIT delivery rule can be represented as

At inspection time t, if
$\{($*the time* t' *that the time series model predicts the inventory position of retailer i reaches its safety stock*$) - t\}$
$\leq L_0 + L_i + st^s_{ij}$,
then the supplier issues an order of Q_i *to the outside source.* (8)

Because the objective of the supplier is to deliver the ordered quantity without delay to retailer i whenever he places an order, the service level of the supplier can be defined as

$$\overline{SL}_{new}(s^s_{ij}) = \overline{SL}_{old}(s^s_{ij}) + StepSize[\max(0, \text{arrival time of ordered quantity from the outside source to the supplier} - \text{retailer's order placement time}) - \overline{SL}_{old}(s^s_{ij})] \quad (9)$$

Also, the safety factor selection rule is modified as

$$P\{new\ safety\ factor = s^s_{ij}\} = \frac{e^{-SL_{new}(s^s_{ij})^2}}{\sum_k e^{-SL_{new}(s^s_{ik})^2}} \quad (10)$$

The detailed operational steps of the supplier are as follows:

Supplier

Step 0. The supplier selects safety factor s^s_{ij} initially for each retailer i ($i = 1, 2, \ldots, N$).
Step 1. At inspection time t, if the retailer i's inventory replenishment condition 8 is satisfied, then the supplier issues an order of Q_i to the outside source.
Step 2. If the ordered quantity Q_i is arrives from the outside source and retailer i does not place an order yet, keep Q_i in his warehouse. Otherwise, deliver Q_i to retailer i.
Step 3. After retailer i receives the ordered quantity, Q_i, the retailer updates the service level according to Eq. 9. Then the supplier selects the next safety factor according to the probabilistic rule in Eq. 10. Set $s^s_{ij} =$ the next safety factor. Go to Step 1.

4 Simulation-based experiment

4.1 Simulation environment

The simulated supply chain consists of one supplier and four retailers. Different customer-demand processes are assumed for each retailer. The time interval for inspecting the retailers' inventory position and customer demands is set at 1 day. The length of a simulation run is 5,000 days, and the simulation result obtained from the first 500 days is excluded in measuring the service level in order to minimize the transient effect of the simulation run. Given a specific demand process for each retailer, 20 simulation runs were performed and their averages measured.

Two types of demand process patterns are considered: stationary demands and nonstationary demands. In the case of stationary demands, a normal distribution with mean 50 demands/day and standard deviation 10 demands/day is applied. On the other hand, in the case of nonstationary demands, the mean of the normal distribution is designed to change at every random interval T according to the rule of $mean_j = mean_{j-1} + slope$. In this rule, $slope$ and T are randomly created by uniform distributions $U(-sm, sm)$ and $U(tu/2,\ tu)$, respectively. sm and tu characterize the nonstationarity of demand processes. In this experiment, we set the two parameters as

$$\text{Low mean variation (LMV)} : sm = 1.0 \quad \text{and} \quad tu = 30$$
$$\text{Medium mean variation (MMV)} : sm = 2.0 \quad \text{and} \quad tu = 15$$
$$\text{High mean variation (HMV)} : sm = 4.0 \quad \text{and} \quad tu = 8\ .$$

The coefficient of variation (CV) is also taken into account in the performance evaluation, where the coefficient of variation is defined as the standard deviation divided by the mean. Because mean demand changes with time, performance evaluation with a fixed CV implies the change of the standard deviation of demand. Three coefficients of variation are chosen: $\text{CV} = 0.1$, $\text{CV} = 0.15$, and $\text{CV} = 0.2$.

The supplier's lead time and retailer's transportation lead times are set to 3 and 2 days, respectively. The target service level is set to 95%. Finally, in the two models, the safety factor of the supplier is allowed to have integer values in the range of $[-3, 3]$. In the decentralized model, the safety factor range of the retailers is set to [3, 6] with increment 0.2 (See Bernard [17] for the justification of the safety factor ranges.)

4.2 Results and analysis

Figure 4 shows simulation results for the case of a stationary demand process. In this case, one retailer is assumed. The plotted data in the figure are the service levels of the retailer collected at

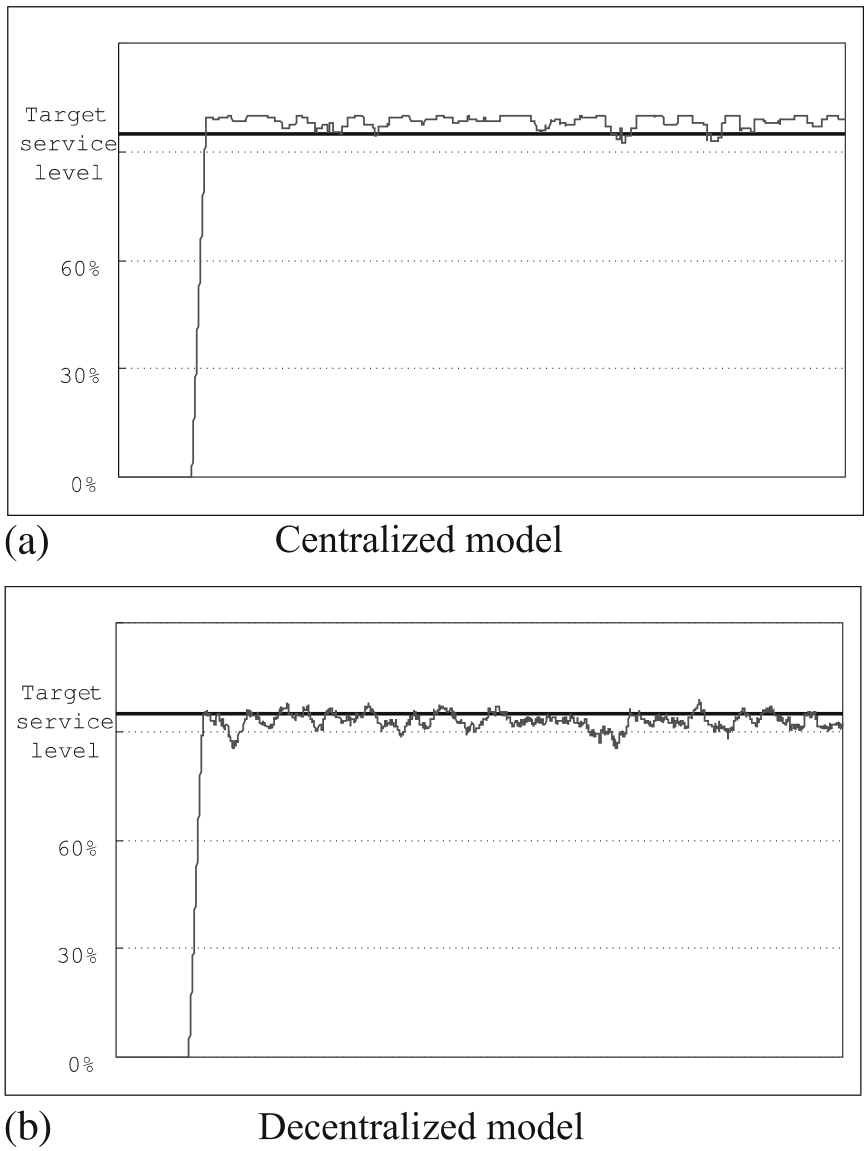

Fig. 4. Sample paths of retailer's service level (stationary demand process and supply chain consisting of one supplier and one retailer)

inventory replenishment times. As shown in the figure, the two inventory-control models rapidly approach the 95% target service level and do not significantly deviate from the target service level.

From the aspect of service-level stabilization, however, we can know that the centralized model shows a more stabilized service level than the decentralized model. This result also experimentally supports the argument that, in most cases, decision-making by a central controller produces better results than decentralized decision making. If the central controller is able to access information on participants in a system, then it is able to coordinate them toward achieving their global objective. On the other hand, pure decentralized controllers are only concerned with their local objectives, which are frequently inconsistent with global objectives.

However, if local controllers in a decentralized system work together according to a global coordination mechanism, it is not impossible for them to generate globally good solutions. This is shown in the decentralized case of Fig. 4. Although, in the decentralized model proposed in this paper, the supplier and retailers independently control their safety factors, the control objective of the supplier is strongly related to that of the retailers. That is, the supplier does not control its safety factor (or safety lead time) for reducing its own inventory-related objective, but for satisfying the orders issued by the retailers. This is an implicit coordination mechanism between the supplier and retailers. Suppose that the retailers find reorder points that can meet the target service level. Then the global objective of satisfying the target service level can be easily achieved provided the supply perfectly replenishes retailer's inventory without delay.

In the case of a nonstationary demand process, the service levels of the four retailers are averaged for each combination of nonstationarity factors (mean demand variation, coefficient of variation) and plotted on a graph. Figure 5 shows the graph. In the graph, the performance of the decentralized model is also evaluated with different β values. As expected, the centralized model shows more stabilized results than the decentralized model in all combinations of the nonstationarity factors.

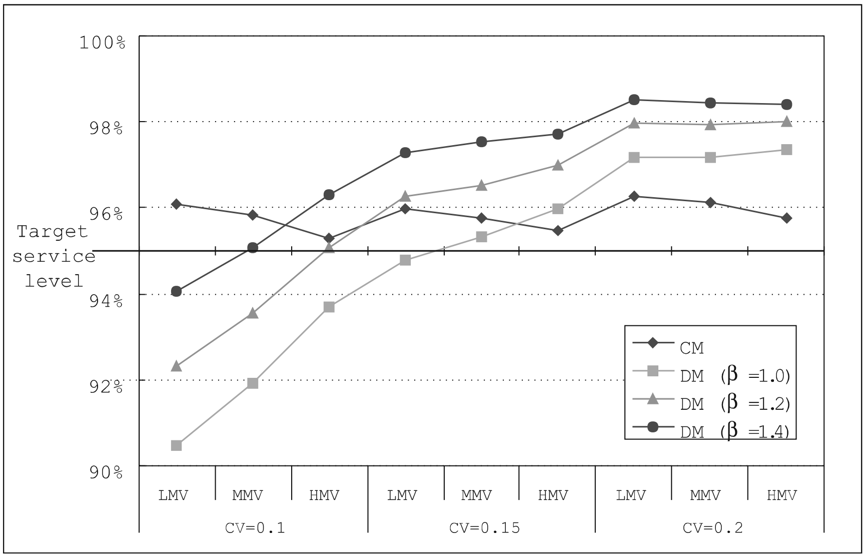

Fig. 5. Averages of four retailers' service levels generated by centralized model (CM) and decentralized model (DM) with three different β values (supply chain consisting of one supplier and four retailers, nonstationary demand process with different combinations of mean demand variation and coefficient of variation)

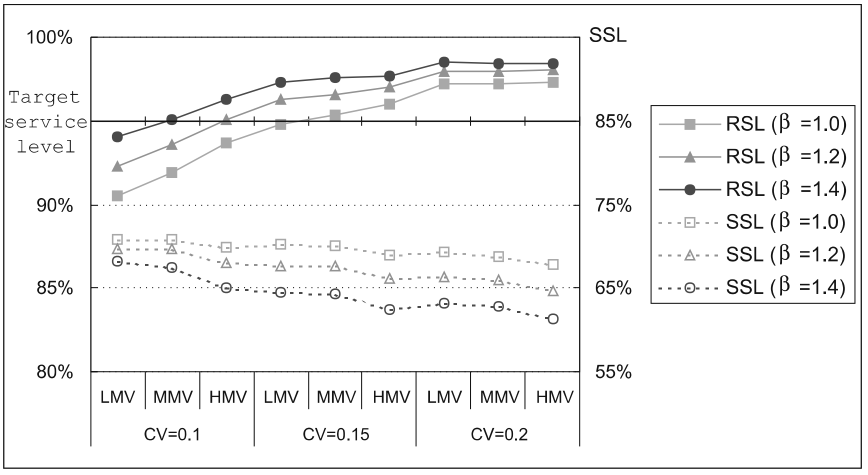

Fig. 6. Averages of four retailers' service levels (RSL) and supplier's service levels (SSL) in decentralized model with three different β values: supply chain consisting of one supplier and four retailers, nonstationary demand process with different combinations of mean demand variation and coefficient of variation

However, the performance of the decentralized model is not so disappointing—its average service levels do not fall below 90%.

In Fig. 5, as demand nonstationarity increases (high mean demand variation, high CV), the average service levels also show increasing trends. This result can be explained based on the fact that high-demand nonstationarity implies a large standard deviation of demand, which enforces the safety stocks of the retailers to increase, resulting in high reorder points. Therefore, the retailers place orders more frequently than in stationary demand situations in order not to lose sales opportunities.

If the supplier has sufficient inventory in advance to meet all the retailers' orders, then the service levels of the retailers may reach 100%. However, because the supplier has to place orders to the outside source, he cannot perfectly follow retailers' efforts. This argument can be verified in Fig. 6, where the supplier's service level decreases as demand nonstationarity increases. Finally, in Fig. 6, we can observe the role of β. Note that safety stock is affected by the formula to approximate the retailer's actual lead time. Because β is an exponential factor in the approximation formula, large β values result in high safety factors.

5 Conclusions

In this paper, we deal with the inventory-control problem of a two-echelon supply-chain system consisting of one supplier and multiple retailers. To cope with the nonstationary demand situation, we propose two adaptive inventory-control models, with the assumption that the supplier is able to access on-line information about customer demand, as well as the inventory position of each retailer. By applying a reinforcement-learning technique, the control parameters of the two inventory-control models are designed to adaptively change as customer-demand patterns change. A simulation-based experiment was conducted to compare the performance of the two inventory-control models. Finally, two important future research topics related to our approach are mentioned as follows: (1) extension of our approach to multiechelon supply chains and (2) incorporation of adaptive control of β in our approach.

Acknowledgement This work was supported by Korea Research Foundation grant KRF-2001-042-E00018.

References

1. Waller M, Johnson M, Davis T (1999) Vendor-managed inventory in the retail supply chain. J Business Logist 20:183–203
2. Sutton R, Barto A (1998) Reinforcement learning. MIT Press
3. Scarf H (1959) Some remarks on Baye's solution to the inventory problem. Naval Res Logist Q 7:591–596
4. Azoury K (1985) Bayes solution to dynamic inventory models under unknown demand distribution. Manage Sci 31:1150–1160
5. Miller B (1986) Scarf's state reduction method, flexibility, and a dependent demand inventory model. Oper Res 34:83–90
6. Lovejoy W (1990) Myopic policies for some inventory models with uncertain demand distributions. Manage Sci 36:724–738
7. Packer A (1967) Simulation and adaptive forecasting as applied to inventory control. Oper Res 15:660–679
8. Moinzadeh K (2002) A multi-echelon inventory system with information exchange. Manage Sci 48:414–426
9. Axsater S (2001) A framework for decentralized multi-echelon inventory control. IIIE Trans 33:91–97
10. Andersson J, Axsater S, Marklund J (1998) Decentralized multi-echelon inventory control. Prod Oper Manage 7:370–386
11. Lee H, Whang S (1999) Decentralized multi-echelon supply chains: incentives and information. Manage Sci 45:633–640
12. Cachon G, Zipkin P (1999) Competitive and cooperative inventory policies in a two stage supply chain. Manage Sci 45:936–953
13. Zhao X, Xie J, Lau R (2001) Improving the supply chain performance: use of forecasting models versus early order commitments. Int J Prod Res 39:3923–3939
14. Zhao X, Xie J (2002) Forecasting errors and the value of information sharing in a supply chain. Int J Prod Res 40:311–335
15. Pontrandolfo P, Gosavi A, Okogbaa O, Das T (2002) Global supply chain management: a reinforcement learning approach. Int J Prod Res 40:1299–1317
16. Brown R (1962) Smoothing, forecasting, and prediction of discrete time series. Prentice-Hall
17. Bernard P (1999) Integrated inventory management. Wiley