Int. J. Production Economics 151 (2014) 48–55

Contents lists available at ScienceDirect

Int. J. Production Economics

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

# Optimal inventory policy for a Markovian two-echelon system with returns and lateral transshipment ☆

Allen H. Tai [a,*], Wai-Ki Ching [b]

[a] Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong
[b] Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong

ARTICLE INFO

Article history:
Received 18 September 2012
Accepted 2 January 2014
Available online 18 January 2014

Keywords:
Lateral transshipment
Local warehouse
Markovian queueing networks
Returns
Two-echelon system

ABSTRACT

In this paper, we consider a Markovian model for a two-echelon inventory/return system. The system consists of a supply plant with infinite capacity and a central warehouse for inventory and returns. There are also a number of local warehouses which are also capable of re-manufacturing of returns to products. To obtain a high service level of handling inventory and returns, lateral transshipment of demands is allowed among the local warehouses. The objective is to minimize the total expected operating cost by choosing the maximum inventory level at the local warehouses. We present some theoretical results for the proposed model. Numerical examples are also given to demonstrate the performance of the system.


## 1. Introduction

The efficiency of product/service delivery is one of the major concerns in many industries including re-manufacturing industry. Since customers are usually scattered over a large regional area, a network of locations (local warehouses) for inventory of products and handling of returns is necessary to maintain a high service level. In our study, Lateral Transshipment (LT) is allowed among the local warehouses to enhance the service level. LTs are also very practical in many organizations having multiple locations linked by computers. Substantial savings can be realized by the sharing of inventory in the local warehouses (Robinson, 1990). A number of research publications have been appeared in this area. Kukreja et al. (2001) developed a single-echelon and multi-location inventory model for slow moving and consumable products. Aggarwal and Moinzadeh (1994) studied the emergency replenishments for a two-echelon model with deterministic lead time. Ching (1997) considered a multi-location inventory system where the process of LT is modeled by Markov-modulated Poisson Process (MMPP). Both numerical algorithm and analytic approximation were developed to solve the steady state probability distribution of the system (Ching, 1997; Ching et al., 1997, 2013). Alfredsson and Verrijdt (1999) considered a two-echelon inventory system for service parts with emergency supply options in terms of LT and direct delivery. Olsson (2009) proposed a single-echelon, two-location system with LT and derived the optimal policy by using stochastic dynamic programming. Paterson et al. (2012) proposed a quasi-myopic approach to the development of a strongly performing enhanced reactive transshipment policy. Service levels are improved while the aggregate costs incurred in managing the system can be significantly reduced. Tiacci and Saetta (2011a) proposed a 'preventive' transshipment policy, which means that transshipment occurs if the inventory level of a warehouse is lower than a predetermined threshold. Tiacci and Saetta (2011b) then extended the policy by considering stock redistribution among warehouses. Minner and Silver (2005) compared two transshipment policies for two warehouses analytically: (i) never transship and (ii) always transship when shortage occurs at one location and stock available at another. Hochmuth and Köchel (2012) suggested using simulation optimization for the control problem of multi-location inventory systems with LT. Summerfield and Dror (2012) discussed several two-stage decentralized inventory problems using the same framework. Graphical taxonomy trees were depicted andstochastic programming models were also applied. Readers are referred to Paterson et al. (2011) for a recent review on LT.

Recently, studies on multi-echelon inventory systems with returns are getting more attention. Hoadley and Heyman (1977) considered a one-period multi-echelon model that allows LT between stocking points at the same multi-echelon level. Returns are handled at the first echelon and the optimal initial stock level at the first echelon was obtained. Lee (1987) considered a similar situation with continuous monitoring on inventory levels. Korugan and Gupta (1998) considered a two-echelon inventory system with return flows, where demand and return rates are mutually

☆A preliminary version of the paper has been presented and published in the Operations Research Proceedings 2005 (Tai and Ching, 2006).

* Corresponding author. Tel.: 852 2766 4694.
E-mail address: allen.tai@polyu.edu.hk (A.H. Tai).

0925-5273/$ - see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ijpe.2014.01.010

independent. Min et al. (2006) proposed a nonlinear mixed-integer programming model and a genetic algorithm that can solve the multi-echelon reverse logistics problem involving product returns. Mitra (2009) considered a two-echelon inventory system with returns. The system consists of one depot and one distributor only. Both deterministic and stochastic demand rates were considered. He and Zhao (2012) proposed a multi-echelon supply chain with both demand and supply uncertainty. Unsold products can be returned from retailers to manufacturer. Vercraene et al. (2014) studied the coordination of manufacturing, remanufacturing and returns acceptance in a hybrid manufacturing/remanufacturing system and proposed several heuristic control rules. Reviews on multi-echelon inventory research with returns can be found in Dekker et al. (2004) and Fleischmann and Minner (2003). However, most of the literature considered handling returns at the upper echelons but few considered handling returns at the lower echelons.

One possible approach to study multi-echelon system is using Markovian model. Gross et al. (1993) obtained the steady-state probabilities for large multi-echelon reparable item inventory systems without considering LT. Saetta et al. (2012) presented an analytical approach for Markovian performance analysis of a serial multi-echelon supply chain without considering returns. Some recent studies on applying Markovian model in inventory control problems are Liu et al. (in press) and Sebnem Ahiska et al. (2013). We remark that none of the studies considered a Markovian multi-echelon model for handling both returns and LT.

In this paper, we propose an inventory/returns model based on the frameworks and analyses in Alfredsson and Verrijdt (1999) and Lee (1987). The model of the system consists of a supply plant with infinite capacity, a central warehouse and a number of local warehouses with re-manufacturing capacity. Here we consider a queueing model for a two-echelon inventory system. Queueing model is a useful tool for many inventory models and manufacturing systems that can assist with long-run decision, see for instance Buzacott and Shanthikumar (1993), Ching et al. (1997, 2003, 2013) and Ching (2001). The rest of the paper is organized as follows. In Section 2, we propose a two-echelon inventory model. In Section 3, we consider a Markovian queueing model for the local warehouses. In Section 4, we consider an aggregated model for the central warehouse. We then give the expected operating cost function in Section 5 and a numerical example is given in Section 6. Finally, concluding remarks are given in Section 7 to conclude the paper.

## 2. The two-echelon system

In this section, we present a two-echelon system based on the frameworks discussed in Alfredsson and Verrijdt (1999) and Lee (1987). The system consists of a supply plant with infinite capacity, a central warehouse (with maximum capacity $C$) and $n$ identical local warehouses (each with maximum inventory level $S$), see Fig. 1. The arrival processes of both the demands and the returns at each local warehouse are assumed to follow independent Poisson

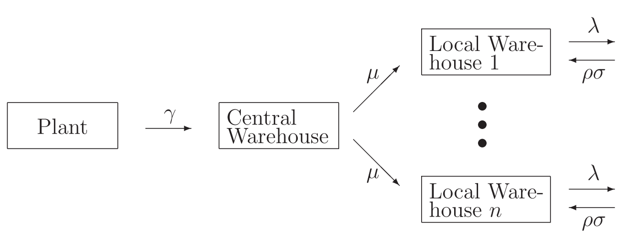


**Fig. 1.** The two-echelon model.

processes with mean rates $\lambda$ and $\sigma$ respectively. If the local warehouse is full, the returns will be disposed. The returns are then inspected and the inspection time is assumed to be negligible. If the returns are repairable, they will be repaired and ready to be delivered again. If the returns are not repairable, they will be disposed (Ching et al., 2003). The probability that a return is repairable is $\rho (0<\rho<1)$. Moreover, the demands are satisfied in the following manner.

(a1) The demand is first filled by the stock at the local warehouse and at the same time a replenishment order is issued to the central warehouse. The replenishment time from the central warehouse to a local warehouse is assumed to be exponentially distributed with mean $\mu^{-1}$. The demand is first-come-first-served and the replenishment is one-for-one. However if a repaired return arrives before the replenishment order is satisfied, the replenishment order will be cancelled.

(a2) If a local warehouse is out of stock, the demand is satisfied by a LT from another randomly chosen local warehouse and at the same time the local warehouse that sourcing the LT will issue a replenishment order to the central warehouse. Here in this situation, we assume that the LT time is negligible when compared with the time for a replenishment order from the central warehouse.

(a3) If unfortunately all the local warehouses are also out of stock, then the demand is satisfied by a direct delivery from the central warehouse and at the same time a replenishment order is issued to the plant from the central warehouse. The replenishment time from the plant to the central warehouse is assumed to be exponentially distributed with mean $\gamma^{-1}$.

(a4) If all the warehouses (including the central warehouse) are out of stock, the demand is satisfied by a direct delivery from the plant with infinite capacity.

We first analyze the inventory in each local warehouse as a simple Markovian queueing system. We then consider an aggregated model for the inventory levels in both the central warehouse and the local warehouses.

## 3. The local warehouses

In this section, we discuss the service level of each local warehouse. For simplicity of discussion, we assume that all the local warehouses are identical and we define the following probabilities:

- $p$ the common starving probability of the local warehouses (the probability that a warehouse has no stock on-hand)
- $\beta$ the common probability that a demand is met from stock on-hand at the local warehouse at which it occurs
- $\alpha$ the common probability that a demand is met by a LT
- $\theta$ the common probability that a demand cannot be met either from stock at the location at which it occurs or by a LT

The probability $\theta$ can be calculated by grouping all the local warehouses together. The aggregated demand arrival rate is $n\lambda$ and the aggregated maximum inventory level is $nS$. We take the aggregated arrival rate of repaired return as $n\rho\sigma$ by neglecting the possibility that some of the local warehouses may be full and do not accept returns. We note that as long as any location has stock, demand can be met and therefore we have (Buzacott and Shanthikumar, 1993)

$$\theta = \frac{(n\lambda)^{nS}}{\prod_{k=1}^{nS}(k\mu+n\rho\sigma)}\left[1+\sum_{j=0}^{nS-1}\frac{(n\lambda)^{nS-j}}{\prod_{l=1}^{nS-j}(l\mu+n\rho\sigma)}\right]^{-1}. \quad (1)$$