INVENTORIES, WATER STORAGE AND QUEUES

N.U. Prabhu
School of Operations Research and Industrial Engineering, Cornell University, USA

Keywords: Base stock inventory model, closed and open networks, input, interarrival times, loss and delay systems, Markovian networks, Markovian switching, Moran's dam model, its extension to continuous time, newsvendor problem, periodic and continuous inspection, queue discipline, queue length, (Q, r) model, service mechanism, service times, shortfall, Solberg model, (s, S) inventory model, switching probabilities, traffic equations, traffic intensity, workload.

Contents

1. Introduction
2. Inventory Models
3. Models for Water Storage
4. Queueing System GI /G /S
5. Queueing Networks
Glossary
Bibliography
Biographical Sketch

Summary

This article attempts to offer a perspective on the subject areas of inventories, water storage and queues, which are three important life support systems. For this purpose some selected models for these systems are surveyed, including classical models that continue to be important, and models that have been the subject of relatively new research. The theoretical as well as practical ideas underlying the models are emphasized, a sketch is given of their analysis and the main results are listed. Inventory models are described in Section 2, dam models in Section 3, the queueing system GI/G/S in Section 4 and queueing networks in Section 5.

1. Introduction

Inventories, water storage and queues are three of the important systems that greatly affect the daily lives of human beings. The study of these systems was initiated at the beginning of the twentieth century and continues to be an active area of research even today, with a vast class of new problems being investigated. Brief descriptions of these systems are as follows.

An inventory is an amount of material stored for the purpose of future sale or production. Traditionally inventories were viewed as a measure of wealth. However surplus stocks were also a principal cause of business failures, and in the 1920s increased emphasis was put on the liquidity of assets such as inventories. With the changing economic conditions of the1970s managers began to recognize the importance of balancing the advantages and disadvantages of carrying inventories. Research has shown that the success of the Japanese automobile industry in the 1980s is due to a production system that efficiently eliminates the need for any significant amount of inventory. Inventory control is increasingly important in today's economy.

The storage of water in a dam is also a type of inventory. Here the inflow of water depends on rainfall, underground seepage, etc., over which the operator of the dam has no control. However, the operator can (and does) try to control the release of water according to the demand, which is either for electric power (expressed in terms of volume of water required to produce it), irrigation or for water to be supplied to a city.

A queueing system is a facility that offers a certain service. Customers arrive at this system and demand this service. If they cannot be served immediately upon arrival they either leave the system (are lost), or else decide to wait (are delayed) and are served according to a prescribed policy. As applied to telephone systems queueing theory (under the title congestion theory) emerged in the early part of the twentieth century. It has continued to be a very active of research throughout, with newer areas of application being investigated in the last thirty years or so; these include computer performance, telecommunication and flexible manufacturing systems.

The quantity of interest in a queueing system is either the number of customers waiting for service or else the workload submitted to the system. This quantity may be compared to material held in stock for future use (measured in discrete or continuous units). This establishes an analogy between queues and inventories.

The above descriptions make it clear that storage systems may be a common term to describe inventories, water storage and queues. The study of storage systems is based on mathematical models and the objective is to study the behavior of stochastic processes that arise from the models and to obtain rules for operating the systems in an optimal manner, given the costs. This subject area is of interest to engineers, mathematicians and economists. The analysis of the models uses techniques of applied mathematics and operations research, and the results of the theory of stochastic processes. The results obtained range from abstract mathematical properties to results meant for practical use.

2. Inventory Models

The need to stock up items in order to meet future needs has been felt since the beginning of history. But it was not until the turn of the twentieth century that analytical techniques were used to solve inventory problems and a theory of inventory control began to develop. The objective of this theory is to determine when to order replenishment and how much to order. In order to do this the system operator has to decide how often the inventory position should be inspected (that is, periodic inspection versus continuous inspection). Thus inventory control is based on an operating policy.

The key components of an inventory model are (1) the pattern of demands for the item, (2) a policy regarding actions to be taken to deal with demands that occur when the system is out of stock (namely, whether to allow backlogs or let sales to be lost), (3) lead time, which is the interval of time between the ordering time and the time of delivery and (4) various costs incurred in operating the system such as procurement costs, inventory carrying (holding) costs, costs associated with deficit, costs of filling orders and costs of data gathering and control procedures (information processing).

In practice the demand pattern and the lead time may be subject to randomness, and the resulting model is then a stochastic model. When randomness is absent, the model is deterministic.

Some of the earliest inventory models are described in the following subsections. The conceptual framework of these models has remained important even in recent developments.

2.1. The (Q, r) Model and the EOQ Formula

There is a single location at which demands occur at a constant rate of λ units per year. It is assumed that the system is never out of stock when a demand occurs. The lead time is a constant τ. There is a fixed ordering cost A and each unit costs C independently of the quantity ordered. Also, let I denote the cost rate per each dollar invested in inventory. The information processing costs are independent of the order size and the reorder rule, and are not included in the total cost of operating the system.

Suppose that a quantity Q is ordered each time the system replenishes its stock. Then the intervals of time between successive replenishments are of length $T = Q/\lambda$. Since the lead time is a constant, the intervals of time between successive deliveries are also of length T, which is called the length of a cycle. During each cycle the system behavior is repeated. The number of cycles during a time interval (0, t] is given by $n = [t/T]$, where for any real number x, $[x]$ is the largest integer less than or equal to x.

Suppose the system starts at time $t = 0$ with a delivery, the inventory level just prior to this delivery being $s(s \geq 0)$. The inventory earning cost per cycle is given by

$$IC\int_0^T (s+Q-\lambda t)dt = ICT \left(s+\frac{1}{2}Q\right). \tag{1}$$

The number of deliveries made during a time interval (0, t] is n, whereas the number of orders placed during this interval is either n or $n + 1$ (this is denoted as m in the following calculations). Thus the total cost over (0, t] is

$$(A+CQ)m + IC\cdot nT\left(s+\frac{1}{2}Q\right)+\eta \tag{2}$$

where η is the inventory carrying cost over (0, $t - nT$], this being less than the quantity given by (1). The long term average annual cost is

$$\lim_{t\to\infty}\frac{1}{t}\left[(A+CQ)\,m+ICnT\left(s+\frac{1}{2}Q\right)+\eta\right]$$
$$=A\frac{\lambda}{Q}+C\lambda+IC\left(s+\frac{1}{2}Q\right). \tag{3}$$

Since $C\lambda$ does not depend on Q, the expression given by (3) is a minimum when $s = 0$ and

$$Q=\sqrt{\frac{2\lambda A}{IC}}\,. \tag{4}$$

Since the optimal inventory level is zero just prior to a delivery, the optimal ordering time is such that the corresponding inventory level (the reorder paint) is r_h, where

$$r_h+\left[\frac{\tau}{T}\right]\;Q-\lambda\tau=0\,. \tag{5}$$

The expression (4) is called the economic order quantity (EOQ) and is one of the earliest and best known results of inventory theory. It was first obtained by Ford Harris of the Westinghouse Corporation in 1915, and was also derived by R.H. Wilson, for which reason it is sometimes called the Wilson Lot Size formula.

The EOQ minimizes the total cost for given A and I. In practice it may be more appropriate to find ways of reducing the Fixed cost A or the inventory carrying cost C. This is a part of the Japanese Just-in-Time (JIT) system, in which A is reduced as much as possible. This will also reduce the EOQ and the average inventory level.

The overall approach of which JIT is a part is known as stockless production, a concept first introduced by the Toyota Motor Company in the 1980s. In this approach to the total manufacturing system inventories are, reduced at every stage of production with a view to increased productivity, improved quality and reduced lead times. The associated information processing is known as Kanban (the Japanese word for a card).

2.2. The Newsvendor Problem

Each day a newsvendor has to decide how many copies of a particular paper to buy for the day's sale. The paper costs c per copy and the selling price is $s(> c)$ per copy. The demand for this newspaper is a random variable D with a distribution function ($D.F.$)F. Any copy not sold at the end of the day is a total loss.

If the vendor buys h copies of the newspaper each day, the gain is $s\min(h, D)-ch$. The expected gain is

$$G(h)=s\int_0^h[1-F(x)]\,dx-ch\,. \tag{6}$$

This is a minimum for $h = \hat{h}$, where

$$F(\hat{h}) = 1 - c / s. \tag{7}$$

Thus $\hat{h}$ is the optimal number of copies the vendor should buy each day.

The newsvendor problem leads to a single period inventory model dealing with a perishable commodity. The obvious extension of this model is to items that can be carried over from period to period. Two such models are described below.

2.3. The (*s*, *S*) Inventory Model

This model is described as follows. Two real numbers s, S are given, where $0 \leq s < S < \infty$. The inventory is inspected at times $n = 0, 1, 2, \ldots$ Let ξ_{n+1} denote the demand over the period $(n, n+1]$; $\xi_1, \xi_2, \ldots$ are independent and identically distributed (IID) random variables. The amount sold always equals the demand (that is, backlogs are allowed). Whenever the inventory level falls below s, an order is placed to bring up the level to S, but otherwise no ordering is done. Thus the amount ordered at time n is

$$\eta_{n+1} = 0, \quad \text{if} \quad s \leq Z_n \leq S, \quad \text{and} \quad = S - Z_n \quad \text{if} \quad Z_n < s \tag{8}$$

where Z_n is the inventory level at time n. Deliveries are made instantaneously (thus the lead time is zero). Therefore Z_n satisfies the recurrence, relation

$$\begin{aligned} Z_{n+1} &= Z_n - \xi_{n+1} \quad \text{if} \quad s \leq Z_n \leq S \\ &= S - \xi_{n+1} \quad \text{if} \quad Z_n < s \ (n \geq 0). \end{aligned} \tag{9}$$

Suppose that an order is placed initially, so that $Z_0 < s$. The D.F. F_n of Z_n, can be expressed in terms of its distribution in between two consecutive replenishments. Thus, let $\Delta = S - s$ and

$$\begin{aligned} H_n(x) &= P\{S - \xi_1 \geq s, S - \xi_1 - \xi_2 \geq s, \ldots, \\ &\quad S - \xi_1 - \xi_2 - \ldots - \xi_{n-1} \geq s, S - \xi_1 - \xi_2 - \ldots - \xi_n < x\} \\ &= P\{S_{n-1} \leq \Delta, S_n > S - x\} (n \geq 1), x \leq S), \end{aligned} \tag{10}$$

where $S_0 = 0$, $S_n = \xi_1 + \xi_2 + \ldots + \xi_n$ = total demand up to time $n (n \geq 1)$. In particular, let $H_n(s) = f_n (n \geq 1)$. Then the intervals of time between successive replenishments are IID with the distribution $\{f_n\}$. From renewal theory it is known that this distribution has mean λ, where

$$\lambda = \sum_0^\infty P\{S_n \leq \Delta\} < \infty \ . \tag{11}$$

A simple probability argument shows that

$$F_n(x) = H_n(x) + \sum_{m=1}^{n-1} f_m F_{n-m}(x) \ \ (n \geq 1) \tag{12}$$

which is a renewal equation for $F_n(x)$. Its unique bounded solution is given by

$$F_n(x) = \sum_{m=1}^{n} u_{n-m} H_m(x) \ \ (n \geq 1) \tag{13}$$

where $u_0 = 1$ and for $n \geq 1$, u_n is the probability that an order is placed at time n. Thus

$$u_n = P\{Z_n < s\} \ (n \geq 0)\,. \tag{14}$$

The probabilities u_n can be obtained from $\{f_n\}$, since (12) yields the relation

$$u_n = \sum_{m=1}^{n} f_m u_{n-m} \ . \tag{15}$$

Therefore the solution (13) is completely determined.

The steady state distribution of Z_n can also be found, using renewal theory. Thus

$$\lim_{n\to\infty} \ u_n = \lambda^{-1} \tag{16}$$

and

$$F(x) = \lim_{n\to\infty} \ F_n(x) = \frac{1}{\lambda}\sum_{1}^{\infty} H_m(x)\ . \tag{17}$$

An optimization problem that arises in this model is to find the values of s, S that minimize the long run expected cost of operating the system. Apart from the costs A, C and I defined in the (Q,r) model there is now a cost p per unit for backordering an item. The total cost at time n is given by

$$-pZ_n 1_{\{Z_n<0\}} + [A + C(S - Z_n)]1_{\{Z_n<s\}} + hZ_n 1_{\{0\leq Z_n \leq S\}} \tag{18}$$

where $h = IC$ = holding cost and for any event E, $1_E = 1$ or 0 according as E occurs or does not occur (the indicator function of E). As an illustration, suppose that the demands ξ_n have density $\mu e^{-\mu x}$. Easy calculations show that

$$\begin{aligned} \sum_{1}^{\infty} H_n(x) &= e^{-\mu(s-x)} && \text{for } x < s \\ &= 1 + \mu(x - s) && \text{for } s \leq x \leq S \end{aligned} \tag{19}$$

and

$$\lambda = 1 + \mu\Delta. \tag{20}$$

The steady state D.F. of the inventory level Z_n is therefore given by

$$\begin{aligned} F(x) &= \tfrac{e^{-\mu(s-x)}}{1+\mu\Delta} \quad \text{for } x < s \\ &= \tfrac{1+\mu(x-s)}{1+\mu\Delta} \qquad \text{for } s \le x \le S. \end{aligned} \tag{21}$$

Using this distribution in (18) the expected long run cost is found to be

$$\frac{1}{\mu(1+\mu\Delta)}\left\{(h+p)e^{-\mu s} + \mu A + h(\mu s - 1) + \frac{1}{2}\mu^2 h\Delta(\Delta + 2s)\right\} + \frac{C}{\mu} \tag{22}$$

expressed as a function of s and Δ. This is a minimum when

$$\Delta = \sqrt{\frac{2A}{\mu h}} \text{ and } e^{-\lambda s} = \frac{h + \sqrt{2A\mu h}}{h+p} \tag{23}$$

if $\sqrt{2A\mu h} < p$ and

$$S = \sqrt{\frac{2A}{\mu h}} \quad \text{and } s = 0 \tag{24}$$

otherwise. It should be noted that Δ is the replenishment when an order is placed and the optimal value of Δ is given by the EOQ formula since μ^{-1} is the rate of demand.

2.4. A Periodic Review Base Stock Inventory Model

A production facility has a finite capacity of c units per period for producing a certain item. Inventory inspection is carried out periodically at times n = 0,1,2,…. There is a target inventory level R (called the base stock inventory level) at the end of each period. The difference between R and the actual inventory level at the end of this period is called the shortfall. The problem is to find the value of R that provides a balance between the end of period holding and backorder costs.

The demands ξ_1, ξ_2,… during the successive periods are IID random variables. Let Z_n denote the inventory level at time $n(\geq 0)$. The operating policy is a follows. At the beginning of the period $(n, n+1]$, the shortfall $V_n = R - Z_n$ and the demand ξ_{n+1} are observed. Then an amount $V_n + \xi_{n+1}$ is produced during this period subject to the capacity constraint. At the end of this period all of the demand is met, allowing for backlogs. Thus we have

$$Z_{n+1} = Z_n + \min(c, V_n + \xi_{n+1}) - \xi_{n+1} \tag{25}$$

or equivalently in terms of the shortfall

$$V_{n+1} = (V_n + \xi_{n+1} - c)^+ \quad (n \geq 0) . \tag{26}$$

From the theory of random walks it is known that as $n \to \infty$, $V_n \to V$ in distribution if, and only if $E(\xi_n) < c$. The distribution of the random variable V is known, but has a complicated form. However, a more useful result is the following. Under certain conditions there exist numbers λ and C ($\lambda > 0$, $0 \leq C \leq 1$) such that

$$P\{V > v\} \sim Ce^{-\lambda v} \qquad (v \to \infty) \tag{27}$$

A consequence of the operating policy is that from time to time the production capacity c may not be fully utilized, which happens if the amount produced $V_n + \xi_{n+1}$ is less than c. If the unused capacity during $(n, n+1]$ is denoted as I_{n+1} then

$$I_{n+1} = (V_n + \xi_{n+1} - c)^- \quad (n \geq 0) . \tag{28}$$

Here and in (27) the notation used is the following. For any real number x,

$$x^+ = \max(0, x) \text{ and } x^- = -\min(0, x) . \tag{29}$$

This model addresses the interaction between inventories, production capacity and quality of service, and is motivated by the current situation faced by manufacturing companies, which are under intense pressure to simultaneously reduce inventories, utilize capacity and provide high customer service levels.

3. Models for Water Storage

Water storage problems were first treated by A. Hazen in 1914, when he studied the discharge of thirteen American rivers. This work was extended by C.E. Sudler in 1927. Both these authors used graphical methods. In a pioneering study carried out during 1938-1956, the British engineer H.E. Hurst investigated the problem of determining the storage required on a stream to give a certain minimum discharge using probability methods. His results were empirical in nature, and he applied them to compute the storage required for the Great Lakes of the Nile River Basin. In 1946 the French engineer P.B.D. Massé gave a formulation of the optimum storage problem for a hydroelectric system as follows. The demand is either for electric power (expressed in terms of the volume of water required to produce it) or for water to be supplied to a city. Supplementary sources exist (a thermal station or an arrangement for "borrowing" water from a nearby dam) in case the entire demand cannot be fully met, but these are available at a cost, and moreover, may be available only up to a certain limit. This problem was also studied by J.D.C. Little in 1955 with, reference to the Grand Coulee dam, and by T.C. Koopmans in 1958. In 1954 P.A.P. Moran formulated his probability theory of a dam, which is described in the next section. C.L. Jarvis applied Moran's model to the Ord River Project in Western Australia in 1963.

-

-

-

Bibliography

Hadley, G. and Whitin, T.M. (1963). *Analysis of Inventory Systems*, 452 pp., Englewood Cliffs, N.J, USA: Prentice Hall. [This classical text serves as a valuable source of historical references.]

Prabhu, N.U. (1965). *Queues and Inventories*: A Study of their Basic Stochastic Processes, 275 pp. New York: John Wiley. [This gives a detailed analysis of classical queueing and inventory models. Moran's dam model and continuous time storage processes.]

Prabhu, N.U. (1997). *Foundations of Queueing Theory*, 206 pp. Boston; MA, USA: Kluwer Academic Publishers. [This contains several problems dealing with applications and a bibliography that includes references to special issues of the journal Queueing Systems: Theory and Applications devoted to topics of current research.]

Prabhu, N.U. (1998). *Stochastic Storage Processes: Queues, Insurance Risk, Dams, and Data Communication*, Second Edition, 206 pp. New York: Springer Verlag. [This monograph contains new material on the queue length and virtual waiting time of single server queues; also included are Markov-modulated storage processes with application to data communication models.]

Roundy, Robin O. and Muckstadt, John A., (2000). Heuristic computation of periodic review base stock inventory policies, *Management Science*, **46**(1), 104-109. [This paper cites references to the earlier work on the base stock inventory model.]

Silver, Edward A., Pyke, David F. and Peterson, Rein (1998). *Inventory Management and Production Planning and Scheduling*, Third Edition, 754 pp. New York: John Wiley. [This is an authoritative treatment of the subject covering the most recent developments.]

Viswanadham, N. and Narahari Y. (1992). *Performance Modeling of Automated Manufacturing Systems*, 592 pp. Englewood Cliffs, NJ, USA: Prentice Hall. [The first text dealing with the mathematical modeling of manufacturing systems using queueing theory]

Biographical Sketch

N.U. Prabhu is Professor Emeritus at Cornell University. He has published many papers dealing with the subject of this article, and related topics such as regenerative phenomena, fluctuation theory and Wiener-Hopf factorization. He is the founding editor of the journal *Queueing Systems: Theory and Applications*, and co-founding editor of the journal *Stochastic Processes and Their Applications*. Previous he was associated professor of statistics at Michigan State University (1964-65); reader in mathematical statistics at the University of Western Australia (1961-64); reader and head of the Department of Statistics at Karnataka University (1952-61); and lecturer in mathematics and statistics at various universities in India (1946-48). He has been a visiting professor at Uppsala University, Sweden, the Technion: Israel Institute of Technology, The University of Melbourne and the Mathematics Research Centre at the University of Wisconsin. Prof. Prabhu is a Fellow of the Institute of Mathematical Statistics and a member of the American Mathematical Society. He is studied in India and received his M.Sc. degree from Manchester University in England in 1957.