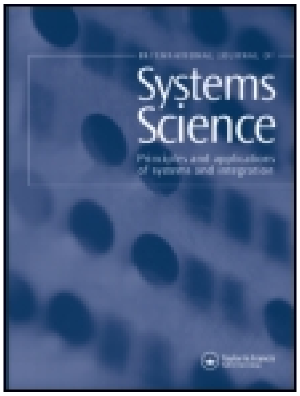

International Journal of Systems Science
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tsys20

Optimal management for infinite capacity N-policy M/G/1 queue with a removable service station

Y.C. Chang [a] & W.L. Pearn [b]
[a] Department of Industrial Engineering and Management , Ching Yun University , Taiwan
[b] Department of Industrial Engineering and Management , National Chiao Tung University , Taiwan

Published online: 19 Apr 2011.

To cite this article: Y.C. Chang & W.L. Pearn (2011) Optimal management for infinite capacity N-policy M/G/1 queue with a removable service station, International Journal of Systems Science, 42:7, 1075-1083, DOI: 10.1080/00207721.2011.570480

To link to this article: http://dx.doi.org/10.1080/00207721.2011.570480

International Journal of Systems Science
Vol. 42, No. 7, July 2011, 1075–1083

Optimal management for infinite capacity N-policy M/G/1 queue with a removable service station

Y.C. Chang[a]* and W.L. Pearn[b]

[a]*Department of Industrial Engineering and Management, Ching Yun University, Taiwan;*
[b]*Department of Industrial Engineering and Management, National Chiao Tung University, Taiwan*

(*Final version received 8 March 2007*)

In this article, we consider an infinite capacity N-policy M/G/1 queueing system with a single removable server. Poisson arrivals and general distribution service times are assumed. The server is controllable that may be turned on at arrival epochs or off at service completion epochs. We apply a differential technique to study system sensitivity, which examines the effect of different system input parameters on the system. A cost model for infinite capacity queueing system under steady-state condition is developed, to determine the optimal management policy at minimum cost. Analytical results for sensitivity analysis are derived. We also provide extensive numerical computations to illustrate the analytical sensitivity properties obtained. Finally, an application example is presented to demonstrate how the model could be used in real applications to obtain the optimal management policy.

Keywords: analytical results; management policy; M/G/1 queue; sensitivity analysis

1. Introduction

In this article, we use a differential technique to study the optimal management policy of an N-policy M/G/1 queue with infinite capacity. The decision-maker can turn a single server on at customers' arrival epochs or off at service completion (departure) epochs. The term 'removable server' represents the system of turning on and turning off the server, depending on the number of customers in the system. The service times are described by an arbitrary probability distribution with mean $1/\mu$ and arrivals entering the service station follow a Poisson process with mean rate λ. The next job order cannot start processing until the previous service has completed and as long as its queue is non-empty, the server will process job orders at this rate. The processing process is independent of the Poisson arrival process of the job orders, which means that regardless of the number of outstanding job orders, they continue to transmit to the processing centre in rate λ. Arriving job orders form a single waiting line at the processing centre based on the order of their arrival. The processing centre can only process one order at a time, and it takes a zero set-up time to restart the production line.

In the literature, the N-policy queueing systems have been extensively studied. A pioneer work in this field is Yadin and Naor (1963), who first introduced the concept of an N-policy, which turns the server on when the number of customers in the system reaches a certain number, N ($N \geq 1$), and turns the server off when no customer is present. After the server is turned off, it may not operate until N customers are present in the system. Past research works may be categorised into two parts according to whether the system capacity is infinite or finite. For the infinite capacity case, we review existing works treating problems with removable servers. The N-policy M/G/1 queueing system with a reliable server was first studied by Heyman (1968), and was investigated extensively by several researchers including Bell (1971, 1972), Kimura (1981), Tijms (1986), Gakis, Rhee, and Sivazlian (1995), Artalejo (1998) and many others. Recently, Wang and Huang (1995a) developed analytic closed-form solutions for the N-policy $M/E_k/1$ queueing system with a reliable server. Wang, Chang, and Sivazlian (1999) investigated the N-policy $M/H_2/1$ queueing system and provided stability conditions, that is, steady-state conditions. For finite capacity queueing systems with a removable server, Hersh and Brosh (1980) considered a queueing system with a Poisson distribution arrival process and exponential distributed service times operating under the N-policy. Teghem (1987) studied an M/G/1 queueing system in which the removable server applies a (v, N) policy. Analytical explicit solutions for the N-policy $M/E_k/1$

*Corresponding author. Email: ycchang@cyu.edu.tw

ISSN 0020–7721 print/ISSN 1464–5319 online

DOI: 10.1080/00207721.2011.570480
http://www.informaworld.com

queueing system were derived and a sensitivity analysis was performed in Wang and Huang (1995b).

The primary objectives of this article are to perform a sensitivity analysis of the optimal management policy, and to demonstrate the connection of mathematical formulation and computer program implementation. First, we derive the analytical closed-form results on the sensitivity analysis. Then, an efficient *S-PLUS* computer program is used to calculate the optimal value of N and other critical system performance measures under optimal operation condition. To examine the effects of the input parameters on the optimal value, we apply differential technique to study system sensitivity. Also, we carry out extensive computational experiments to illustrate our findings. Finally, without having an intimate knowledge of the mathematics implemented in the model, an application example is presented for demonstration purposes.

2. Steady-state results

Let L_N denote the expected number of customers in the N-policy M/G/1 queue with infinite capacity. From Tijms (1986) and Wang and Ke (2000), we have the following analytical closed-form expression, where $\rho = \lambda/\mu$ and $E[S^2]$ is the second moment of the service time:

$$L_N = \frac{N-1}{2} + \rho + \frac{\lambda^2 E[S^2]}{2(1-\rho)}. \tag{1}$$

Notations for the idle period, the busy period and the busy cycle are defined as follows: (1) the idle period, the length of time the server is turned off per cycle, is denoted by I, (2) the busy period, the length of time when the server is turned on and in operation and customers are being served per cycle, is denoted by B, (3) the busy cycle, from the beginning of the last idle period to the beginning of the following next idle period, is denoted by C. The expected lengths of the idle period, the busy period and the busy cycle are denoted by $E[I]$, $E[B]$ and $E[C]$, respectively. The busy cycle is the sum of the idle period and the busy period, $C = I + B$, or $E[C] = E[I] + E[B]$. Using the results stated in Wang and Ke (2000), we have the long-run fraction of time, for the server is in idle, busy, respectively, and the number of busy cycles per unit time:

$$\frac{E[I]}{E[C]} = 1 - \rho, \tag{2}$$

$$\frac{E[B]}{E[C]} = \rho, \tag{3}$$

$$\frac{1}{E[C]} = \frac{\lambda(1-\rho)}{N}. \tag{4}$$

Empty probability, that there is no customer in the system and no station is in service (the service station is turned off), is given by

$$P_{00} = \frac{1-\rho}{N}. \tag{5}$$

Stability conditions for a stable queueing system are given by Equation (5) with $0 < P_{00} < 1$. With simple algebraic manipulations, we obtain the following inequality, where $\rho = \lambda/\mu$, which is sufficient for stationary conditions,

$$0 < \rho < 1. \tag{6}$$

3. Optimal management policy

We define: (1) $C_h \equiv$ holding cost per unit time for each customer present in the system, (2) $C_a \equiv$ cost per unit time for performing an auxiliary task by the service station in the idle period, (3) $C_o \equiv$ operating cost per unit time for the service station in operation, (4) $C_s \equiv$ start-up cost per unit time for activating the service station while the service station is turned off and (5) $C_d \equiv$ shut down cost per unit time for removing the service station from the service. Our objective is to determine the optimal value of the management parameter N, say N*, to minimise the total expected cost function. Utilising the definition of each cost element, the total expected cost function per unit time per customer is given by (see also Wang and Ke 2000)

$$TC(N) = C_h L_N + C_a \frac{E[I]}{E[C]} + C_o \frac{E[B]}{E[C]} + (C_s + C_d)\frac{1}{E[C]}. \tag{7}$$

We should note that the last two terms of Equation (1) are not function of the decision variable N. Likewise, we note from Equations (2)–(3) that, terms $E[I]/E[C]$, and $E[B]/E[C]$ do not involve the decision variable N. Omitting those cost terms not a function of the decision variable N, the optimisation problem in Equation (7) is equivalent to minimising the following equation:

$$\tilde{T}C(N) = C_h \frac{N-1}{2} + (C_s + C_d)\frac{\lambda(1-\rho)}{N}. \tag{8}$$

Discarding the fixed cost $-(1/2)C_h$ of the first term, Equation (8) reduces to the following expression, subject to $0 < \rho < 1$, and $N = 1, 2, \ldots$

$$\hat{T}C(N) = C_h \frac{N}{2} + (C_s + C_d)\frac{\lambda(1-\rho)}{N}. \tag{9}$$

Since N is a positive integer, $N = 1, 2, \ldots,$ the optimal value N* minimising $TC(N)$ can be determined from the following two inequalities,

$$\hat{T}C(N^*-1) \geq \hat{T}C(N^*),$$
$$\hat{T}C(N^*+1) \geq \hat{T}C(N^*). \tag{10}$$

From Equation (10), the necessary conditions for N^* to be optimal reduce to

$$(N^*-1)N^* \leq \frac{2\lambda(C_s+C_d)(1-\rho)}{C_h} \leq N^*(N^*+1). \tag{11}$$

The optimal value N^* may be determined by giving a particular value of $2\lambda(C_s+C_d)(1-\rho)/C_h$. Note that there might be two simultaneous solutions for Equation (11) which minimise the total expected cost function $TC(N)$. For example, we set a particular value of $2\lambda(C_s+C_d)(1-\rho)/C_h=42$ in Equation (11) and solve for N^* to obtain $N^*=6$ or 7. If N is treated as a continuous variable greater than zero, we present two methods to solve for the optimal of N, say N^*, and convexity of $TC(N)$ will be proved. Note that the *S-PLUS* computer program we used allows one to plot $TC(N)$ versus N to illustrate the convexity property (Figure 8).

Method 1: Differentiating $TC(N)$ with respect to N and setting the result equal to zero yields

$$\frac{C_h}{2} - (C_s+C_d)\frac{\lambda(1-\rho)}{N^2} = 0.$$

Thus, the optimal value of N is approximately given by

$$N^* = \left(\frac{2\lambda(1-\rho)(C_s+C_d)}{C_h}\right)^{1/2}. \tag{12}$$

Differentiate $TC(N)$ with respect to N twice and then substitute

$$N^* = \left(\frac{2\lambda(1-\rho)(C_s+C_d)}{C_h}\right)^{1/2} \text{ to obtain}$$

$$\frac{d^2TC(N^*)}{dN^2} = \sqrt{\frac{C_h^3}{2\lambda(C_s+C_d)(1-\rho)}} > 0, \quad \text{for } \rho < 1, \tag{13}$$

which implies that $TC(N)$ is a concave upward (convex) function and achieves a global minimum when (Wang and Ke 2000)

$$N^* = \left(\frac{2\lambda(1-\rho)(C_s+C_d)}{C_h}\right)^{1/2}. \tag{14}$$

Method 2: From Equation (9) we have the following inequality

$$\hat{T}C(N) = C_h\frac{N}{2} + (C_s+C_d)\frac{\lambda(1-\rho)}{N}$$
$$\geq \sqrt{2\lambda C_h(C_s+C_d)(1-\rho)}, \tag{15}$$

which gives a lower bound of $\hat{T}C(N)$ and indicates that $\hat{T}C(N)$ is a concave upward function with lower bound $\sqrt{2\lambda C_h(C_s+C_d)(1-\rho)}$. Equality in Equation (15) holds when

$$C_h\frac{N}{2} = (C_s+C_d)\frac{\lambda(1-\rho)}{N}. \tag{16}$$

With some algebraic manipulations, we obtain

$$N^* \approx \left(\frac{2\lambda(1-\rho)(C_s+C_d)}{C_h}\right)^{1/2}. \tag{17}$$

Note that the expressions of N^* in Equations (14) and (17) are the same. If N^* is not an integer, the optimal value of N is one of the two integers closest to N^*, the expression may rewrite as

$$N^* = \left(\frac{2\lambda(1-\rho)(C_s+C_d)}{C_h}\right)^{1/2} + \varepsilon. \tag{18}$$

where $\varepsilon \in (-1, 1)$ is a constant.

4. Analytical results for sensitivity analysis

An important part of any modelling study is that of sensitivity analysis, which determines how changes in model parameters would affect system performance. Sensitivity coefficients, defined as the partial derivatives of the model output with respect to the input parameters, are useful in assessing the reliability of the output from a complex model with many uncertainty parameters. A system analyst often concern with how the system performance can be affected by the changes of the input parameters in the recommended queueing service model. Sensitivity study on the queueing model with critical input parameters may provide some answers to this question. In the following, we conduct some sensitivity investigations on the optimal value N^* based on changes in values of the cost parameters C_h, C_a, C_o, C_s, C_d and system parameters λ and μ.

We note that the terms $E[B]/E[C]$, and $E[I]/E[C]$ do not involve the decision variable N. Therefore, we may set the relative cost parameters C_a and C_o to be some fixed constants, say, zero. Further, from Equation (18), it is easy to see that

$$N^* \propto \sqrt{(C_s+C_d)/C_h}.$$

We perform some algebraic manipulation with respect to system parameters λ and μ. By differentiating N^* with respect to λ, we obtain

$$\frac{\partial N^*}{\partial \lambda} = \frac{(1-2\rho)\sqrt{(C_s + C_d)}}{\sqrt{2\lambda C_h(1-\rho)}}. \tag{19}$$

Setting the last equation to be 0 then solving for λ, we find $\lambda = \mu/2$ (note that $\lambda < \mu$ is required). By differentiating $\partial N^*/\partial\lambda$ with respect to λ again and substituting $\lambda = \mu/2$, we can easily show that

$$\frac{\partial^2 N^*}{\partial \lambda^2}\Big|_{\lambda=\mu/2} = -2\sqrt{\frac{2(C_s + C_d)}{C_h \mu^3}} < 0. \tag{20}$$

The above result implies that N^* is a concave downward function with respect to λ, which achieves its maximum at $\lambda = \mu/2$. By differentiating N^* with respect to μ to achieve

$$\frac{\partial N^*}{\partial \mu} = \frac{\rho^2\sqrt{C_s + C_d}}{\sqrt{2\lambda C_h(1-\rho)}} > 0, \tag{21}$$

for any μ with $\lambda < \mu$. Thus, N^* is increasing in μ. We summarise the analytical results for the sensitivity analysis as follows.

(1) N^* increases in λ for $\rho < 1/2$ and decreases in λ for $\rho > 1/2$.
(2) N^* increases in μ.
(3) Cost parameters C_a and C_o do not affect N^*.
(4) N^* is proportional to $\sqrt{(C_s + C_d)/C_h}$. In other words, N^* increases in C_s and C_d whereas decreases in C_h.

Sensitivity analyses indicate the effects of changes in key input parameters on the optimal solution. The derivatives in the calculations may be treated as the 'changing rate' with respect to each changing parameters. A small derivative may result in a small 'changing rate', which means it would not affect the optimal decision value significantly. Consequently, more effort should be made to obtain accurate estimates for those parameters with large 'changing rate'. Further, the optimal management policy is insensitive to the cost elements C_a and C_o in the total expected cost function per unit time per customer. Hence, poor estimates of those cost elements do not affect the optimal value N^*.

The results reveal some interest properties of the M/G/1 queueing system with a removable service station. For low-traffic intensity service systems with $\rho < 1/2$, when arrival customers increase, we should raise the threshold N^* to start serving waiting customers. On the other hand, for high-traffic intensity service systems with $\rho > 1/2$, when arrival customers increase, we should reduce the threshold N^* to start serving waiting customers to maintain low cost. For the service station, as long as it can serve in a faster rate, the system manager should increase the threshold N^*. Operating cost per unit time for the service station in operation and cost per unit time for performing an auxiliary task by the service station may treat as fixed cost to the service system and they would not affect the decision variable N.

We should note that N-policy is used because of expensive start-up and shut down cost per cycle (relative to holding cost), they affect N^* in the following manner. For the same cost ratio (cost per cycle relative to holding cost), we would obtain the same value N^*. As start-up cost per unit time for activating the service station or removable cost per unit time increases, one should increase the threshold N^* to prevent high set-up and shut down costs. When holding cost per unit time for each customer present in the system increase, one should decrease the threshold N^* to avoid high holding cost.

5. Numerical computations

We now perform an extensive numerical study to illustrate these sensitivity analysis results. Our findings reveal that the optimal management policy is sensitive to some input parameters, such as the cost coefficients C_h, C_s, C_d and system parameters λ and μ. Therefore, we may set the insensitivity cost elements C_a and C_o equal to zero. Further, incremental rather than accounting costs are considered since the latter often include such non-incremental elements as overhead. In our investigation, holding cost C_h is set to be 5(0.75)80 or 5(1)85 to cover various levels of, from low to high holding costs. Equation (18) suggests that $N^* \propto \sqrt{(C_s + C_d)/C_h}$. We may treat $(C_s + C_d)$, say C_c, as the cost per cycle, without loss of generality, we assume C_s and C_d to be equal since only the sum of them is concerned. Dealing with the system parameters, we note that $0 < \rho < 1$ is sufficient for steady-state condition. A queueing system may be characterised by $\rho = \lambda/\mu$, which represents the traffic intensity. In our study, a widespread range of ρ is covered.

The sensitivity calculations demonstration may now focus on the four critical input parameters: C_h, C_c, λ and μ. We group them into six possible pairs: C_h and C_c; C_h and λ; C_h and μ; C_c and λ; C_c and μ; λ and μ, under consideration simultaneously in order to study the interaction of these key factors. Individual affection on the optimal solution is examined as well. We consider the following experimental design of system parameters for sensitivity analysis on the optimal value N^* based on changes in considerable input values. We calculate the optimal value N^* for the

Table 1. Parameters settings for various system parameters combinations.

C_h	C_c	λ	μ	Parameter setting (1)	Parameter setting (2)
(1)	(2)	0.3	0.8	5(0.75)80	100(50)5100
(1)	(2)	0.3	0.8	5(0.75)80	100, 3000, 5000
(1)	(2)	0.3	0.8	5, 20, 50	100(50)5100
(1)	1600	(2)	1	5(1)85	0.05(0.01)0.85
(1)	1600	(2)	1	5(1)85	0.25, 0.55, 0.85
(1)	1600	(2)	1	5, 10, 50	0.05(0.01)0.85
(1)	1600	0.8	(2)	5(1)85	1(0.1)10
(1)	1600	0.8	(2)	5(1)85	1, 3, 5
(1)	1600	0.8	(2)	5, 10, 50	1(0.1)10
5	(1)	(2)	1	100(50)4100	0.05(0.01)0.85
5	(1)	(2)	1	100(50)4100	0.25, 0.55, 0.85
5	(1)	(2)	1	100, 2000, 4000	0.05(0.01)0.85
5	(1)	0.8	(2)	100(50)4100	1(0.1)10
5	(1)	0.8	(2)	100(50)4100	1, 3, 5
5	(1)	0.8	(2)	100, 2000, 4000	1(0.1)10
5	1600	(1)	(2)	0.05(0.009)0.95	1(0.09)10
5	1600	(1)	(2)	0.05(0.009)0.95	1, 3.07, 5.05
5	1600	(1)	(2)	0.32, 0.68, 0.95	1(0.09)10

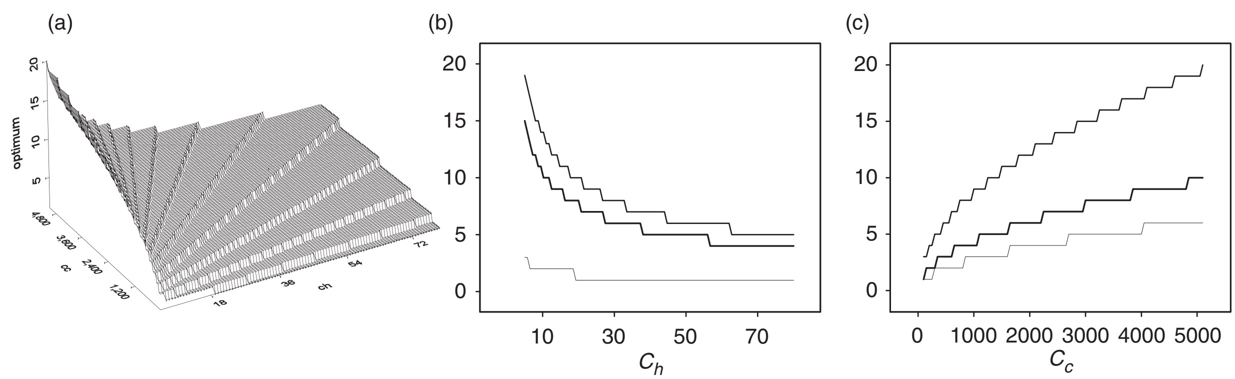

Figure 1. (a) Surface plot of N^* versus $C_h = 5(0.75)80$ and $C_c = 100(50)5100$; (b) plots of N^* versus $C_h = 5(0.75)80$, $C_c = 100, 3000, 5000$ (bottom to top in plot) and (c) plots of N^* versus $C_h = 5, 20, 50$ and $C_c = 100(50)5100$ (top to bottom in plot).

parameters settings summarised in Table 1, which cover a widespread range of applications dealing with the referred queueing model and plot the results in Figures 1–6. For example, rows 2–4 list the parameters settings for various combinations of C_h and C_c. The specified range $C_h = 5(0.75)80$, which means that C_h is set to the range [5, 80] in incremental steps of size 0.75, and $C_c = 100(50)5100$ are considered in which the ratio C_c/C_h covers widespread cost relationships. Rows 3 and 4 are chosen to examine the sensitivity of N^* versus C_h or C_c once a time. In row 3, for various C_h increase from 5 to 80 by 0.75, three levels of $C_c = 100$, 3000 and 5000 are selected. In row 4, three levels of $C_h = 5$, 20 and 50 and $C_c = 100(50)5100$ are considered. Figure 1(a) plots the surface of N^* versus C_h and C_c. Figure 1(b) and (c) shows the cross-section, which plot the curves of N^* versus C_h and C_c, respectively.

We observe from Figure 1(a)–(c) that: (1) N^* increases in C_c but decreases in C_h, (2) N^* increases in the ratio C_c/C_h. Figure 2(a)–(c) reveals that: (1) N^* increases in λ for $\rho < 1/2$ and decreases in λ for $\rho > 1/2$, (2) N^* decreases in C_h. From Figure 3(a)–(c) we observe that: N^* increases in μ but decreases in C_h. We observe from Figure 4(a)–(c) that: (1) N^* increases in C_c, (2) N^* increases in λ for $\rho < 1/2$ and decreases in λ for $\rho > 1/2$. Figure 5(a)–(c) reveals that: N^* increases both in C_c and in μ. From Figure 6(a)–(c) we observe that: (1) N^* increases in λ for $\rho < 1/2$ and decreases in λ for $\rho > 1/2$. The 'local maximum' $\mu/2$ moves from left to right as μ increases. If μ is large enough, one could see that N^* increases in λ, (2) N^* increases in μ.

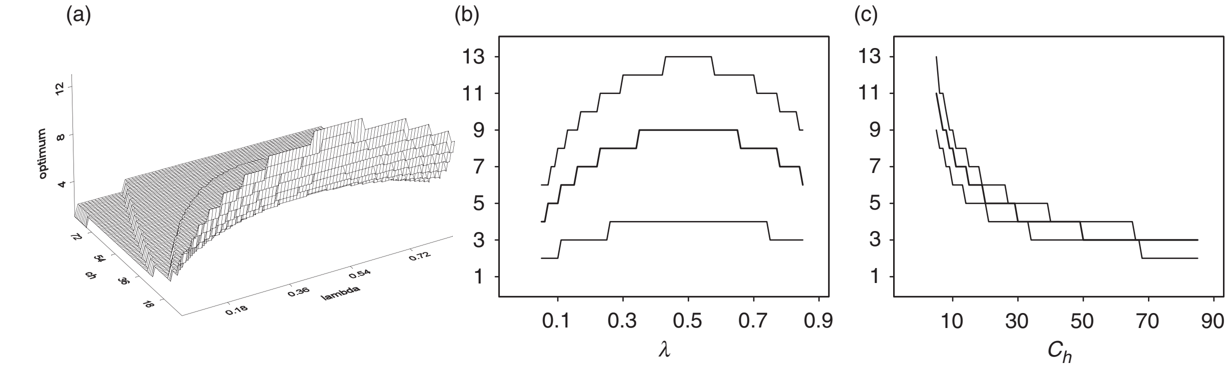

Figure 2. (a) Surface plot of N^* versus $C_h = 5(1)85$ and $\lambda = 0.05(0.01)0.85$; (b) plots of N^* versus $C_h = 5, 10, 50$ and $\lambda = 0.05(0.01)0.85$ (top to bottom in plot) and (c) plots of N^* versus $C_h = 5(1)85$ and $\lambda = 0.25, 0.55, 0.85$ (bottom to top in plot).

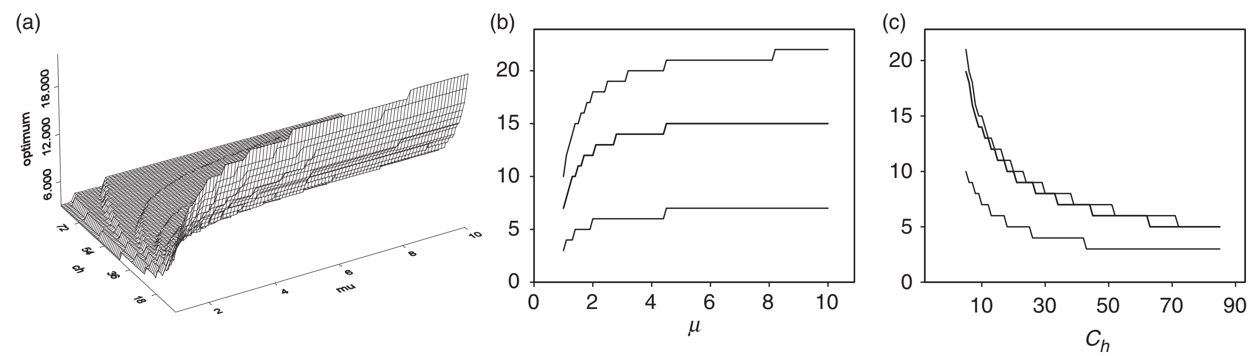

Figure 3. (a) Surface plot of N^* versus $C_h = 5(1)85$ and $\mu = 1(0.1)10$; (b) plots of N^* versus $C_h = 5, 10, 50$ and $\mu = 1(0.1)10$ (top to bottom in plot) and (c) plots of N^* versus $C_h = 5(1)85$ and $\mu = 1, 3, 5$ (bottom to top in plot).

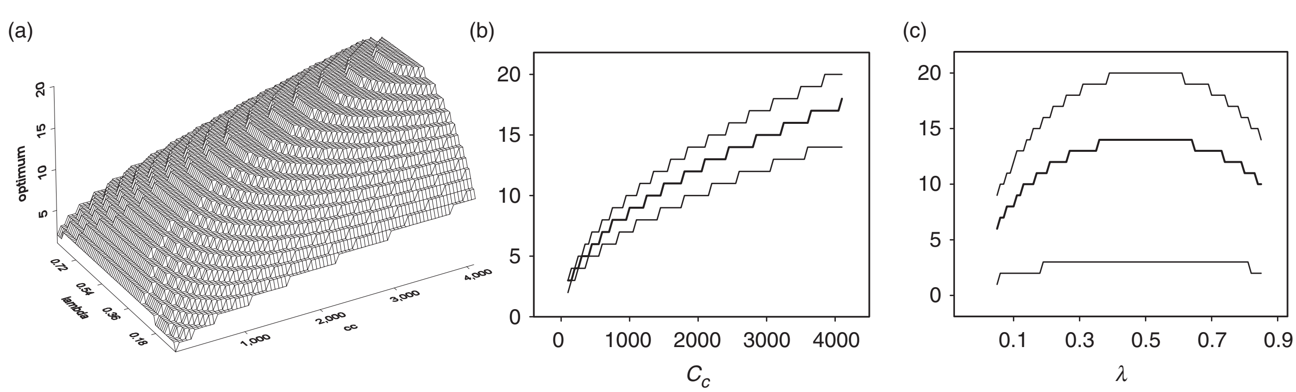

Figure 4. (a) Surface plot of N^* versus $C_c = 100(50)4100$ and $\lambda = 0.05(0.01)0.85$; (b) plots of N^* versus $C_c = 100(50)4100$, $\lambda = 0.25, 0.55, 0.85$ (bottom to top in plot) and (c) plots of N^* versus $C_c = 100, 2000, 4000$ and $\lambda = 0.05(0.01)0.85$ (bottom to top in plot).

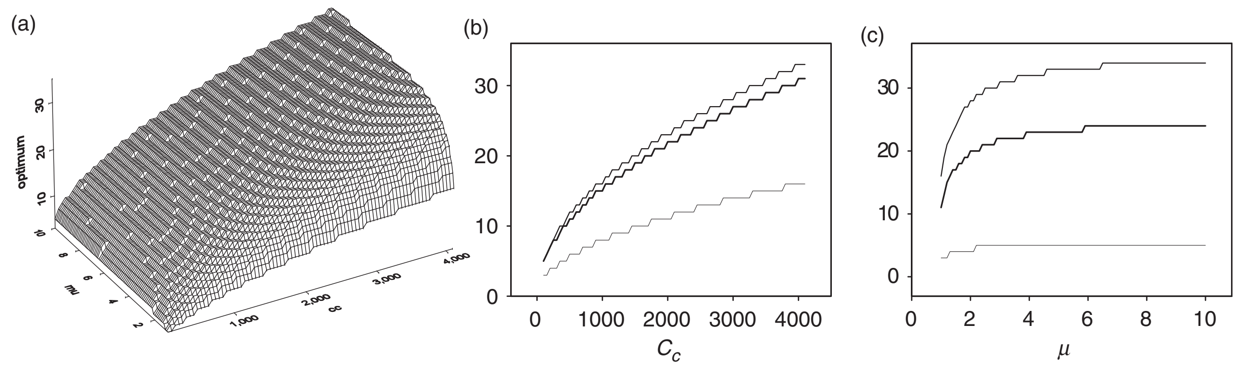

Figure 5. (a) Surface plot of N^* versus $C_c = 100(50)4100$ and $\mu = 1(0.1)10$; (b) plots of N^* versus $C_c = 100(50)4100$ and $\mu = 1, 3, 5$ (bottom to top in plot) and (c) plots of N^* versus $C_c = 100, 2000, 4000$ and $\mu = 1(0.1)10$ (bottom to top in plot).

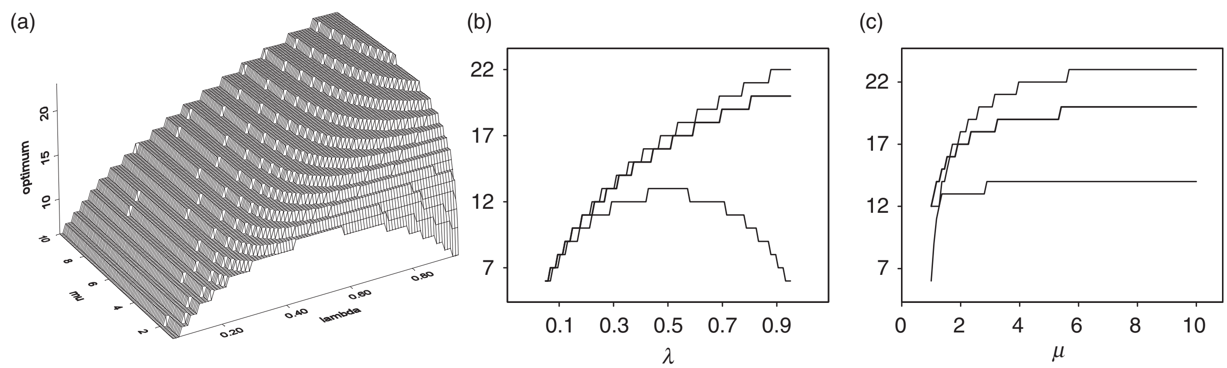

Figure 6. (a) Surface plot of N^* versus $\lambda = 0.05(0.009)0.95$ and $\mu = 1(0.09)10$; (b) plots of N^* versus $\lambda = 0.05(0.009)0.95$ and $\mu = 1, 3.07, 5.05$ (bottom to top in plot) and (c) plots of N^* versus $\lambda = 0.32, 0.68, 0.95$ and $\mu = 1(0.09)10$ (bottom to top in plot).

6. An application example

A routine maintenance in computer communication systems is now presented for illustrative purposes. In addition to transmitting and receiving data, processors in communication systems also perform a variety of testing and processing data, which is considered the processor's primary activity, while the maintenance is considered a secondary activity. The way in which the maintenance is scheduled relative to data management and processing is dependent upon system requirements. Since maintenance activity is often divided into small tasks, whenever the processor finds that there are no primary jobs in the system to service, it begins to work on a maintenance task. For the purpose of our research work, however, we think of a communication network as a network whose purpose is to interconnect a set of applications that are implemented on processing centre and managed by an N-policy. The N-policy applied to control the queueing system is probability due to expensive start-up and shut down cost per cycle or to fully utilise the server.

In the idle period, the processor keeps working on this maintenance task until the processor finds that N or more primary jobs have accumulated in the system, it resumes working on primary jobs until there are no job orders in the system. An N-policy model that activates the server when there are N customers waiting for service and deactivates the server when there are no customers in the system can be also defined as a queueing system in which the idle time of the server may be utilised for other secondary jobs, for our case to work on a maintenance task. The N-policy systems are easily particularised to model many practical situations where the server's effort is divided between primary and secondary customers by specifying an appropriate server scheduling discipline. From the perspective of primary customers, work performed on secondary customers is equivalent to perform an auxiliary task by the server in the so-called idle period with fixed cost.

Data management and data processing have higher priority over the maintenance activity. However, the maintenance tasks are never pre-empted. When primary jobs are being severed, the system behaves as a typical single-queue, single-server system. When primary jobs are absent from the system, the server

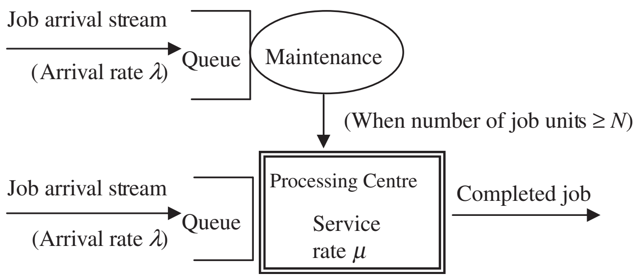

Figure 7. A schematic diagram of the routine operations in computer communication systems.

(processor) performs a maintenance task until finding at least N primary jobs in the system. Figure 7 shows a schematic diagram of the routine operations in computer communication systems.

The program output is shown as follows:

$N = 7$,
$TC(N) = 157.5714$,
$E[I] = 8.75$,
$E[B] = 35$,
$1/E[C] = 0.02285714$.

System characteristics calculations for the model do not require complicated intermediate functions to be implemented, and most of the system performance measures usually of interest can be calculated in a straightforward way. In the example investigated, input system parameters the job order stream arrival rate $\lambda = 0.8$ job/s, the job order processing (service) rate $\mu = 0.8$ job/s and cost element holding cost per second for each job present in the system set to $C_h = 10$, and cost per second for performing an auxiliary task by the service station or cost per second for keeping the server off set to $C_a = 0$, the cost per second for keeping the processor (service station) operating set to $C_o = 0$, the start-up cost for turning the processor on set to $C_s = 800$, the removable cost per second for removing the processor set to $C_d = 800$. The summary of the model inputs are tabulated in Table 2.

The *S-PLUS* computer program gives the expected length of idle period $E[I] = 8.75$ s, the expected length of processing (busy) period $E[B] = 35$ s and the expected number of busy cycles per second $1/E[C] = 0.02$. The value of N for the optimal management policy is $N^* = 7$ units of job orders, and the corresponding minimum expected cost is found to be $TC(N^*) = 157.57$. Figure 8 plots the expected cost $TC(N)$ versus $N = 1(1)30$. It shows that the minimum expected cost indeed occurs when $N = 7$, and the tendency of $TC(N)$ versus N could be easily observed. We summarise the model outputs in Table 3. We have given an example to illustrate how a system analyst can use computer program such as *S-PLUS* to calculate system performance measures, the optimal value of N and its minimum expected cost. The application example demonstrates the levels of detail that are appropriate for building a model and using that model for performance projection. The example illustrates the relationship between modelling concepts, evaluation

Table 2. Model input parameter values.

System parameters and cost elements	Notation	Value
Job order stream arrival rate	λ	0.8
Job order processing (service) rate	μ	1
Holding cost per second per job order present	C_h	10
Cost per second for performing an auxiliary task	C_a	0
Cost per second for keeping the processor on	C_o	0
Start-up cost per second for turning the processor on	C_s	800
Shut-down cost per second for turning the processor off	C_d	800

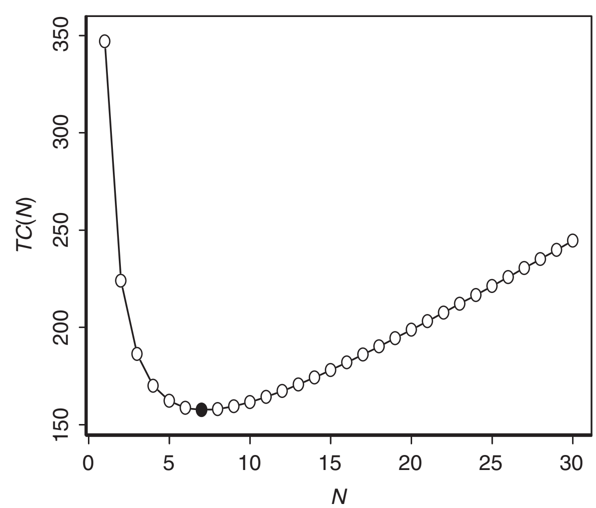

Figure 8. Plot of $TC(N)$ versus N for $N = 1(1)30$.

Table 3. Model output for system performance measures.

System performance measures	Notation	Value
Optimal management policy	N^*	7
Minimum expected cost	$TC(N^*)$	157.57
Expected length of idle period	$E[I]$	8.75
Expected length of busy period	$E[B]$	35
Number of busy cycles per second	$1/E[C]$	0.02

algorithms and implementation. It also indicates how such implementation can save the cost by the analyst.

Notes on contributors

Y.C. Chang received his PhD degree in Industrial Engineering and Management from National Chiao Tung University, Taiwan. Currently, he is an Associate Professor of Quality Management and Operations Research at the Department of Industrial Engineering and Management, Ching Yun University, Taiwan. His research interests include process capability analysis and queueing systems management.

W.L. Pearn is a Professor of Operations Research and Quality Assurance at National Chiao Tung University, Taiwan. He received his PhD degree in Operations Research from the University of Maryland, College Park, MD, USA. He worked at AT&T Bell Laboratories as a quality research scientist before joining National Chiao Tung University. His research interests include process capability, network optimisation, and production management. His publications appeared in *Journal of the Royal Statistical Society*, *Journal of Quality Technology*, *Journal of Applied Statistics*, *Statistics, Journal of the Operational Research Society*, *European Journal of Operations Research*, *Operations Research Letters*, *Omega, Networks*, *International Journal of Productions Research*, and others.

References

Artalejo, J.R. (1998), 'Some Results on the M/G/1 Queue with *N*-policy', *Asia-Pacific Journal of Operational Research*, 15, 147–157.

Bell, C.E. (1971), 'Characterization and Computation of Optimal Policies for Operating an M/G/1 Queueing System with Removable Server', *Operations Research*, 19, 208–218.

Bell, C.E. (1972), 'Optimal Operation of an M/G/1 Priority Queue with Removable Server', *Operations Research*, 21, 1281–1289.

Gakis, K.G., Rhee, H.K., and Sivazlian, B.D. (1995), 'Distributions and First Moments of the Busy and Idle Periods in Controllable M/G/1 Queueing Models with Simple and Dyadic Policies', *Stochastic Analysis and Applications*, 13, 47–81.

Hersh, M., and Brosh, I. (1980), 'The Optimal Strategy Structure of an Intermittently Operated Service Channel', *European Journal of Operational Research*, 5, 133–141.

Heyman, D.P. (1968), 'Optimal Operating Policies for M/G/1 Queueing System', *Operations Research*, 16, 362–382.

Kimura, T. (1981), 'Optimal Control of an M/G/1 Queueing System with Removable Server via Diffusion Approximation', *European Journal of Operational Research*, 8, 390–398.

Teghem Jr, J. (1987), 'Optimal Control of a Removable Server in an M/G/1 Queue with Finite Capacity', *European Journal of Operational Research*, 31, 358–367.

Tijms, H.C. (1986), *Stochastic Modeling and Analysis: A Computational Approach*, New York, NY: Wiley.

Wang, K.-H., Chang, K.-W., and Sivazlian, B.D. (1999), 'Optimal Control of a Removable and Non-reliable Server in an Infinite and Finite $M/H_2/1$ Queueing System', *Applied Mathematical Modeling*, 23, 651–666.

Wang, K.-H., and Huang, H.-M. (1995a), 'Optimal Control of an $M/E_k/1$ Queueing System with a Removable Service Station', *Journal of the Operational Research Society*, 46, 1014–1022.

Wang, K.-H., and Huang, H.-M. (1995b), 'Optimal Control of an $M/E_k/1$ Queueing System with Finite Capacity', *Microelectronics and Reliability*, 35, 1023–1030.

Wang, K.-H., and Ke, J.-C. (2000), 'A Recursive Method to the Optimal Control of an M/G/1 Queueing System with Finite Capacity and Infinite Capacity', *Applied Mathematical Modeling*, 24, 899–914.

Yadin, M., and Naor, P. (1963), 'Queueing Systems with a Removable Service Station', *Operational Research Quarterly*, 14, 393–405.