*Appl. Math. Lett.* Vol. 1, No. 3, pp. 235–239, 1988
Printed in Great Britain

0893-9659/88 \$3.00 + 0.00
Pergamon Press plc

# A STATE DEPENDENT M/M/1 QUEUE: A PRODUCTION-INVENTORY MODEL

C. KNESSL[1], B. J. MATKOWSKY[2], Z. SCHUSS[3], C. TIER[1]

[1]Dept. of Mathematics, Statistics and Computer Science,
University of Illinois at Chicago
[2]Dept. of Engineering Sciences and Applied Mathematics,
Northwestern University
[3]Dept. of Mathematical Sciences, Tel Aviv University

## 1. INTRODUCTION

We consider an extension of a model of production and inventory control given in [1,2]. The model incorporates a continuous-level production rate with a base-stock level inventory policy subject to fluctuating demand. The inventory level at time $t$ is denoted by $Y(t)$ and is the time (or effort) required to produce the items currently in stock given a production rate of unity. Negative inventory levels $Y(t) < 0$, reflect orders that cannot be satisfied by current inventory and are backlogged. We assume that the inventory is produced at a rate which depends on the current inventory level and may differ from unity. Thus the production rate is controlled by the producer, which introduces a feedback mechanism for inventory control. We define the base-stock level $M$ as the level of inventory at which production stops. Production only restarts when $Y(t) < M$. Orders or demands arrive randomly with an inter-arrival time that is exponentially distributed with a parameter that depends on the current inventory level. The size of the orders is exponentially distributed. The dependence of the arrival rate on the current inventory level can have several interpretations. For example, a company might reject orders or discourage arrivals when the backlog of orders becomes sufficiently large or alternatively customers may not place orders if there is a backlog. As discussed in [1], this model is appropriate if production is highly automated and easily controlled.

The deficit below the base-stock level is defined by $U(t) = M - Y(t)$ and is the quantity of interest here [1,2]. In queueing theory, $U(t)$ is the unfinished work (virtual waiting time or buffer content) in the system at time $t$. In terms of $U(t)$ the inventory control model described above is an $M/M/1$ queue with state-dependent arrival and service rate. This state-dependence is the new feature introduced into the model described in [1,2]. We note that the methods used to analyze state-independent systems, e.g. Laplace transforms and generating functions, cannot in general be extended to state-dependent systems. Such systems also arise in other applications such as storage and dam theory and delays in computer and communications systems [3-6].

We now formulate our model in terms of the deficit level $U(t)$ including the state-dependence of the arrival and service rates on $U(t)$. The state-dependent arrival rate of orders is defined by

$$Pr\{\text{order arrives in } (t, t+\Delta t) | U(t) = w\} = \lambda(w)\Delta t + o(\Delta t). \tag{1.1}$$

The instantaneous production rate also depends on the current deficit level. Thus, if $U(t) = w$, the rate at which the deficit level is reduced at time $t$ is defined by

$$\text{production rate} = \begin{cases} r(w), & w > 0 \\ 0, & w = 0 \end{cases} \tag{1.2}$$

Typeset by $\mathcal{A}_M\mathcal{S}$-TEX

where we assume that $r(w) > 0$ if $w > 0$. Thus, the deficit level is decreased (or the inventory level is increased) at a rate of $r(w)$ units per unit time when $U(t) = w$ (or $Y(t) = M - w$). When $U(t) = 0$, the inventory level is $Y(t) = M$ and production stops. The order size is exponentially distributed with mean $1/\mu$. We define the instantaneous inventory turnover intensity (traffic intensity) by

$$\alpha(w) = \frac{\lambda(w)/\mu}{r(w)} = \frac{\text{expected total demand due to orders}}{\text{production rate}}. \tag{1.3}$$

We consider only systems with $\alpha(w) < 1$ if $w > w_0$ for some $w_0$. In terms of the inventory level, this means that the production capacity is not overtaken by the demand due to orders.

Our objective is to compute several long run operating characteristics (performance measures) for the above production facility. Specifically, we compute the stationary density of $U(t)$ including the facility utilization and the mean time for $U(t)$ to reach or exceed a given value. In addition, we analyze the production busy period of the facility. We compute the mean length of a production period and the mean time for a production busy period to end, given the current value of $U(t)$ (i.e. the residual production period). We also compute the mean residual total production until the end of a production period given the current value of $U(t)$ and the mean total production during a production period, i.e. the integral of $U(t)$ during a busy period. Finally, we compute the distribution of the maximum of $U(t)$ during a production period, i.e. the distribution of the minimum inventory level.

## 2. Performance Measures

We now formulate and solve problems for the performance measures described in Section 1 for the case when $r(w)$ and $\lambda(w)$ are smooth functions of $w$.

i) *Stationary Density.* We define the density function of $U(t)$ by

$$p(w) = \lim_{t\to\infty} \frac{\partial}{\partial w} Pr\{U(t) \le w\}, \quad A = \lim_{t\to\infty} Pr\{U(t) = 0\}. \tag{2.1}$$

We expect the stationary limits to exist if the turnover intensity $\alpha(w) \le 1 - \delta < 1$ for some $\delta > 0$ for all $w$ sufficiently large. Then $p(w)$ and $A$ satisfy the forward equation [7]

$$(r(w)p(w))' - \lambda(w)p(w) + \int_0^w \lambda(w-z)p(w-z)\mu e^{-\mu z} dz = -A\lambda(0)\mu e^{-\mu w} \tag{2.2}$$

and the boundary and normalization conditions

$$\lambda(0)A = r(0+)p(0+), \quad \int_0^\infty p(w)dw + A = 1 \tag{2.3}$$

where $r(0+) = \lim_{w\downarrow 0} r(w) > 0$. We note that the integro-differential equation (2.2) is a generalization of the Takacs equation in queueing theory [2,7]. By multiplying (2.2) by $e^{\mu w}$ and differentiating we obtain

$$(r(w)p(w))'' + [(\mu r(w) - \lambda(w))p(w)]' = 0 \tag{2.4}$$

with the boundary condition at $w = 0$

$$\lim_{w\to 0}[(r(w)p(w))' + (\mu r(w) - \lambda(w))p(w)] = 0. \tag{2.5}$$

The solution of (2.4)-(2.5) with the normalization condition in (2.3) is

$$p(w) = \frac{\lambda(0)A}{r(w)} e^{-\mu w} e^{\mu \int_0^w \alpha(s)ds} \tag{2.6}$$

where

$$A = \left[1 + \lambda(0) \int_0^\infty \frac{e^{-\mu z} e^{\mu \int_0^z \alpha(s)ds}}{r(z)} dz \right]^{-1}. \tag{2.7}$$

We observe that $p(w)$ has local maxima at points $w_0$ where $\alpha(w_0) = 1$ and $\alpha'(w_0) < 0$. Thus, in contrast to the state-independent situation, the density of $p(w)$ may have, e.g., a bimodal character.

ii) *Mean Time to Reach the Value K.* We compute the mean time for the deficit level to reach or exceed $K$, given $w$ units of deficit level. This is the mean time for the inventory level $Y(t)$ to reach $M - K$ given $Y(0) = M - w$. First, we define the stopping times

$$\tilde{n} = \inf[t : U(t) \geq K]. \tag{2.8}$$

The mean time, $n(w)$, to reach or exceed $K$ is

$$n(w) = E[\tilde{n} | U(0) = w] \tag{2.9}$$

where $E$ denotes expectation. The mean time $n(w)$ satisfies [7]

$$-r(w)n'(w) - \lambda(w)n(w) + \lambda(w) \int_0^{K-w} n(w+z)\mu e^{-\mu z} dz = -1, \; 0 \leq w < K, \tag{2.10}$$

and $n(w) = 0$, $w \geq K$. As we show, $\lim_{w \uparrow K} n(w) = n(K^-) > 0$ so that $n(w)$ has a discontinuity at $w = K$. Equation (2.10) can be transformed as above into

$$Ln(w) \equiv n''(w) + \left[\frac{(r(w)/\lambda(w))'}{r(w)/\lambda(w)} + \frac{\lambda(w)}{r(w)} - \mu\right] n'(w) = -\frac{\mu}{r(w)} - \frac{\lambda'(w)}{\lambda(w)r(w)}. \tag{2.11}$$

We obtain a boundary condition at $w = 0$ by evaluating (2.10) at $w = 0$, noting that $r(0) = 0$, and subtracting the result from the limit of (2.10) as $w \downarrow 0$, thus yielding $n'(0) = 0$. By evaluating (2.10) as $w \uparrow K$, we obtain $r(K)n'(K^-) + \lambda(K)n(K^-) = 1$. The solution of (2.11) that satisfies these two boundary conditions is

$$n(w) = \frac{e^{\mu K} e^{-\mu \int_0^K \alpha(s)ds}}{\lambda(0)} + \int_0^K \frac{e^{\mu(K-t)} e^{-\mu \int_t^K \alpha(s)ds}}{r(t)} dt \tag{2.12}$$

$$- \int_w^K \frac{\lambda(z)}{r(z)} e^{\mu z} e^{-\mu \int_0^z \alpha(s)ds} \int_0^z e^{\mu \int_0^t \alpha(s)ds} \frac{d}{dt}\left[\frac{e^{-\mu t}}{\lambda(t)}\right] dt\, dz.$$

From (2.12), we observe that $n(K^-) > 0$. Thus, the mean first passage time $n(w)$ is discontinuous at $w = K$. This is a consequence of the fact that the deficit level $U(t)$ can jump across the boundary at $w = K$ without actually hitting it. From (2.12), one can easily show that if $\alpha(w) < 1$ for $w \geq K$, then $n(w)$ is exponentially large in $K$.

iii) *Length of a Production Period.* We now compute the mean residual length and mean length of a production period. We define the residual length of a production period, or the amount of time for the deficit level to reach zero, given $w$ units of deficit level, by

$$\tilde{\tau} = \inf[t : U(t) = 0] \tag{2.13}$$

and the mean residual length of a production period by

$$\tau(w) = E[\tilde{\tau}|U(0) = w]. \tag{2.14}$$

The mean length of a production period $T$ is then given by

$$T = \int_0^\infty \tau(w)\mu e^{-\mu w} dw. \tag{2.15}$$

We first compute $\tau(w)$ which satisfies [7]

$$\hat{L}\tau(w) = -1, \quad \tau(0) = 0 \tag{2.16}$$

where $\hat{L}$ is the integro-differential operator on the left side of (2.10) with the upper limit on the integral replaced by $\infty$. The solution of (2.16) satisfying the condition of no exponential growth as $w \to \infty$ is

$$\tau(w) = \int_0^w \frac{\lambda(z)}{r(z)} \int_z^\infty \left[\frac{\mu}{\lambda(t)} + \frac{\lambda'(t)}{\lambda^2(t)}\right] e^{-\mu(t-z)} e^{\mu \int_z^t \alpha(s)ds}\, dt\, dz\,. \tag{2.17}$$

Using (2.17), we find that the mean length of a production period is

$$T = \int_0^\infty \frac{e^{-\mu z}}{r(z)} e^{\mu \int_0^z \alpha(s)ds}\, dz\,. \tag{2.18}$$

iv) *Total Production.* The residual total production is defined by

$$\tilde{u} = \int_0^{\tilde{\tau}} U(t)dt \tag{2.19}$$

where $\tilde{\tau}$ is the residual length of a production period defined by (2.13). The mean residual total production $u(w)$ is then

$$u(w) = E[\tilde{u}|U(0) = w]. \tag{2.20}$$

The mean total production during a production period, $P$, is defined by

$$P = \int_0^\infty u(w)\mu e^{-\mu w} dw\,. \tag{2.21}$$

The mean residual total production $u(w)$ satisfies [7]

$$\hat{L}u(w) = -w, \tag{2.22}$$

so that

$$u(w) = \int_0^w \frac{\lambda(z)}{r(z)} \int_z^\infty e^{-\mu(t-z)} e^{\mu \int_z^t \alpha(s)ds} \left[\frac{\mu t}{\lambda(t)} - \frac{1}{\lambda(t)} + \frac{\lambda'(t)t}{\lambda^2(t)}\right] dt\, dz\,. \tag{2.23}$$

The mean total production during a production period is constructed using (2.23) in (2.21) and repeated integration by parts. We obtain

$$P = \int_0^\infty \frac{ze^{-\mu z}}{r(z)} e^{\mu \int_0^z \alpha(s)ds}\, dz. \tag{2.24}$$

v) *Distribution of the Maximum.* We now compute the distribution of the maximum of $U(t)$ during a production period which corresponds to the minimum of the inventory level $Y(t)$ during a production period. We define the stopping times

$$\tau^* = \inf[t : U(t) = 0 \quad \text{or} \quad U(t) \geq K] \tag{2.25}$$

and the probability that $U(t)$ reaches or exceeds $K$ before the end of the current production period by

$$q(w) = Pr\{U(\tau^*) \geq K | U(0) = w\}. \tag{2.26}$$

The distribution of the maximum of $U(t)$ during a production period, $M(K)$, is then given by

$$M(K) = Pr\{\max_{[0,\tilde{\tau}]} U(t) > K\} = \int_0^K q(w)\mu e^{-\mu w} dw + e^{-\mu K} \tag{2.27}$$

where $\tilde{\tau}$ is the length of a production period. The conditional probability $q(w)$ satisfies

$$\hat{L}q(w) = 0; \quad q(0) = 0; \quad q(w) = 1, \quad w \geq K. \tag{2.28}$$

The solution of (2.28) is

$$q(w) = \frac{\int_0^w \mu\alpha(z) e^{\mu z} e^{-\mu \int_0^z \alpha(s)ds}\, dz}{\int_0^K \mu\alpha(z) e^{\mu z} e^{-\mu \int_0^z \alpha(s)ds}\, dz + e^{\mu K} e^{-\mu \int_0^K \alpha(s)ds}}, \tag{2.29}$$

from which we obtain

$$M(K) = \left[ e^{\mu K} e^{-\mu \int_0^K \alpha(s)ds} + \int_0^K \mu\alpha(z) e^{\mu z} e^{-\mu \int_0^z \alpha(s)ds}\, dz \right]^{-1}. \tag{2.30}$$

ACKNOWLEDGEMENTS

This research was supported in part by NSF Grants DMS-87-03110 and DMS-86-20267, and grants from the U.S.-Israel Binational Science Foundation and the Israel Academy of Sciences. C. K. was partially supported by an N.S.F. Postdoctoral Fellowship.

## REFERENCES

1. D.P. Gaver, Jr., *Operating characteristics of a simple production, inventory control model,*, Oper. Res. 9 (1961), 635-649.
2. D.P. Heyman and M.J. Sobel, "Stochastic Models in Operations Research Vol.1," McGraw-Hill, New York, 1982.
3. J. Keilson and N.D. Mermin, *The second order distribution of integrated shot noise*, Trans. I.R.E. 5 (1959), 75-77.
4. D.P. Gaver and R.G. Miller, *Limiting distributions for some storage problems*, Studies in Applied Probability and Management Science eds. K. J. Arrow, S. Karlin and H. Scarf (1962), Stanford University Press, Stanford, Calif..
5. J.W. Cohen, "The Single Server Queue," North-Holland, New York, 1982.
6. L. Kleinrock, "Queueing Systems I, II," Wiley, New York, 1975, 1976.
7. C. Knessl, B.J. Matkowsky, Z. Schuss and C. Tier, *Asymptotic analysis of a state-dependent M/G/1 queueing system*, SIAM J. Appl. Math. 46 (1986), 483-505.

[1]Dept. of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, Chicago, IL 60680, [2]Dept. of Engineering Sciences and Applied Mathematics, Northwestern University, Evanston, IL 60208, [3]Dept. of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel