Methodology and Computing in Applied Probability, 7, 97–118, 2005

Computation of the Optimal Policy for the Control of a Compound Immigration Process through Total Catastrophes

EPAMINONDAS G. KYRIAKIDIS* kyriak@aegean.gr
Department of Financial and Management Engineering, University of the Aegean, 31 Fostini Str., 82100, Chios, Greece

THEODOSIS D. DIMITRAKOS dimitheo@aegean.gr
Department of Statistics and Actuarial Science, University of the Aegean, Karlovassi, 83200, Samos, Greece

Received December 12, 2003; Revised October 29, 2004; Accepted October 29, 2004

Abstract. In this paper we consider a Markov decision model introduced by Economou (2003), in which it was proved that the optimal policy in the problem of controlling a compound immigration process through total catastrophes is of control-limit type. We show that the average cost of a control-limit policy is unimodal as a function of the critical point. This result enables us to design very efficient algorithms for the computation of the optimal policy as the bisection procedure and a special-purpose policy iteration algorithm that operates on the class of control-limit policies.

Keywords: compound Poisson process, Markov decision process, control-limit policies, minimum average cost

AMS 2000 Subject Classification: Primary 90C40; Secondary 60J25

1. Introduction

The Markov decision process is an appropriate mathematical model for the optimal control of a Markov process under various optimization criteria. The most widely used optimization criteria are the minimization of the expected total discounted cost and the minimization of the expected long-run average cost per unit time. The standard algorithms for the computation of the optimal policy are the policy-iteration algorithm, the value-iteration algorithm and the linear programming algorithm. We refer to Chapter 3 of Tijms's (1994) book for the description of these algorithms and for various applications. Surveys of real applications of Markov decision processes, in which the theoretical results have had some influence on actual decision making, are given in White (1985, 1988).

In many Markov decision problems it can be shown that the optimal policy initiates the controlling action if and only if the state of the process (e.g., number of customers in a queue, degree of deterioration of a machine, size of a biological population) is greater

*Corresponding author. P. Zervou 13, P. Psychico, 15452, Athens, Greece.

than or equal to a critical level. Such a policy is usually called control-limit policy and the critical level the control-limit. This characterization of the structure of the optimal policy may accelerate considerably its computation since in many problems it is possible to design efficient algorithms, which operate on the class of control-limit policies and converge to the optimal policy. We refer to the papers of Abakuks (1979), Federgruen and So (1989), Benyamini and Yechiali (1999) and Love et al. (2000) in which suitable algorithms for the determination of the optimal policy were developed.

A standard method that often leads to the proof of the optimality of a control-limit policy is the method of successive approximations. This method first establishes the optimality of a control-limit policy for the corresponding finite-horizon discounted cost problem. Then, this structural property is transferred to the infinite-horizon discounted cost problem and, finally, to the average-cost problem. Economou (2003) introduced a Markov decision process in continuous time for the control of a compound immigration process through total catastrophes. Using the method of successive approximations, he proved that the average-cost optimal policy is of the control-limit type. In this article we prove that in Economou's model the average cost of a control-limit policy is a unimodal function of the critical point. This result enables us to develop very efficient algorithms for the determination of the optimal policy as the bisection procedure or a tailor-made policy iteration algorithm.

In Section 2, the description of the model is given.

2. The Model

Consider a population of individuals, which grows stochastically according to a compound Poisson process with rate $\lambda > 0$ and group-size distribution $\{ g_j : j = 1, 2, \ldots\}$ with $g_1 > 0$. We assume that the individuals are damaging in some sense. For example, the individuals may be insects, which destroy a crop or spread a disease. We refer to such individuals as pests. The damage done by the pests is represented by a cost c_i, $i \geq 0$, for each unit of time during which the population size is i. It is assumed that the sequence $\{c_i\}$ is non-decreasing and $c_i \leq Ai^m$ for some constant $A > 0$ and integer m. Furthermore, we introduce the following conditions on the group-size distribution: (i) The probability generating function $G(z) = \sum_{j=1}^{\infty} g_j z^j$ of $\{g_j : j = 1, 2, \ldots\}$ has radius of convergence strictly greater than 1 and (ii) the group-size distribution has finite moments of all orders up to m.

We suppose that there is a controller who observes the evolution of the population continuously and may take an action that introduces total catastrophes whenever a new state is entered. Spraying the crop or the insects with some insecticide may be appropriate controlling action for the population growth in the above case. It is assumed that the catastrophe rate is equal to $\mu > 0$. Thus, when the controlling action is taken, the length of time until the occurrence of a catastrophe is exponentially distributed with mean μ^{-1}. Moreover, the controlling action gives rise to costs due to labour, materials, risk etc. Let the cost of taking controlling action be k per unit time, where k is a positive constant.

A stationary policy f is defined by a sequence $\{f_i\}$, $i \geq 0$, where f_i is the action taken when the process is at state i. It is assumed that $f_i = 1$ when the controlling action is being taken and $f_i = 0$ when the controlling action is not being taken. If the stationary policy $f \equiv \{f_i\}$, $i \geq 0$, is employed, our assumptions imply that we have a continuous-time Markov chain model for the population growth of the pests with state space $S = \{0, 1, \ldots\}$ and the following transitions in a time interval $(t, t + \delta t)$:

$$i \to i+j \quad \text{with probability } \lambda g_j \delta t + o(\delta t), i \geq 0, \ j \geq 1,$$

$$i \to 0 \qquad \text{with probability } \mu f_i \delta t + o(\delta t), i \geq 1.$$

The expected long-run average cost per unit time of a policy π is defined as the limit as $t \to \infty$ of the expected cost incurred in the time interval $[0, t]$ divided by t, given that the policy π is used. Our goal is to find a policy that minimizes the expected long-run average cost per unit time among all policies. The decision epochs include the epochs at which an arrival of a group of pests occurs. We place no restrictions on the class of allowable policies. Thus, the action chosen by a policy, when the process is at state $i \geq 1$, may depend on the history of the process or it may be randomized in the sense that it chooses action $a \in \{0, 1\}$ with some probability that depends on a. It seems reasonable that the optimal policy belongs to the class of control-limit policies $\{P_n, n = 1, 2, \ldots\}$, where P_n is the stationary policy under which the controlling action is taken if and only if the population size of the pests is greater than or equal to n.

Economou (2003) studied the above problem in the case in which the sequence $\{c_i\}$ is bounded above. He transformed the original model to an equivalent discrete-time Markov decision model using the usual uniformization technique (see e.g., Serfozo (1979)) and then, using the method of successive approximations (see e.g., Ross (1992)), proved that the average-cost optimal policy is either the one that keeps the catastrophe mechanism off at all times or it is of control-limit type. In the last section of the paper, he proposed an algorithm for the approximate determination of the stationary probabilities under the policy P_n. A disadvantage of the uniformization technique is that the proof of the above result only holds in the case in which the set of allowable policies consists solely of stationary policies.

In the next section of the present paper we will show that the optimal policy is of control-limit type without using the uniformization technique. The same result will be obtained in the case in which the sequence $\{c_i\}$ is unbounded. The optimality of a control-limit policy is valid even if we introduce a fictitious cost r incurred for each unit of time the process is occupying the state 0. In Section 4 of the paper it is shown that for each $n = 1, 2, \ldots$ the policy P_n is optimal when the fictitious parameter r belongs to a suitable interval of the real line. A consequence of this result is that the average cost of the policy P_n is unimodal as a function of $n \geq 1$. This enables us to construct very efficient algorithms for the computation of the optimal policy as the bisection procedure and a special-purpose policy iteration algorithm that operates on the class of control-limit policies. In Section 5 we present the description of the algorithms and we give some numerical results. Furthermore, we propose a method for the exact computation of the stationary probabilities under the policy P_n. In Section 6 we consider a more general model in which the controlling action introduces binomial catastrophes. An explanation for a technical manipulation in the proof of Proposition 3 is given in the appendix.

3. The Form of the Optimal Policy

In this section and in Section 4 we introduce a fictitious cost r incurred each unit of time the process is occupying the state 0. We assume that r can be any real number. The main result of the paper, Proposition 4, is a corollary of Theorem 1, which is based on the variation of the parameter r over the entire real line.

We prove below that the average-cost optimal policy is of control-limit type. We first consider the case in which the sequence $\{c_i\}$ is bounded above and then the case in which the sequence $\{c_i\}$ tends to infinity.

Case 1: Suppose that the sequence $\{c_i\}$ is bounded above. Following the usual line of proof (see e.g., Kyriakidis (1999a), Van der Duyn Schouten and Vanneste (1995)) we first consider the corresponding finite-horizon discounted cost problem. The relevant theory can be found in Chapter 7 of Ross's (1992) book and in Chapter 5 of Heyman and Sobel's (1984) book. The minimum n-step expected discounted cost $V_\alpha(i,n)$, if $i \geq 1$ is the initial state, satisfies for $n = 1, 2, \ldots$ the following equations (see e.g., Heyman and Sobel (1984), p. 202).

$$V_\alpha(i,n) = \min\left\{ \frac{c_i + \lambda\sum_{j=1}^{\infty} g_j V_\alpha(i+j, n-1)}{\alpha+\lambda}, \frac{c_i + k + \lambda\sum_{j=1}^{\infty} g_j V_\alpha(i+j, n-1) + \mu V_\alpha(0, n-1)}{\alpha+\lambda+\mu} \right\}, i \geq 1,$$

with initial value

$$V_\alpha(i,0) = 0, i \geq 1.$$

It can be easily proved by induction on n that $\{V_\alpha(i,n)\}$ is non-decreasing in $i \geq 1$, for all $n = 0, 1, \ldots$. The limit of $V_\alpha(i,n)$ as $n \to \infty$ is equal to the minimum total expected discounted cost $V_\alpha(i)$, with initial state i (see Theorem 7.3 on p. 158 in Ross (1992)). Hence, it is deduced that $\{V_\alpha(i)\}, i \geq 1$, is non-decreasing in i.

Assume now that the initial state is $i \geq 1$. Using the same arguments as in Kyriakidis (1999a) we have that

$$V_\alpha(i) \leq B\left(\frac{\lambda}{\mu} + 1\right) + V_\alpha(0), i \geq 1, \tag{1}$$

where,

$$B = \frac{\sup_{i \geq 1} c_i + k}{\lambda}.$$

For i = 1, 2, . . . we have that

$$\begin{aligned}|V_\alpha(i) - V_\alpha(0)| &\leq |V_\alpha(i) - V_\alpha(1)| + |V_\alpha(1) - V_\alpha(0)| \\ &= V_\alpha(i) - V_\alpha(1) + |V_\alpha(1) - V_\alpha(0)| \\ &= V_\alpha(i) - V_\alpha(0) + V_\alpha(0) - V_\alpha(1) + |V_\alpha(1) - V_\alpha(0)| \\ &\leq B\left(\frac{\lambda}{\mu} + 1\right) + V_\alpha(0) - V_\alpha(1) + |V_\alpha(1) - V_\alpha(0)|,\end{aligned}$$

where the first equality follows from the monotonicity of the sequence $\{V_\alpha(i)\}, i \geq 1$, and the second inequality follows from (1). Therefore, the condition of Theorem 7.7 in Ross (1992) is satisfied, since the quantity $|V_\alpha(i) - V_\alpha(0)|$ is uniformly bounded in $i \geq 0$ and $\alpha > 0$. It follows that there exists a bounded sequence $\{h_i\}$, i = 0, 1,... and a constant G such that

$$h_i = \lim_{n\to\infty}[V_{\alpha_n}(i) - V_{\alpha_n}(0)],\ i \geq 0, \tag{2}$$

for some sequence $\alpha_n \to 0$, and

$$h_i = \min\left\{\frac{c_i - G}{\lambda} + \sum_{j=1}^{\infty} g_j h_{i+j},\ \frac{c_i + k - G}{\lambda + \mu} + \sum_{j=1}^{\infty} \frac{\lambda g_j}{\lambda + \mu} h_{i+j} + \frac{\mu}{\lambda + \mu} h_0\right\},\ i \geq 1. \tag{3}$$

Furthermore, there exists a stationary optimal average-cost policy that chooses at each state $i \geq 1$ the minimizing action on the right-hand side in (3). Hence the optimal policy takes controlling action at $i \geq 1$ if and only if

$$\frac{\lambda}{\mu}k + \lambda h_0 + G \leq c_i + \lambda \sum_{j=1}^{\infty} g_j h_{i+j} \tag{4}$$

From (2) it follows that the sequence $\{h_i\}$, $i \geq 1$, is non-decreasing in i, since $\{V_\alpha(i)\}$, $i \geq 1$, is non-decreasing in i. From the monotonicities of the sequences $\{c_i\}$ and $\{h_i\}$, $i \geq 1$, we have the following result:

PROPOSITION 1 *If the sequence* $\{c_i\}$ *is bounded above, the optimal policy is either the policy that never takes controlling action or the control-limit policy* P_{x^*}, *where*, $x^* = \min\{\text{integer } i \geq 1 \text{ such that (4) holds}\}$.

Note that the above result generalizes Theorem 2.3 in Economou (2003), which holds only in the case in which the admissible policies are only the stationary ones.

Case 2: Suppose that $c_i \to \infty$, as $i \to \infty$. The stationary probabilities under the policy P_n, $n \geq 1$, are asymptotically geometrically distributed since the probability generating function of $\{g_j : j = 1, 2, \ldots\}$ has radius of convergence strictly greater than 1 (see Theorem 3.1 in Economou (2003)). This result together with the inequality $c_i \leq Ai^m$, that was introduced in the previous section, guarantee the finiteness of the average cost of every control-limit policy. Note also that the process under the policy that never introduces catastrophes does not have an equilibrium distribution. Hence the average cost in this case is infinite for every initial state since $\{c_i\}$ tends to infinity (see Lemma 2.2.2. in Bather (1976)). Furthermore, as in the case in which the sequence $\{c_i\}$ is bounded above it can be shown that the minimum total expected discounted cost $V_\alpha(i), i \geq 1$, is non-decreasing in i.

Sennott (1989) proved that in a semi-Markov decision process an average-cost optimal stationary policy exists if five assumptions are satisfied. The transition rates and the cost structure of the present model are such that these assumptions are satisfied. Specifically, the first assumption states that there exist $\delta > 0$ and $\varepsilon > 0$, such that, for every state and action, there is a probability of at least ε that the transition time will be greater than δ. The second assumption states that there exists $B > 0$ such that $T(i, a) \leq B$ for every i and a, where $T(i, a)$ is the one-step expected time if action a is taken in state i. In the present model the first assumption holds with $\delta = 1$, $\varepsilon = \exp[-(\lambda + \mu)]$ and the second assumption holds with $B = (\lambda + \mu)^{-1}$. The other assumptions deal with properties of $V_\alpha(i), i \geq 0$, and are satisfied if the following condition holds (see Proposition 2 and Proposition 3 in Sennott (1989)):

Condition: There exists a stationary policy f with average cost b that has the following properties:

(P1) f induces an irreducible ergodic embedded Markov chain on S and

$$\sum_{i \in S} p_i(f) C(i, f(i)) < \infty \text{ and } \sum_{i \in S} p_i(f) T(i, f(i)) < \infty,$$

where, $p_i(f)$, $i \in S$, are the steady-state probabilities of the embedded Markov chain under the policy f and $C(i, a)$ is the one-step expected cost if action a is taken in state i.

(P2) There exists $\varepsilon > 0$ and a finite subset G of S such that $d(i) \geq b + \varepsilon$ for $i \in S - G$, where $d(i)$ is the minimum cost rate with respect to action a in state i. Furthermore, for each $i \in G$, there exists a stationary policy $f^{(i)}$ such that $c_{0i}(f^{(i)}) < \infty$, where $c_{0i}(f^{(i)})$ represents the expected cost of a first passage from 0 to i under the policy $f^{(i)}$.

In the present model the above condition holds if we choose $f = P_1$. The property (P1) is satisfied since the steady-state probabilities under P_1 are asymptotically geometrically distributed. The property (P2) holds if we choose any $\varepsilon > 0$, $G = \{0, 1, \ldots, j\}$ with $c_j \geq b + \varepsilon$ and $f^{(i)} = P_j$, $i \in G$.

Hence, in the present model, Sennott's five assumptions are satisfied. From Theorem 2 in Sennott (1989) it follows that there exists a constant G and a non-decreasing sequence $\{h_i\}$, $i \geq 1$, such that the optimal policy prescribes at state $i \geq 1$ the minimizing action on the right-hand side of (3). This remark leads us to the following result.

PROPOSITION 2 *If $c_i \to \infty$, as $i \to \infty$, then the optimal policy is control-limit.*

4. The Form of the Average Cost Function Under a Control-Limit Policy

Let $T_{i0}^{(n)}$ and $C_{i0}^{(n)}$, $i \geq 0$, be the expected time and cost, respectively, until the process under the policy P_n, $n \geq 1$, reaches the state 0, given that the initial state is i. Let also G_n

denote the expected long-run average cost per unit time under the policy P_n. The process under the policy P_n is a regenarative process, where the successive entries into state 0 can be taken as regenerative epochs between successive cycles. From a well-known regenerative argument (see Proposition 5.9 in Ross (1992)) it follows that G_n is equal to the expected cost of a cycle divided by the expected time of a cycle. Hence,

$$G_n = \frac{\frac{r}{\lambda} + \sum_{i=1}^{\infty} g_i C_{i0}^{(n)}}{T_{00}^{(n)}} \tag{5}$$

Let $h_i^{(n)}$, $i \geq 0$, be the relative values associated with the policy P_n, $n \geq 1$, defined by (see relation (3.1.7) in Tijms (1994))

$$h_i^{(n)} = C_{i0}^{(n)} - G_n T_{i0}^{(n)} \tag{6}$$

Clearly,

$$h_0^{(n)} = 0, \tag{7}$$

since $G_n = C_{00}^{(n)} / T_{00}^{(n)}$, by the usual regenerative argument. Note also that

$$T_{i0}^{(n)} = \mu^{-1}, \quad i \geq n. \tag{8}$$

In the next proposition a necessary and sufficient condition is given for the optimality of the policy P_n.

Proposition 3 *The policy P_n, $n \geq 1$, is optimal if and only if*

$$\frac{c_i - G_n}{\lambda} + \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} \leq \frac{c_i + k - G_n + \lambda \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} + \mu h_0^{(n)}}{\lambda + \mu}, \; 1 \leq i \leq n-1, \tag{9}$$

and

$$\frac{c_i + k - G_n + \lambda \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} + \mu h_0^{(n)}}{\lambda + \mu} \leq \frac{c_i - G_n}{\lambda} + \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)}, \quad i \geq n. \tag{10}$$

Proof: Suppose that the policy P_n is optimal. According to relation (1) in Nobel and Tijms (1999), the numbers $h_i^{(n)}$, $i \geq 0$, and G_n satisfy the following system of linear equations:

$$h_i^{(n)} = \frac{c_i - G_n}{\lambda} + \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)}, \quad 0 \leq i \leq n-1,$$

$$h_i^{(n)} = \frac{c_i + k - G_n + \lambda \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} + \mu h_0^{(n)}}{\lambda + \mu}, \quad i \geq n.$$

Assume that for some $\hat{i}$, $0 \le \hat{i} \le n-1$, inequality (9) does not hold. Thus, in view of relation (2) in Nobel and Tijms (1999), the stationary policy f with

$$f(i) = 0, 0 \le i \le n-1, i \ne \hat{i},$$
$$f(\hat{i}) = 1,$$
$$f(i) = 1,\ i \ge n,$$

has smaller average cost than the average cost of the policy P_n. This is a contradiction. Hence inequality (9) holds. Similarly, it can be shown that inequality (10) is valid.

Suppose that the policy P_n satisfies the inequalities (9) and (10). Hence the numbers G_n and $h_i^{(n)}$, $i \ge 0$, satisfy the following equations:

$$h_i^{(n)} = \min\left\{\frac{c_i - G_n}{\lambda} + \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)},\ \frac{c_i + k - G_n + \lambda \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} + \mu h_0^{(n)}}{\lambda + \mu}\right\},\ i \ge 1. \tag{11}$$

We distinguish two cases:

Case 1: Suppose that the sequence $\{c_i\}$ is bounded above. Then, the numbers $h_i^{(n)}$, $i \ge 0$, are bounded above and hence from Theorem 7.6 in Ross (1992) it follows that the policy P_n is optimal.

Case 2: Suppose that $c_i \to \infty$, as $i \to \infty$. From Proposition 2 it follows that there exists some positive integer u such that the policy P_u is optimal.

Let X_N and a_N be the state of the process and the action chosen, respectively, after the N-th transition. We assume that X_0 and a_0 are the initial state and initial action, respectively. Also, let $T(i, a)$ and $C(i, a)$ be the expected time and cost, respectively, until a transition occurs when action a is chosen in state i. From the proof of Theorem 7.6 in Ross (1992), it follows that

$$G_n \le \frac{E_u\left[h_{X_N}^{(n)} - h_{X_0}^{(n)}\right] + E_u \sum_{i=1}^{N} C(X_{i-1}, a_{i-1})}{E_u \sum_{i=1}^{N} T(X_{i-1}, a_{i-1})}, \tag{12}$$

where E_u represents the conditional expectation given that the policy P_u is employed. Relation (12) holds with equality for $n = u$. Using the definition of $h_{X_N}^{(n)}$, relation (12) can be written as

$$G_n \le \frac{E_u C_{X_N 0}^{(n)} - G_n E_u T_{X_N 0}^{(n)} - h_{X_0}^{(n)} + E_u \sum_{i=1}^{N} C(X_{i-1}, a_{i-1})}{E_u \sum_{i=1}^{N} T(X_{i-1}, a_{i-1})},$$

or, equivalently,

$$G_n \le \frac{\sum_{j=0}^{\infty} p_{X_0 j}^{(N)} C_{j0}^{(n)} - G_n \sum_{j=0}^{\infty} p_{X_0 j}^{(N)} T_{j0}^{(n)} - h_{X_0}^{(n)} + E_u \sum_{i=1}^{N} C(X_{i-1}, a_{i-1})}{E_u \sum_{i=1}^{N} T(X_{i-1}, a_{i-1})},$$

where, $p_{X_0 j}^{(N)}$ is the probability that the state of the process under the policy P_u after N transitions will be j. Letting $N \to \infty$ and using the bounded convergence theorem (see appendix for an explanation in detail), we can interchange the order of limit and summation in the numerator of the above expression. Hence,

$$G_n \leq \frac{\sum_{j=0}^{\infty} p_j C_{j0}^{(n)} - G_n \sum_{j=0}^{\infty} p_j T_{j0}^{(n)} - h_{X_0}^{(n)}}{\lim_{N\to\infty} E_u \sum_{i=1}^{N} T(X_{i-1}, a_{i-1})} + \lim_{N\to\infty} \frac{E_u \sum_{i=1}^{N} C(X_{i-1}, a_{i-1})}{E_u \sum_{i=1}^{N} T(X_{i-1}, a_{i-1})},$$

where, p_j, $j \geq 0$, are the equilibrium probabilities of the corresponding embedded Markov chain of the process under the policy P_u. The numerator of the first term of the expression on the right-hand side in the above relation is finite since p_j, $j \geq 0$, are asymptotically geometrically distributed (see Theorem 3.1 in Economou (2003)). Since the denominator, as $N \to \infty$, tends to infinity, we deduce that this term is zero. The second term is the expected long-run average cost per unit time G_u of the policy P_u. Hence $G_n \leq G_u$, and thus, the policy P_n is optimal. ■

The results of Lemmas 1, 2 and 3 will be used in the proof of Theorem 1.

LEMMA 1 *For each fixed $n \geq 1$, the sequence $C_{i0}^{(n)}$, $i \geq n$, is non-decreasing in i.*

Proof: Let $X(s)$, $s \geq 0$, be the population size at time s of the (uncontrolled) compound immigration process. Conditioning on the time until the occurrence of the catastrophe we obtain that

$$C_{i0}^{(n)} = \int_0^{\infty} \left[\int_0^t E\{c_{X(s)} + k \mid X(0) = i\} ds \right] \mu e^{-\mu t} dt$$

$$= \int_0^{\infty} \left[\int_0^t E\{c_{X(s)+i} \mid X(0) = 0\} ds \right] \mu e^{-\mu t} dt + k\mu^{-1}, \tag{13}$$

where the first equality follows from the cost structure of the problem as specified in Section 2 and the second equality follows from the fact that the distribution of $X(s)$, if $X(0) = i$, coincides with the distribution of $X(s) + i$, if $X(0) = 0$. From the monotonicity of the sequence $\{c_i\}$ and the above expression we deduce that $C_{i0}^{(n)}$ is non-decreasing in $i \geq n$. ■

LEMMA 2 *Assume that the policy P_n, $n \geq 1$, is optimal for some fixed value R of the parameter r. Then, it is impossible for the policy P_n to be optimal for all $r \geq R$.*

Proof: Suppose that the policy P_n is optimal for all $r \geq R$. Then the numbers $h_i^{(n)}$, $i \geq 0$, satisfy (10) at $i = n$ for all $r \geq R$. Hence,

$$\frac{c_n + k - G_n + \mu h_0^{(n)}}{\lambda + \mu} + \lambda \sum_{j=1}^{\infty} \frac{g_j}{\lambda + \mu} h_{n+j}^{(n)} \leq \frac{c_n - G_n}{\lambda} + \sum_{j=1}^{\infty} g_j h_{n+j}^{(n)}.$$

Using (6), (7) and (8) the above relation reduces to

$$(\lambda+\mu)G_n \leq \mu c_n + \lambda\mu\sum_{j=1}^{\infty} g_j C_{n+j,0}^{(n)} - \lambda k.$$

From (5), it follows that $G_n \to \infty$, as $r \to \infty$. This contradicts the above inequality because its right-hand side does not depend on r and the coefficient of G_n is positive. ■

Let $P_{n(r)}$ denote the optimal control-limit policy, where $n(r)$ is the optimal critical number. In view of Propositions 1 and 2 the policy $P_{n(r)}$ is the overall optimal policy.

LEMMA 3 *The optimal critical value n(r) is non-decreasing with respect to* $r \in (-\infty, +\infty)$.

Proof: Suppose that

$$r_1 < r_2 \tag{14}$$

Let n_1, n_2 be the optimal critical numbers that correspond to the parameters r_1 and r_2, respectively. Let also G_1 and G_2 the average costs of the policies P_{n_1} and P_{n_2}, respectively. Relation (14) implies that

$$G_1 < G_2 \tag{15}$$

We want to show that $n_1 \leq n_2$. Assume to the contrary that $n_1 > n_2$. From the optimality of P_{n_1} and the relation (9) with $i = n_1 - 1$ we deduce that

$$\frac{c_{n_1-1} - G_1}{\lambda} + \sum_{j=1}^{\infty} g_j h_{n_1-1+j}^{(n_1)} \leq \frac{c_{n_1-1} + k - G_1 + \lambda\sum_{j=1}^{\infty} g_j h_{n_1-1+j}^{(n_1)} + \mu h_0^{(n_1)}}{\lambda+\mu}.$$

From (6), (7) and (8) we deduce that the above relation is equivalent to

$$G_1(\lambda+\mu) \geq \mu c_{n_1-1} + \lambda\mu\sum_{j=1}^{\infty} g_j C_{n_1-1+j,0}^{(n_1)} - \lambda k. \tag{16}$$

From the optimality of P_{n_2} and the relation (10) with $i = n_2$ we obtain that

$$\frac{c_{n_2} + k - G_2 + \lambda\sum_{j=1}^{\infty} g_j h_{n_2+j}^{(n_2)} + \mu h_0^{(n_2)}}{\lambda+\mu} \leq \frac{c_{n_2} - G_2}{\lambda} + \sum_{j=1}^{\infty} g_j h_{n_2+j}^{(n_2)}.$$

From (6), (7) and (8) we deduce that the above relation is equivalent to

$$G_2(\lambda+\mu) \leq \mu c_{n_2} + \lambda\mu\sum_{j=1}^{\infty} g_j C_{n_2+j,0}^{(n_2)} - \lambda k. \tag{17}$$

From (15), (16) and (17) we have that

$$\mu c_{n_2} + \lambda\mu\sum_{j=1}^{\infty} g_j C_{n_2+j,0}^{(n_2)} - \lambda k > \mu c_{n_1-1} + \lambda\mu\sum_{j=1}^{\infty} g_j C_{n_1-1+j,0}^{(n_1)} - \lambda k.$$

However the left-hand side of the above relation is smaller or equal to its right-hand side since $c_{n_2} \leq c_{n_1-1}$ and $C^{(n_2)}_{n_2+j,0} \leq C^{(n_1)}_{n_1-1+j,0}$. This is a contradiction. ■

In Theorem 1 below it is shown that for each $n = 1, 2, \ldots$ the policy P_n is optimal when the fictitious parameter r belongs to a suitable interval.

THEOREM 1 *These exists an increasing sequence* $\{R_n\}$, $n \geq 1$, *with* $R_1 = -\infty$, $R_2 > -\infty$, *and* $\lim_{n\to\infty} R_n = +\infty$ *such that the policy* P_n, $n \geq 1$, *is optimal for all* $r \in [R_n, R_{n+1}]$, *where* $R_{n+1} = \sup\{w : w \geq R_n$ *and the policy* P_n *is optimal for all* $r \in [R_n, w]\}$.

Proof: The proof is by induction on n. We first establish that a number $R > -\infty$ exists such that the policy P_1 is optimal for all $r \leq R$. In view of Proposition 3 it suffices to show that the numbers $h^{(1)}_i$, $i \geq 1$, and G_1 satisfy the inequality

$$\frac{c_i + k - G_1 + \lambda\sum_{j=1}^{\infty} g_j h^{(1)}_{i+j} + \mu h^{(1)}_0}{\lambda+\mu} \leq \frac{c_i - G_1}{\lambda} + \sum_{j=1}^{\infty} g_j h^{(1)}_{i+j}, \ i \geq 1.$$

Using (6), (7) and (8) the above relation reduces to

$$(\lambda+\mu)G_1 \leq \mu c_i + \lambda\mu\sum_{j=1}^{\infty} g_j C^{(1)}_{i+j,0} - \lambda k, \quad i \geq 1.$$

From (5), we have that $G_1 \to -\infty$, as $r \to -\infty$. Thus, there exists a number $R > -\infty$ such that the above inequalities hold for all $r \leq R$. From Lemma 2, it follows that $R_2 < +\infty$, where $R_2 = \sup\{w : w \geq R$ and the policy P_1 is optimal for all $r \leq w\}$.

Suppose that there exists a sequence $R_1 < R_2 < \ldots < R_n$, such that the policy P_s, $1 \leq s \leq n$, is optimal for all $r \in [R_s, R_{s+1}]$ with $R_{n+1} = \sup\{w : w \geq R_n$ and the policy P_n is optimal for all $r \in [R_n, w]\} < +\infty$. We will show that the policy P_{n+1} is optimal for $r = R_{n+1}$. Consider some $\varepsilon > 0$ such that the policy P_n is not optimal for $r = R_{n+1} + \varepsilon$. From Proposition 3 it follows that one of the following two cases occurs.

Case 1: For some i with $1 \leq i \leq n-1$,

$$\frac{c_i + k - G_n + \lambda\sum_{j=1}^{\infty} g_j h^{(n)}_{i+j} + \mu h^{(n)}_0}{\lambda+\mu} < \frac{c_i - G_n}{\lambda} + \sum_{j=1}^{\infty} g_j h^{(n)}_{i+j}$$

Using (6) and (7) we deduce that the above inequality is equivalent to $\psi_i(R_{n+1} + \varepsilon) > 0$ with

$$\psi_i(r) = \mu c_i - \mu G_n + \lambda\mu\sum_{j=1}^{\infty} g_j\left(C^{(n)}_{i+j,0} - G_n T^{(n)}_{i+j,0}\right) - \lambda k.$$

Since G_n, as given in (5), is increasing in r, we deduce that $\psi_i(r)$ is decreasing in r. Thus,

$$0 < \psi_i(R_{n+1} + \varepsilon) < \psi_i(R_{n+1}) \leq 0,$$

where the last inequality follows from the optimality of P_n for $r = R_{n+1}$. Clearly, this is a contradiction and therefore Case 2 must arise.

Case 2: For some i with $i \geq n$,

$$\frac{c_i - G_n}{\lambda} + \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} < \frac{c_i + k - G_n + \lambda \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} + \mu h_0^{(n)}}{\lambda + \mu}.$$

Using (6), (7) and (8) we deduce that the above inequality is equivalent to $\psi_i(R_{n+1} + \varepsilon) < 0$ with

$$\psi_i(r) = \mu c_i - (\lambda + \mu) G_n + \lambda \mu \sum_{j=1}^{\infty} g_j C_{i+j,0}^{(n)} - \lambda k.$$

From Lemma 1 and the monotonicity of $\{c_i\}$ it follows that the sequence $\psi_i(r)$, $i \geq n$, is non-decreasing in i. Thus,

$$\psi_n(R_{n+1} + \varepsilon) \leq \psi_i(R_{n+1} + \varepsilon) < 0.$$

By the definition of R_{n+1} and the above analysis, there exists a sequence $\{\varepsilon_k\} \downarrow 0$ with $\psi_n(R_{n+1} + \varepsilon_k) < 0$. From the continuity of $\psi_n(r)$ in r, it follows that $\psi_n(R_{n+1}) \leq 0$. However $\psi_n(R_{n+1}) \geq 0$ because the policy P_n is optimal for $r = R_{n+1}$. Thus $\psi_n(R_{n+1}) = 0$. The last equality means that for $r = R_{n+1}$ the right-hand side of (11) attains the same value at the state $i = n$ for both actions 0 and 1. Therefore, the action prescribed, for each state $i \geq 0$, by the policy P_{n+1} minimizes the right-hand side of (11). Thus the policy P_{n+1} is optimal for $r = R_{n+1}$. Define

$$R_{n+2} = \sup\{w : w \geq R_{n+1} \text{ and the policy } P_{n+1} \text{ is optimal for all } r \in [R_{n+1}, w]\}.$$

From Lemma 2, it follows that $R_{n+2} < +\infty$. It remains to be shown that $\lim_{n\to\infty} R_n = +\infty$. Assume to the contrary that $\lim_{n\to\infty} R_n = R^* < +\infty$. Choose $n^* > n(R^*)$, where $n(R^*)$ the optimal critical number given that the value of the fictitious parameter is R^*. From Lemma 3 it follows that $n(R^*) \geq n(R_{n^*})$. Consider some state i such that $n^* > i \geq n(R_{n^*})$. For $r = R_{n^*}$ at state i the optimal policy chooses action 1. This contradicts the optimality of the policy P_{n^*} for $r = R_{n^*}$. ■

REMARK We point out that, when the sequence $\{C_i\}$ is bounded above, Theorem 1 is valid without the conditions (i) and (ii), that we introduced in Section 2 on the group-size distribution $\{g_j : j = 1, 2, \ldots\}$, since these conditions are not needed in the proof of Proposition 3 in this case.

LEMMA 4 *The average cost G_n of a control-limit policy P_n, $n \geq 1$, can be expressed as*

$$G_n = rA_n + \widetilde{G}_n,$$

where, A_n is non-increasing in n and $\widetilde{G}_n$ is independent of r.

Proof: From (5) we see that G_n can be written as

$$G_n = \frac{r}{\lambda T_{00}^{(n)}} + \widetilde{G}_n,$$

where, $\widetilde{G}_n$ is independent of r. It suffices to show that $T_{00}^{(n)}$, $n \geq 1$, is non-decreasing in n. Note that

$$T_{00}^{(n)} = T_{0E_n}^{(n)} + \mu^{-1}, \tag{18}$$

where $E_n = \{n, n+1, \ldots\}$ and $T_{0E_n}^{(n)}$ is the expected time until entry into the set of states E_n, given that the initial state is 0 and the policy P_n is used. For $n \geq 2$, since $E_n \subset E_{n-1}$, from the structure of the model it follows that

$$T_{0E_n}^{(n)} \geq T_{0E_{n-1}}^{(n)}, \tag{19}$$

and

$$T_{0E_{n-1}}^{(n)} = T_{0E_{n-1}}^{(n-1)}. \tag{20}$$

From (18), (19) and (20) we deduce that

$$T_{00}^{(n)} \geq T_{0E_{n-1}}^{(n)} + \mu^{-1} = T_{0E_{n-1}}^{(n-1)} + \mu^{-1} = T_{00}^{(n-1)}. \qquad \blacksquare$$

Using Lemma 4, the following result, which is useful for the computation of the optimal control-limit policy, can be proved in the same way as Lemma 5.2 in Federgruen and So (1989).

PROPOSITION 4 *For any fixed r, the average cost G_n, $n \geq 1$, is unimodal as a function of n.*

5. The Computation of the Optimal Policy

Proposition 4 enables us to develop two efficient algorithms in order to compute the optimal policy. In this section we assume that the parameter r takes the value 0, i.e., we consider the model described in Section 2. The critical point n^* of the optimal control-limit policy P_{n^*} can be found by the standard bisection procedure, which determines the minimum within the set $\{G_n : n \geq 1\}$.

Let N be a positive integer such that $G_N < G_{N+1}$. We state the procedure below.

Bisection Procedure

Step 1. Initialize $n_1 = 1$, $n_2 = N$.

Step 2. Let n be the integer part of $(n_1+n_2)/2$. If $G_n < G_{n+1}$, set $n_2 = n$; if $G_n > G_{n+1}$ set $n_1 = n + 1$; if $G_n = G_{n+1}$ and $G_{n-1} \leq G_n$, set $n_2 = n$; if $G_n = G_{n+1}$ and

$G_{n-1} > G_n$, set $n_1 = n+1$. If $n_2 - n_1 > 1$, repeat Step 2. Otherwise, choose n^* such that $G_{n^*} = \min\{G_{n_1}, G_{n_2}\}$. The optimal policy is P_{n^*} and its average cost is G_{n^*}.

Using a well-known result (see Theorem 5.10 in Ross (1992)), the average costs G_n in Step 2 of the algorithm can be expressed in terms of the stationary probabilities p_i, $i \geq 0$ and the cost rates under the policy P_n as follows:

$$G_n = \sum_{i=0}^{n-1} c_i p_i + \sum_{i=n}^{\infty} (c_i + k) p_i$$

Economou (2003) proposed an algorithm for the approximate determination of p_i, $i \geq 0$. A method for the exact computation of these probabilities is presented below.

By conditioning on the first transition out of state i, $0 \leq i \leq n-1$, we obtain

$$T^{(n)}_{n-1,0} = \lambda^{-1} + \mu^{-1}, \tag{21}$$

$$T^{(n)}_{i0} = \lambda^{-1} + \sum_{j=1}^{n-i-1} g_j T^{(n)}_{i+j,0} + \mu^{-1} \sum_{j=n-i}^{\infty} g_j, \quad 0 \leq i \leq n-2. \tag{22}$$

Since the process under the policy P_n is regenerative we deduce that (see Theorem 5.8 in Ross (1992))

$$p_0 = \frac{\text{expected time in state 0 during one cycle}}{\text{expected time of one cycle}} = \frac{\lambda^{-1}}{T^{(n)}_{00}}.$$

The denominator of the above expression can be computed from (21) and (22). The other stationary probabilities p_i, $i \geq 1$, can be found from the following balance equations:

$$p_i = \sum_{j=0}^{i-1} g_{i-j} p_j, \; 1 \leq i \leq n-1,$$
$$p_i = \frac{\lambda}{\lambda + \mu} \sum_{j=0}^{i-1} g_{i-j} p_j, \; i \geq n.$$

An alternative way for the computation of the optimal policy is the development of a special-purpose policy iteration algorithm that generates a sequence of improved control-limit policies. Proposition 4 guarantees that this sequence converges to the optimal policy. The design of the algorithm is based on the embedding technique of Tijms (see Tijms (1994), p. 234). Similar algorithms have been developed in queueing, inventory and maintenance models (see pp. 234–248 in Tijms (1994), Nobel and Tijms (1999)) and also, in other pest control models (see Kyriakidis (1993, 1999b)). The design of the algorithm is described below.

Equation (13) gives the expected cost $C_{i0}^{(n)}$ until the process under the policy P_n reaches the state 0, given that the initial state is $i \geq n$. According to a well-known property of Laplace transforms (see e.g., relation (C.2) on p. 362 in Tijms (1994)), we have that

$$C_{i0}^{(n)} = \int_0^\infty E\{c_{X(t)+i} \mid X(0) = 0\} e^{-\mu t} dt + k\mu^{-1}$$

Let $p_{0i}(t)$, $i \geq 0$, be the probability that the population size $X(t)$ at time t of the (uncontrolled) compound immigration process is i given that $X(0) = 0$. The above expression takes the following form:

$$C_{i0}^{(n)} = \int_0^\infty \left\{ \sum_{j=0}^{\infty} p_{0j}(t) c_{j+i} \right\} e^{-\mu t} dt + k\mu^{-1}.$$

By interchanging the order of the integral and the series we obtain

$$C_{i0}^{(n)} = \sum_{j=0}^{\infty} p_{0j}^{*} c_{j+i} + k\mu^{-1}, \tag{23}$$

where $p_{0j}^{*} = \int_0^\infty p_{0j}(t) e^{-\mu t} dt$, is the Laplace transform of $p_{0j}(t)$ at the point μ. Let $N(t)$ be the number of arrivals up to time $t \geq 0$ of the groups of pests and let

$$\phi_n(j) = P[X(t) = j \mid N(t) = n, X(0) = 0], \quad j \geq 0, n \geq 1.$$

The probabilities $p_{0j}(t)$ can be computed by conditioning on the variable $N(t)$ as follows:

$$p_{0j}(t) = e^{-\lambda t} I(j = 0) + \sum_{n=1}^{\infty} \phi_n(j) e^{-\lambda t} \frac{(\lambda t)^n}{n!},$$

where, $I(j = 0) = 0$, if $j \neq 0$, and $I(j = 0) = 1$, if $j = 0$. By integrating the above equation we obtain the following expression for p_{0j}^{*} after some simplifications

$$p_{0j}^{*} = \frac{I(j = 0)}{\lambda + \mu} + \sum_{n=1}^{\infty} \phi_n(j) \frac{\lambda^n}{(\lambda + \mu)^{n+1}}. \tag{24}$$

The conditional probabilities $\phi_n(j)$ can be computed recursively from the convolution formula:

$$\phi_n(j) = \sum_{\ell=1}^{j-1} \phi_{n-1}(\ell) g_{j-\ell}, \quad j \geq n, \quad n \geq 2,$$

with

$$\phi_n(j) = 0,\ j < n,\ n \geq 1,$$

$$\phi_1(j) = g_j,\ j \geq 1.$$

It is now clear that the expected costs $C_{i0}^{(n)}$, $i \geq n$, can be computed from (23) and (24). By conditioning on the first transition out of state i, $0 \leq i \leq n-1$, we obtain

$$C_{i0}^{(n)} = \lambda^{-1}c_i + \sum_{j=1}^{n-i-1} g_j C_{i+j,0}^{(n)} + \sum_{j=n-i}^{\infty} g_j C_{i+j,0}^{(n)}, \quad 0 \leq i \leq n-1. \tag{25}$$

The average cost G_n, under the policy P_n can be computed from the expression $G_n = C_{00}^{(n)}/T_{00}^{(n)}$.

We consider now the so-called policy-improvement quantity $Q_n(i;a)$ associated with the policy P_n, which is defined by:

$$Q_n(i;a) = C(i,a) - G_nT(i,a) + \sum_{j=0}^{\infty} p_{ij}\ (a)h_j^{(n)},\ i \geq 0,\ a \in \{0,1\}$$

where $p_{ij}(a)$ is the probability that the next state of the process will be j, given that the present state is i and action a is chosen and $T(i, a)$ and $C(i, a)$ are the corresponding expected time and cost, respectively. When the controlling action is being taken, a takes the value 1, and when the controlling action is not being taken, a takes the value 0.

Suppose that there exists an integer $\tilde{n}$ such that $1 \leq \tilde{n} < n$ and $Q_n(i;1) < h_i^{(n)}$, $\tilde{n} \leq i < n$. According to Theorem 3.2.1 in Tijms (1994), the control-limit policy $P_{\tilde{n}}$, characterized by the critical point $\tilde{n}$, achieves smaller average cost than the policy P_n. Similarly, if there exists an integer $\tilde{n}$ such that $n < \tilde{n}$ and $Q_n(i;0) < h_i^{(n)}$, $n \leq i < \tilde{n}$, according to Theorem 3.2.1 in Tijms (1994), the control-limit policy $P_{\tilde{n}}$ characterized by the critical point $\tilde{n}$, achieves smaller average cost than the policy P_n. When it is not possible to find such a $\tilde{n}$, in view of Proposition 4 the average cost G_n is a minimum within the set $\{G_i : i \geq 1\}$.

The above remarks lead us to develop the following special-purpose policy iteration algorithm, which generates a sequence of improved control-limit policies.

Special-Purpose Policy Iteration Algorithm

Step 1. (initialization). Choose an initial integer n, $n \geq 1$.

Step 2. (value-determination step). For the current policy P_n compute the average cost G_n from the equation $G_n = C_{00}^{(n)}/T_{00}^{(n)}$.

Step 3. (policy improvement step).

(3a) Find, if it exists, the smallest integer $\tilde{n}$ such that $1 \leq \tilde{n} < n$ and $Q_n(i;1) < h_i^{(n)}$ for all $i \in \{\tilde{n}, \tilde{n}+1, \ldots, n-1\}$ and go to Step 2 with n replaced by $\tilde{n}$. Else go to Step (3b).

(3b) Find, if it exists, the largest integer $\tilde{n}$ such that $n < \tilde{n}$ and $Q_n(i;0) < h_i^{(n)}$ for all $i \in \{ n, n+1, \ldots, \tilde{n} - 1\}$ and go to Step 2 with n replaced by $\tilde{n}$.

Step 4. (convergence test). If it is not possible to find an integer $\tilde{n}$ such that (3a) or (3b) is satisfied, then the algorithm is stopped. The final policy is P_n and the corresponding average cost is G_n.

The quantity $Q_n(i;1)$ in Part (3a) of the algorithm is given by:

$$Q_n(i;1) = \frac{c_i + k - G_n + \lambda \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} + \mu h_0^{(n)}}{\lambda + \mu}$$

and the quantity $Q_n(i;0)$ in Part (3b) is given by:

$$Q_n(i;0) = \frac{c_i - G_n}{\lambda} + \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)}$$

The quantities $h_i^{(n)}$, $i \geq 0$, can be computed by the equation $h_i^{(n)} = C_{i0}^{(n)} - G_n T_{i0}^{(n)}$, in which $T_{i0}^{(n)}$ and $C_{i0}^{(n)}$ can be found from (8), (21), (22), (23), (25).

In a great number of examples that we have tested, we note that the special-purpose policy iteration algorithm achieves a very large improvement at the first iteration. For every initial control-limit policy P_n, $n \geq 1$, the average costs of the policies that are generated after the first iteration do not differ significantly from the average cost of the optimal policy. Moreover, there is evidence that the number of iterations of the special-purpose policy iteration algorithm is not especially influenced by the initial critical point.

We give as illustration two numerical examples. In the first example, we assume that $\lambda = 2$, $\mu = 5$ and that the group-size distribution is $(g_1, g_2, g_3) = (0.6, 0.2, 0.2)$. We set $c_i = \sqrt{i}$, $i \geq 0$. The point $n = 20$ is chosen as the initial point in Step 1 of the special-purpose policy iteration algorithm. In Table 1 below, we present the optimal critical point n^* obtained by the algorithms and the corresponding average cost G_{n^*} for different values of the parameter k. We also present the number of iterations required by the special-purpose algorithm and the CPU time (in seconds) of the corresponding

Table 1. The optimal values of n^*, the corresponding average cost G_{n^*}, the number of iterations and the C.P.U. times (in seconds) for different values of k.

k	Optimal Critical Point n^*	Average Cost G_{n^*}	Number of Iterations	CPU Times
10	5	2.28	4(5)	2(3.9)
20	9	3.06	4(6)	2.16(5.3)
50	19	4.35	2(6)	1.08(6.4)
100	31	5.58	4(6)	2.5(5.9)

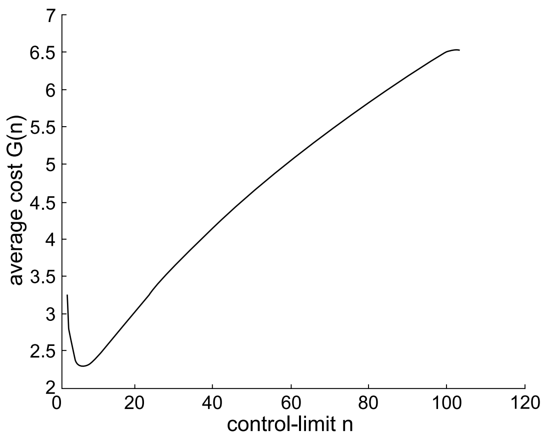
Matlab program that we run on a PC Acer Aspire 1605DLC. In parentheses we present the number of iterations required by the standard bisection procedure and the CPU time.

Note that the optimal critical point n^* increases as k increases and that in all cases the number of iterations and the CPU time required by the special-purpose policy iteration algorithm are smaller than the number of iterations and the CPU time required by the bisection procedure.

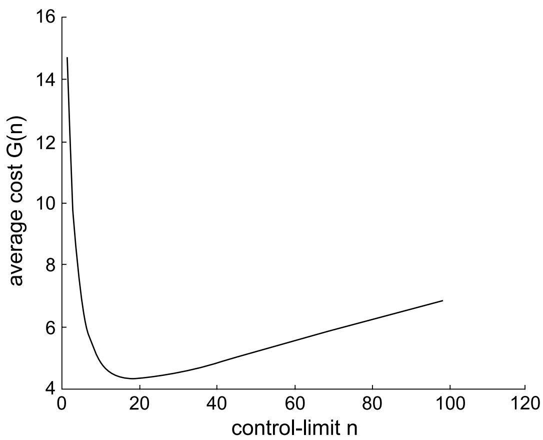
In the second example, we assume that $\lambda = 10$, $\mu = 8$ and that the group-size distribution is $(g_1, g_2, g_3, g_4, g_5) = (0.2, 0.2, 0.2, 0.2, 0.2)$ (see Scenario 4 in Economou (2003)). We set $c_i = 0.5i$, $i \geq 0$ and $k = 30$. In Table 2 below, we give the values of the critical points and the average costs of the successive policies generated by the special-purpose policy iteration algorithm, for various values of the initial critical point.

Table 2. The successive critical points and the corresponding average costs that are generated by the algorithm for different values of the initial critical point.

Initial Critical Point	Successive Critical Points	Successive Average Costs
1	1, 34, 12, 17	18.54, 11.92, 10.36, 10.03
20	20, 16, 17	10.13, 10.04, 10.03
50	50, 9, 19, 17	15.02, 11.12, 10.08, 10.03
70	70, 6, 22, 16, 17	19.44, 12.57, 10.27, 10.04, 10.03
90	90, 5, 24, 15, 17	23.83, 13.59, 10.47, 10.07, 10.03



Graph 1. Plot of G_n versus n for the first example when $k = 10$.



Graph 2. Plot of G_n versus n for the first example when $k = 50$.

For the first example the plots of G_n versus n are given in Graphs 1 and 2 above if k is equal to 10 and 50, respectively. The graphs confirm the result of Proposition 4.

6. Discussion

In this paper we considered the problem of controlling a compound immigration process, which represents a pest population, by the introduction of total catastrophes. It was assumed that the cost rate caused by the pests is an increasing function of their population size and that the cost rate of the controlling action is constant. Using the method of successive approximations we proved the existence of a control-limit policy that minimizes the expected long-run average cost per unit time. Furthermore, we showed that the average cost of a control-limit policy is unimodal as a function of the critical point. This result guarantees that the standard bisection method and a special-purpose policy iteration algorithm converge to the optimal policy.

A more general model is the one in which the pest population is controlled through binomial catastrophes instead of total ones. Populations under the influence of binomial catastrophes were studied by Brockwell et al. (1982). In this case, if the stationary policy $\{f_i\}$, $i \geq 0$, is employed, the transitions of the process in a time interval $(t, t + \delta t)$ are

$$i \to i + j \quad \text{with probability } \lambda g_j \,\delta t + o(\delta t),\ i \geq 0,\ j \geq 1,$$

$$i \to j \quad \text{with probability } \binom{i}{j} p^j (1-p)^{i-j} f_i \,\delta t + o(\delta t),\ i \geq 1,\ 0 \leq j \leq i.$$

The number p, $0 \leq p < 1$, represents the probability of survival of a pest during a catastrophe. If $p = 0$, the model reduces to that introduced in Section 2. It seems again intuitively reasonable that the optimal policy is of control-limit type. However it seems difficult to prove this result.

The optimal policy may be computed by applying the Aproximating Sequence Method introduced by Sennott (see Chapter 10 in Sennott (1999)). This method approximates an infinite-state Markov decision process by a sequence of finite-state Markov decision processes. The value iteration algorithm is used to obtain the optimal policy of each finite-state Markov decision process. The convergence of these policies to the optimal policy of the original infinite-state Markov decision process is guaranteed under suitable assumptions.

After applying the Sennott's method in the modified model, a great number of examples that we tested provide strong evidence that the optimal policy is also of control-limit type. As an illustration, we give a numerical example. We assume that $\lambda = 3$, $p = 0.3$ and that the group-size distribution is $(g_1, g_2, g_3) = (0.4, 0.2, 0.4)$. We set $c_i = i$, $i \geq 0$ and $k = 5$. The critical point n^* of the optimal policy is $n^* = 4$. The number of iterations of the algorithm is 61 and the average cost of the optimal policy is found to be $G = 12.7$.

Appendix

The equations

$$\lim_{N\to\infty}\sum_{j=0}^{\infty} p_{X_0 j}^{(N)} C_{j0}^{(n)} = \sum_{j=0}^{\infty} p_j C_{j0}^{(n)}, \tag{26}$$

$$\lim_{N\to\infty}\sum_{j=0}^{\infty} p_{X_0 j}^{(N)} T_{j0}^{(n)} = \sum_{j=0}^{\infty} p_j T_{j0}^{(n)}, \tag{27}$$

where $p_j = \lim_{N\to\infty} p_{X_0 j}^{(N)}, j \geq 0$, were used in Case 2 in the proof of Proposition 3. These equations are consequences of the following theorem (see p. 351 in Tijms (1994)) which is a special case of the bounded convergence theorem.

THEOREM 2 *Let* $\{p_m, m = 0, 1, \ldots\}$ *be a sequence of non-negative numbers. Suppose that the numbers* a_{nm}, $n, m = 0, 1, \ldots$ *are such that* $\lim_{n\to\infty} a_{nm} = a_m$ *exists for all* $m = 0, 1, \ldots$. *If there is a finite constant* $M > 0$ *such that* $|a_{nm}| \leq M$ *for all* n, m *and if* $\sum_{m=0}^{\infty} p_m < \infty$ *then*

$$\lim_{n\to\infty}\sum_{m=0}^{\infty} a_{nm}\, p_m = \sum_{m=0}^{\infty} a_m p_m.$$

To explain (26) we first see that, conditioning on the time until the occurrence of the catastrophe, the expected cost $C_{j0}^{(n)}, j \geq n$ can be expressed as

$$C_{j0}^{(n)} = \int_0^{\infty}\left[\int_0^t E\{c_{X(s)} + k \mid X(0) = j\}ds\right]\mu e^{-\mu t}dt,$$

where $X(s)$ is the population size at time s of the compound immigration process. Since it was assumed in Case 2 in the Proof of Proposition 3 that $c_i \leq Ai^m$, for some $A > 0$ and integer m, from the above expression it follows that

$$C_{j0}^{(n)} \leq A\int_0^\infty \left[\int_0^t E\{(X(s))^m \mid X(0)=j\}ds\right]\mu e^{-\mu t}dt + k\mu^{-1}.$$

From the definition of the compound Poisson process it follows that the above inequality can be written equivalently as

$$C_{j0}^{(n)} \leq A\int_0^\infty \left[\int_0^t E\{\left(j+Y_1+\ \dots\ +Y_{N(s)}\right)^m\}ds\right]\mu e^{-\mu t}dt + k\mu^{-1},$$

where, $N(s) \sim$ Poisson(λs) and the Y_i's are i.i.d. random variables with $P(Y_i = \ell) = g_\ell$, $\ell = 1, 2, \dots$. Note that the r.h.s. of the above inequality is a polynomial $P(j)$ in j of degree m, since it was assumed in Section 2 that the group-size distribution $\{g_\ell : \ell = 1, 2, \dots\}$ has finite moments of all orders up to m.

The series in the left-hand side of (26) is written as

$$\sum_{j=0}^{\infty} p_{X_0 j}^{(N)} j^2 P(j)\frac{C_{j0}^{(n)}}{j^2 P(j)}. \tag{28}$$

Note that from Theorem 3.1 in Economou (2003) it follows that, for sufficiently large N and j, $p_{X_0 j}^{(N)}$ is approximately equal to $c\theta^{-j}$, with $c > 0$ and $0 < \theta < 1$. Hence, the quantity $p_{X_0 j}^{(N)} j^2 P(j)$ for sufficiently large N and j is bounded. We also have that $\sum_j C_{j0}^{(n)} j^{-2} [P(j)]^{-1} < \infty$, since $C_{j0}^{(n)} < P(j)$. Hence, Theorem 2 can be applied if we take the limit as $N \to \infty$ of (28) and by interchanging the order of limit and summation we obtain the right-hand side of (26).

The equation (27) can be justified similarly.

Acknowledgments

We would like to thank Dr. A. Economou and Dr. A. N. Yannacopoulos for helpful comments. Special thanks are also due to the referee whose comments improved the presentation of the paper.

References

A. Abakuks, "An optimal hunting policy for a stochastic logistic model," *Journal of Applied Probability* vol. 16 pp. 319–331, 1979.

J. Bather, "Optimal stationary policies for denumerable Markov chains in continuous time," *Advances in Applied Probability* vol. 8 pp. 144–158, 1976.

Z. Benyamini and U. Yechiali, "Optimality of control limit maintenance policies under non-stationary deterioration," *Probability in the Engineering and Informational Sciences* vol. 13 pp. 55–70, 1999.

P. J. Brockwell, J. Gani, and S. I. Resnick, “Birth, immigration and catastrophe processes,” *Advances in Applied Probability* vol. 14 pp. 709–731, 1982.

A. Economou, “On the control of a compound immigration process through total catastrophes,” *European Journal of Operational Research* vol. 147 pp. 522–529, 2003.

A. Federgruen and K. C. So, “Optimal time to repair a broken server,” *Advances in Applied Probability* vol. 21 pp. 376–397, 1989.

D. P. Heyman and M. J. Sobel, *Stochastic Models in Operations Research*, vol. 2, McGraw-Hill: New York, 1984.

E. G. Kyriakidis, “A Markov decision algorithm for optimal pest control through uniform catastrophes,” *European Journal of Operational Research* vol. 64 pp. 38–44, 1993.

E. G. Kyriakidis, “Characterization of the optimal policy for the control of a simple immigration process through total catastrophes,” *Operations Research Letters* vol. 24 pp. 245–248, 1999a.

E. G. Kyriakidis, “Optimal control of a truncated general immigration process through total catastrophes,” *Journal of Applied Probability* vol. 36 pp. 461–472, 1999b.

C. E. Love, Z. G. Zhang, M. A. Zitron, and R. Guo, “A discrete semi-Markov decision model to determine the optimal repair/replacement policy under general repairs,” *European Journal of Operational Research* vol. 125 pp. 398–409, 2000.

R. D. Nobel and H. C. Tijms, “Optimal control for an $M^X/G/1$ queue with two service modes,” *European Journal of Operational Research* vol. 113 pp. 610–619, 1999.

S. M. Ross, *Applied Probability Models with Optimization Applications*, Dover: New York, 1992.

L. Sennott, “Average cost semi-Markov decision processes and the control of queueing systems,” *Probability in the Engineering and Informational Sciences* vol. 3 pp. 247–272, 1989.

L. Sennott, *Stochastic Dynamic Programming and the Control of Queueing Systems*, Wiley: New York, 1999.

R. F. Serfozo, “An equivalence between continuous and discrete time Markov decision processes,” *Operations Research* vol. 27 pp. 616–620, 1979.

H. C. Tijms, *Stochastic Models: An Algorithmic Approach*, Wiley: New York, 1994.

F. A. Van Der Duyn Schouten and S. G. Vanneste, “Maintenance optimization of a production system with buffer capacity,” *European Journal of Operational Reseach* vol. 82 pp. 323–338, 1995.

D. J. White, “Real applications of Markov decision processes,” *Interfaces* vol. 15 pp. 73–83, 1985.

D. J. White, “Further real applications of Markov decision processes,” *Interfaces* vol. 18 pp. 55–61, 1988.