Ecology Letters, (2003) **6**: 843–849 doi: 10.1046/j.1461-0248.2003.00493.x

REPORT

Marine reserves and optimal harvesting

Michael G. Neubert
Biology Department, MS-34, Woods Hole Oceanographic Institution, Woods Hole, MA 02543-1049, USA
Correspondence: E-mail: mneubert@whoi.edu

Abstract

Advocates of no-take marine reserves emphasize their conservation benefits. Critics counter that reserves would decrease fisheries yield. Analysis of a spatially explicit harvesting model, however, shows that no-take marine reserves are always part of an optimal harvest designed to maximize yield. The optimal harvest generates a spatial source–sink structure with source populations placed in reserves. The sizes and locations of the optimal reserves depend on a dimensionless length parameter. For small values of this parameter, the maximum yield is obtained by placing a large reserve in the centre of the habitat. For large values of this parameter, the optimal harvesting strategy is a spatial 'chattering control' with infinite sequences of reserves alternating with areas of intense fishing. Such a chattering strategy would be impossible to actually implement, but in these cases an approximate yet practicable policy, utilizing a small number of reserves, can be constructed.

Keywords

Bioeconomics, fisheries, marine protected areas, marine reserves, maximum principle, maximum sustainable yield, optimal control theory, source–sink.

Ecology Letters (2003) 6: 843–849

INTRODUCTION

Advocates of no-take marine reserves emphasize their ability to protect habitat, conserve biodiversity and defend endangered stocks from overexploitation (Halpern & Warner 2002; Lubchenco *et al.* 2003). Their use as a fisheries management tool, however, has remained controversial (Peterson 2002; Shipp 2002). Sceptics argue that the use of no-take reserves would decrease yields relative to optimally managed fisheries. Modelling studies to date (see Gerber *et al.* 2003 for references) have shown that reserves are yield-neutral or produce minor improvements when compared with the best spatially uniform effort-control policies. In some cases, the establishment of a reserve decreases yield. When reserves do increase yield, it is usually when the fishery is already overexploited (Gerber *et al.* 2003).

Most of these studies have used spatially implicit models (National Research Council 2001; Gerber *et al.* 2003, but see Gaines *et al.* 2003 for an exception). The principal advantage of these models is that they are mathematically tractable whereas spatially explicit models often are not. Spatially implicit models, however, do have disadvantages. For example, they usually must assume that reserves would protect a fixed fraction of the stock from harvesting. Actual reserves, on the contrary, protect a fixed area in space through which fish move, and so require spatially explicit models. As spatially explicit models are more difficult to analyse, studies using them have resorted to a comparative method. In this method, the yield produced by the best harvesting strategy from one class of strategies is compared with the yield produced by the best strategy from a different class. The first class typically includes only spatially uniform harvests. The second class usually contains strategies that employ a small number of reserves (usually only one). The problem with the comparative method is that there is no guarantee that either class contains the harvesting strategy that would maximize the yield over *all* feasible strategies.

Here, I take a different approach. Rather than calculating the yield that results from a specified number and arrangement of reserves, I derive the yield-maximizing spatial harvesting strategy in a spatially explicit model in which no reserves are imposed. The model is a combination of the Schaefer harvest model (Schaefer 1957), which is at the foundation of many bioeconomic analyses (Clark 1990), and Fisher's equation (Fisher 1937), a fundamental model in spatial population dynamics (Kot 2001). As it turns out, the optimal harvesting strategy that emerges always includes a no-take reserve. In some cases, the optimal strategy has

©2003 Blackwell Publishing Ltd/CNRS

infinitely many reserves arranged in an intricate network. While the latter policy is clearly impractical, its form suggests a simple approximate strategy, employing only two reserves, that is nearly optimal. I compare this nearly optimal strategy with the best spatially uniform harvesting strategy to determine how much of an improvement it actually represents.

THE MODEL

Let $N(X,T)$ be the population density of the stock at location X and time T. I assume that N is governed by the partial differential equation

$$\frac{\partial N}{\partial T} = rN\left(1 - \frac{N}{K}\right) + D\frac{\partial^2 N}{\partial X^2} - qE(X)N \tag{1}$$

for $0 < x < L$, and the boundary conditions

$$N(0,T) = N(L,T) = 0. \tag{2}$$

Equation 1 states that the rate of change of population density at a given location is controlled by three processes: population growth, movement and harvesting. The first term on the right side of eqn 1 represents logistic population growth. The parameter r is the intrinsic growth rate of the stock and K is the environmental carrying capacity. The second term describes the movement of the population as by diffusion; the parameter D is the diffusion coefficient. The final term describes harvesting. In this model, the harvest rate at a given location is proportional to the product of the stock size, N, and the effort applied at that location, $E(X)$. The proportionality constant q, called the 'catchability coefficient' (Clark 1990) is the per capita harvest rate per unit fishing effort.

Equation 1 holds for all locations inside the habitat, which I assume is a line segment of length L. The boundary conditions (eqn 2) complete the model by setting the population density to zero at the habitat boundaries. In effect, these boundary conditions imply that individuals that cross the boundary immediately die. Thus the model describes a population living in a patch of suitable habitat, of length L, surrounded by unsuitable habitat. There is a 'critical habitat size' L_{cr}; if $L < L_{cr}$ the stock cannot survive even in the absence of harvesting (Kot 2001). For model 1–2,

$$L_{cr} = \pi\sqrt{\frac{D}{r}}. \tag{3}$$

For the subsequent analyses, I will assume $L > L_{cr}$.

I consider a population at equilibrium, which implies $\partial N/\partial T = 0$. We can, therefore, ignore the dependence of N on T [by writing $N(X,T) = N(X)$] and rewrite eqns 1 and 2 as

$$D\frac{d^2N}{dX^2} = -rN\left(1 - \frac{N}{K}\right) + qE(X)N, \tag{4}$$

$$N(0) = N(L) = 0. \tag{5}$$

The equilibrium yield is then given by the integral

$$Y[N(X), E(X)] = \int_0^L qE(X)N(X)dX. \tag{6}$$

If we take the perspective of a sole owner who owns complete rights to exploit the population (Clark 1990), our objective is to find the spatial distribution of effort $E(X)$ that maximizes the yield Y. Because of real limitations on fishing effort, it is reasonable to assume the existence of an upper limit on $E(X)$. We thus try to maximize the yield over all effort functions that satisfy

$$0 \le E(X) \le E_{max}. \tag{7}$$

ANALYSIS AND RESULTS

The first step in our analysis is to rescale the variables using

$$u = \frac{N}{K}, \quad t = rT, \quad x = \frac{X}{\sqrt{D/r}},$$
$$h(x) = \frac{qE(x)}{r}, \quad y = \frac{Y}{K\sqrt{rD}}, \quad \ell = \frac{L}{\sqrt{D/r}}. \tag{8}$$

The new dependent variable u measures the stock density relative to the carrying capacity. The new variable t measures time on a demographic scale; one unit of t is the time in which the population would increase by a factor of e if it were growing exponentially at the rate r. The new spatial variable x measures distance using L_{cr}/π as the unit of length. On this scale, the critical habitat size is $\ell_{cr} = \pi$.

Using the new variables, the fishery management problem is to maximize

$$y[u(x), h(x)] = \int_0^\ell h(x)u(x)dx, \tag{9}$$

over all harvest strategies that satisfy

$$0 \le h(x) \le h_{max}, \tag{10}$$

subject to the state equations

$$\frac{d^2u}{dx^2} = -u(1-u) + h(x)u, \tag{11}$$

and the boundary conditions

$$u(0) = u(\ell) = 0. \tag{12}$$

One advantage of rescaling the variables is immediately apparent: it reduces the number of parameters from six (r, K, D, q, L, and E_{max}) to two (ℓ and h_{max}). As ℓ and h_{max}

©2003 Blackwell Publishing Ltd/CNRS

are the only remaining parameters, we know that the carrying capacity has no qualitative effect on the solution. The other parameters only affect the solution through their effects on ℓ and $h_{\max}$, which simplifies our analysis tremendously. By varying only ℓ and $h_{\max}$, one can find the optimal harvest for every possible combination of the original six parameters.

In the analysis that follows, I explore the effect of ℓ on the optimal strategy. Small values of ℓ correspond to species living in spatially restricted habitats with large diffusion coefficients or small population growth rates. I assume that $h_{\max} \geq 1$, which implies that the stock would be driven to extinction if it were harvested everywhere at the maximum rate.

The maximization problem (eqns 9–12) is an example of a Lagrange problem in optimal control theory (Kot 2001). It can be solved by applying Pontryagin's Maximum Principle (Pontryagin *et al.* 1962), which converts the maximization problem into a two-point boundary value problem (see Appendix). The boundary value problem must be solved numerically to obtain the optimal harvest $\hat{h}(x)$ and the corresponding optimal stock size $\hat{u}(x)$. We can, however, obtain some important properties of the optimal harvest without a numerical solution.

First, because $h(x)$ appears linearly in both the population dynamic eqn 11 and in the yield (eqn 9), $\hat{h}(x)$ must be a piecewise constant function (a so-called 'bang–bang control'). More precisely, at any given location x, $\hat{h}(x)$ takes on only one of three values: 0, 1/2, or $h_{\max}$. Further, $\hat{h}(x)$ always equals $h_{\max}$ near the habitat boundaries. Fish near the boundary have a relatively high probability of dispersing outside the suitable habitat and perishing; they should be harvested as quickly as possible to maximize yield.

Second, one can show that the optimal harvest always has at least one reserve; i.e. $\hat{h}(x)$ must equal zero on some part of the habitat. If $\hat{h}(x)$ did not have a reserve, it would have to take one of the three forms: either $\hat{h}(x)$ would equal $h_{\max}$ everywhere, or it would equal 1/2 everywhere, or it would include segments with $\hat{h}(x) = 1/2$ and others with $\hat{h}(x) = h_{\max}$. However, none of these three possibilities are optimal. The first case can be eliminated because it results in extinction of the stock and zero yield. A detailed analysis (see Appendix) shows that if $\hat{h}(x) = 1/2$ somewhere, then $\hat{h}(x)$ must equal 0 somewhere else, and so the second and third cases can also be eliminated. The only remaining possibility is that $\hat{h}(x)$ includes at least one reserve. Note that reserves are not built into this model. Rather, they emerge as a result of the optimization procedure.

To determine the optimal size and placement of reserves, and to determine what harvest rate should be employed outside the reserves, one must solve a two-point boundary value problem (defined by eqns 15–23 in the Appendix). Figure 1 shows numerical solutions of this problem for $h_{\max} = 1$ and a range of habitat lengths. For $\ell < 6$ (approximately twice the critical habitat size) the maximum yield is obtained by setting $\hat{h} = h_{\max}$ near the boundaries and by placing a no-take reserve in the centre of the habitat (Fig. 1a, b). For $\ell \lesssim \ell \lesssim 7.1$, reserves alternate with areas of maximal harvesting, and the number of reserves increases as ℓ increases (Fig. 1c–e). Further numerical solutions show that the larger the maximal harvest rate $h_{\max}$, the larger the reserves should be for any given habitat size.

If $\ell \gtrsim 7.1$, it can be shown that the optimal harvest consists of three parts (Fig. 1f; also see the Appendix). The central portion is given by a *singular arc*, on which $\hat{h} = 1/2$. [The harvest rate $h = 1/2$ is also optimal for the Schaefer model, the non-spatial version of model 9–12 (Clark 1990).] On either side of the singular arc the optimal harvest rate switches infinitely often between 0 and $h_{\max}$. The distances between switches shrink rapidly near the central singular arc. They shrink so rapidly that the infinite number of switches occur in a finite length of habitat and so rapidly that only a few of the switches can be seen at the resolution of Fig. 1f. These rapidly varying portions of the solution to an optimal control problem are called *chattering arcs* (Zelikin & Borisov 1994).

DISCUSSION

Model 4–6 is the simplest spatially explicit model that captures the central processes of density-dependent growth, movement and spatial harvesting. The processes that govern actual fish populations are vastly more complicated. For example, the life cycle of many marine fish can be divided into two stages, one mobile and one relatively sedentary, but model 4 ignores stage structure and assumes that all individuals are identical. The habitats of many fish are approximately two-dimensional and spatially heterogeneous, but model 4 assumes a homogeneous, one-dimensional, patch of habitat embedded in a larger uninhabitable region. In addition, the diffusion term in model 4 captures the stochastic component of larval movement, but fish dispersal actually reflects the interaction of stochasticity, stage-specific behaviour and advection by currents.

Each of these 'real-world' complications is important and should be investigated, but many of them will not change the qualitative nature of the optimal solution. For example, advection will break the spatial symmetry of the solution, making reserves larger upstream. The best harvest policy, however, will still be a bang–bang control combining no-take reserves, maximally harvested areas and singular arcs with intermediate harvest rates.

The boundary conditions (eqn 12) are important. They reflect the assumed discontinuity between the habitat patch and its uninhabitable surroundings. If this discontinuity

©2003 Blackwell Publishing Ltd/CNRS

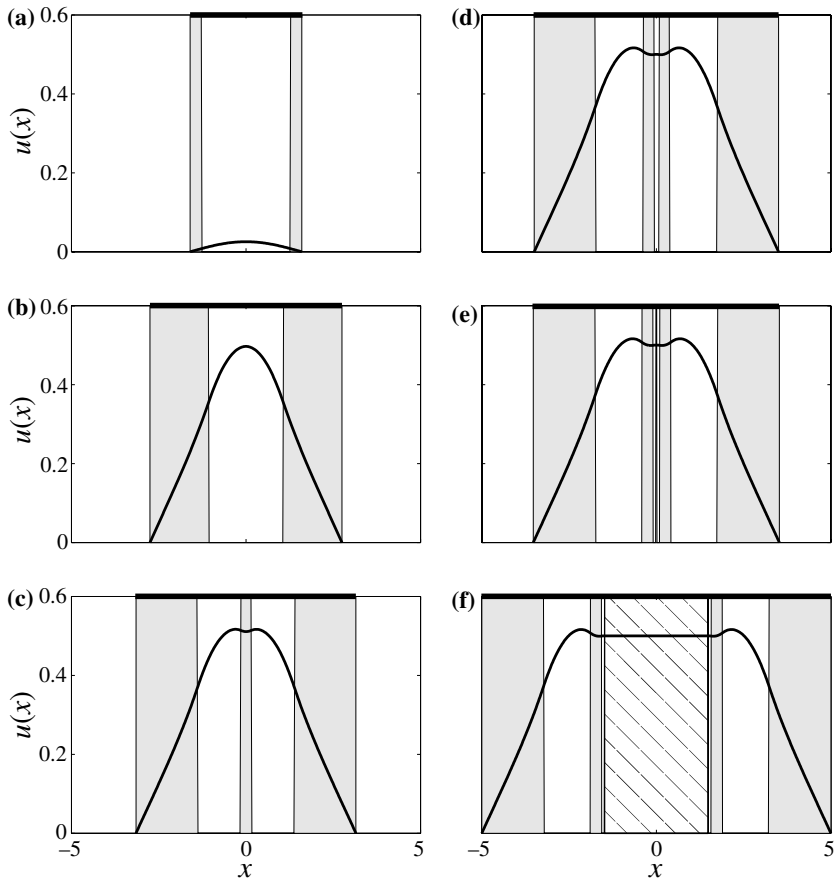

Figure 1 Solutions to the optimal control problem (eqns 9–12) for various values of the dimensionless habitat length ℓ (dark bars). Shaded areas represent different optimum harvest levels: $\hat{b} = 0$ (white), $\hat{b} = 1/2$ (hatched) and $\hat{b} = b_{\max} = 1$ (grey). The dimensionless length and the optimal number of reserves in each panel is (a) $\ell = 3.2$, 1 reserve; (b) $\ell = 5.5$, 1 reserve; (c) $\ell = 6.3$, 2 reserves; (d) $\ell = 7$, 3 reserves; (e) $\ell = 7.05$, 4 reserves; and (f) $\ell = 10$, infinitely many reserves. To facilitate comparisons, the x-axis has been shifted to place the centre of the habitat at the origin.

were eliminated by using the no-flux boundary conditions $\mathrm{d}u/\mathrm{d}x = 0$ at $x = 0$ and $x = \ell$, the optimal policy would be the spatially homogeneous policy that sets $\hat{b}(x) = 1/2$ for all x. No-flux boundary conditions are also called 'reflecting' boundary conditions because they mimic a process whereby individuals about to leave the habitat through the boundary instantaneously reverse direction. These boundary conditions are unrealistic for most marine fish because dispersing larvae are at the mercy of currents and cannot change direction of their own accord.

Other modelling studies have examined the effects of spatial heterogeneity in the form of an intrinsic source–sink structure in the environment (Crowder *et al.* 2000). A 'source' is a region in which emigration exceeds immigration and a 'sink' is a region where immigration exceeds emigration (Pulliam 1988). Sources and sinks also appear in the optimal harvest strategy here, but they are produced by the optimal harvest rather than by intrinsic environmental heterogeneity. The yield-maximizing policy requires $\hat{b} = b_{\max}$ wherever $\mathrm{d}^2\hat{u}/\mathrm{d}x^2 > 0$ (Fig. 1). In these areas the diffusion term in (eqn 1) is positive and augments the net rate of change of population density. Thus areas of maximum harvesting are sinks. In contrast, $\hat{b} = 0$ wherever $\mathrm{d}^2\hat{u}/\mathrm{d}x^2 < 0$. Thus the locations of no-take reserves are sources. Note again, however, that there is no intrinsic source–sink structure in model 9–12. In the absence of harvesting, $\mathrm{d}^2u/\mathrm{d}x^2 < 0$ at all places, and every location is a source. Just as the existence of reserves in the optimal harvest emerges as a result of the optimization procedure, so does the existence of a source–sink structure.

For large ℓ, the yield is maximized by an intricate network of marine reserves, alternating with areas of maximal harvesting and a large central area where effort controls must be employed. Networks of reserves have been advocated by many marine ecologists as being better than single reserves in that they would 'provide significantly greater protection for marine communities than a single reserve' and 'protect against catastrophes' (Lubchenco *et al.* 2003). In the present study, I have shown that such marine reserve networks may also produce better yields than single reserves.

A harvest policy with infinitely many exceedingly small reserves, like the optimal policy for large ℓ, could never actually be implemented. Spatial management policies must have much simpler geometries to be practical. Fortunately, a simple approximate solution is available. The chattering in

©2003 Blackwell Publishing Ltd/CNRS

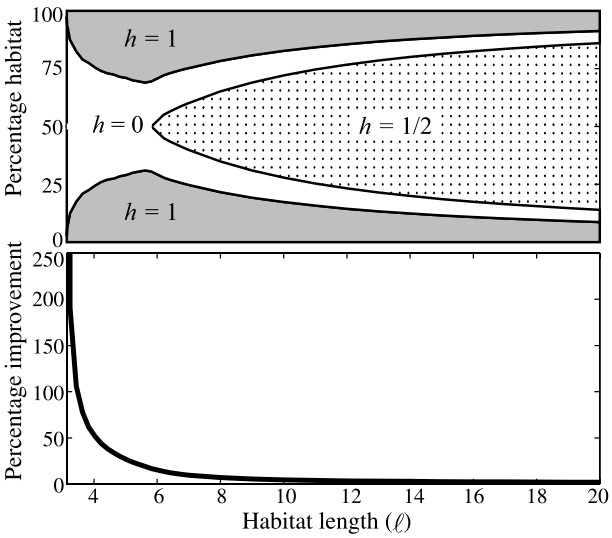

Figure 2 The best near-optimal harvesting strategy vs. dimensionless habitat length ℓ (top), and the improvement of the best near-optimal strategy over the best spatially uniform strategy (bottom).

the optimal trajectory for large ℓ produces tiny oscillations around $u = 1/2$ (Fig. 1f). This suggests an approximate solution for $\ell > 6$, with only two reserves separated by a central region with $h = 1/2$ and bordered by regions of maximal harvesting near the habitat boundaries. The top panel of Fig. 2 shows the best such approximate solution as a function of ℓ. As ℓ increases, the fraction of the habitat that should be set aside as reserve decreases and the intermediate harvest intensity should be used over a larger fraction of the habitat. The bottom panel of Fig. 2 shows the advantage of the approximate policy over the best spatially uniform policy. As ℓ shrinks to ℓ_{cr}, the advantage increases without bound. For large ℓ however, the advantage practically disappears. Thus, for relatively sedentary species, with high population growth rates, living in large habitats, i.e. for species that are thriving, spatially uniform policies are likely to be close to optimal with respect to yield and reserves will be less important.

For small values of ℓ, $\ell \lesssim 2\ell_{cr}$, the yield-maximizing harvest policy for model 4–6 is to create a large central no-take reserve and to exert maximal fishing effort outside the reserve. Further numerical studies show that the larger the maximal effort h_{max} is, the larger the reserve should be. 'Small ℓ' species have small population growth rates, large diffusion coefficients, and live in spatially restricted habitats. 'Large $h_{\max}$' species are relatively easy to catch (they have high catchability coefficients) and are subject to intense fishing pressure. These characteristics make small ℓ, large $h_{\max}$ species more vulnerable and therefore more likely to be in need of large reserves for conservation purposes. The goals of conservation and exploitation are often at odds but my analysis of this simple model suggests that, at least in this fortunate circumstance, they should employ the same means to achieve their ends.

ACKNOWLEDGEMENTS

The author thanks Claudia Augusto Martins, Hal Caswell, Michael Fogarty, Lara Gulmann, Christine Hunter, Tin Klanjscek, Petra Klepac, Mark Kot, Mercedes Pascual, Andrew Solow and John Steele for helpful discussions and suggestions. He acknowledges support from the National Science Foundation (DEB-0235692, DEB-9973212, DEB-9973518, OCE-0083976), the Environmental Protection Agency (R-82908901-0), and the Woods Hole Oceanographic Institution. WHOI contribution 10969.

REFERENCES

Clark, C.W. (1990). *Mathematical Bioeconomics: The Optimal Management of Renewable Resources*. John Wiley and Sons, New York.

Crowder, L.B., Lyman, S.J., Figueira, W.F. & Priddy, J. (2000). Source-sink population dynamics and the problem of siting marine reserves. *Bull. Mar. Sci.*, 66, 799–820.

Fisher, R.A. (1937). The wave of advance of advantageous genes. *Eugenics*, 7, 355–369.

Gaines, S.D., Gaylord, B. & Largier, J.L. (2003). Avoiding current oversights in marine reserve design. *Ecol. App.*, 13, 532–546.

Gerber, L.R., Botsford, L.W., Hastings, A., Possingham, H.P., Gaines, S.D., Palumbi, S.R. *et al.* (2003). Population models for marine reserve design: a retrospective and prospective synthesis. *Ecol. App.*, 13, S47–S64.

Halpern, B.S. & Warner, R.R. (2002). Marine reserves have rapid and lasting effects. *Ecol. Lett.*, 5, 361–366.

Kelley, H.J., Kopp, R.E. & Moyer, H.G. (1967). Singular extremals. In: *Topics in Optimization: Mathematics in Science and Engineering*, Vol. 31 (ed. Leitmann, G.). Academic Press, New York, pp. 63–101.

Kot, M. (2001). *Elements of Mathematical Ecology*. Cambridge University Press, Cambridge, UK.

Lewis, R.M. (1980). Definitions of order and junction conditions in singular optimal control problems. *SIAM J. Control Optim.*, 18, 21–32.

Lubchenco J., Palumbi, S.R., Gaines, S.D. & Andelman S. (2003). Plugging a hole in the ocean: the emerging science of marine reserves. *Ecol. Appl.*, 13, S3–S7.

National Research Council. (2001). *Marine Protected Areas: Tools for Sustaining Ocean Ecosystems*. National Academies Press, Wasington, DC.

Peterson, C.C. (2002). *Testimony before the Subcommittee on Fisheries Conservation Wildlife and Oceans Committee on Resources United States House of Representatives* [WWW document]. URL http://resourcescommittee.house.gov/107cong/fisheries/2002may23/peterson.htm.

Pontryagin, L.S., Boltyanskii, V.S., Gamkrelidze, R.V. & Mischenko, E.F. (1962) *The Mathematical Theory of Optimal Processes*. Wiley-Interscience, New York.

©2003 Blackwell Publishing Ltd/CNRS

Pulliam, H.R. (1988). Sources, sinks, and population regulation. *Am. Nat.*, 132, 652–661.

Robbins, H.M. (1967). A generalized Legendre-Clebsch condition for the singular cases of optimal control. *IBM. J. Res. Dev.*, 11, 361–372.

Schaefer, M.B. (1957). Some aspects of the dynamics of populations important to the management of commercial marine fisheries. *J. Fish. Res. Board Can.*, 14, 669–681.

Shipp, R.L. (2002). *No Take Marine Protected Areas (nMPAs) as a Fishery Management Tool A Pragmatic Perspective. Testimony before the Subcommittee on Fisheries Conservation Wildlife and Oceans Committee on Resources United States House of Representatives.* [WWW document]. URL http://resourcescommittee.house.gov/107cong/fisheries/2002may23/shipp.htm).

Zelikin, M.I. & Borisov, V. (1994). *Theory of Chatterning Control with Applications to Astronautics, Robotics, Economics, and Engineering.* Birkhauser, Boston, MA.

APPENDIX

The first step in solving the maximization problem (eqns 9–12) is to convert eqn 11 into a system of two first-order ordinary differential equations by setting $\mathrm{d}u/\mathrm{d}x = v$. The problem then becomes a standard problem in optimal control, namely to maximize the yield (eqn 9) over all non-negative harvest functions $h(x)$ subject to the new state equations

$$\frac{\mathrm{d}u}{\mathrm{d}x} = v, \tag{13}$$

$$\frac{\mathrm{d}v}{\mathrm{d}x} = hu - u(1-u), \tag{14}$$

subject to the control constraint (eqn 10) and the boundary conditions (eqn 12).

Let $\mathbf{u}(x) = [u(x), v(x)]^{\mathrm{T}}$, $\boldsymbol{\lambda}(x) = [\lambda_1(x), \lambda_2(x)]^{\mathrm{T}}$, let $[\hat{\mathbf{u}}(x), \hat{h}(x)]$ be an optimal pair, and construct the Hamiltonian

$$H = \lambda_1 v - \lambda_2 u(1-u) + hu(1+\lambda_2), \tag{15}$$

$$= H_0(\boldsymbol{\lambda}, \mathbf{u}) + hH_1(\boldsymbol{\lambda}, \mathbf{u}). \tag{16}$$

Then, by the Maximum Principle, $\hat{\mathbf{u}}(x)$, $\hat{h}(x)$, and $\boldsymbol{\lambda}(x)$ satisfy the system of ordinary differential equations,

$$\frac{\mathrm{d}\hat{u}}{\mathrm{d}x} = \frac{\partial}{\partial \lambda_1} H[\boldsymbol{\lambda}(x), \hat{\mathbf{u}}(x), \hat{h}(x)] \tag{17}$$

$$\frac{\mathrm{d}\hat{v}}{\mathrm{d}x} = \frac{\partial}{\partial \lambda_2} H[\boldsymbol{\lambda}(x), \hat{\mathbf{u}}(x), \hat{h}(x)] \tag{18}$$

$$\frac{\mathrm{d}\lambda_1}{\mathrm{d}x} = -\frac{\partial}{\partial u} H[\boldsymbol{\lambda}(x), \hat{\mathbf{u}}(x), \hat{h}(x)] \tag{19}$$

$$\frac{\mathrm{d}\lambda_2}{\mathrm{d}x} = -\frac{\partial}{\partial v} H[\boldsymbol{\lambda}(x), \hat{\mathbf{u}}(x), \hat{h}(x)], \tag{20}$$

the maximum condition

$$\max_{h \in [0, h_{\max}]} H[\boldsymbol{\lambda}(x), \hat{\mathbf{u}}(x), h(x)] = H[\boldsymbol{\lambda}(x), \hat{\mathbf{u}}(x), \hat{h}(x)], \tag{21}$$

the boundary conditions

$$u(0) = u(\ell) = 0, \tag{22}$$

and the transversality conditions

$$\lambda_2(0) = \lambda_2(\ell) = 0. \tag{23}$$

As the Hamiltonian is linear in h, the maximum condition requires

$$\hat{h} = \begin{cases} 0, & \text{if } H_1[\boldsymbol{\lambda}(x), \hat{\mathbf{u}}(x)] < 0 \\ h_{\max}, & \text{if } H_1[\boldsymbol{\lambda}(x), \hat{\mathbf{u}}(x)] > 0. \end{cases} \tag{24}$$

If

$$H_1[\boldsymbol{\lambda}(x), \hat{\mathbf{u}}(x)] = 0 \tag{25}$$

on some interval (i.e. if $\lambda_2 = -1$), then the maximum condition does not provide any information about $\hat{h}$ on that interval. In this case one can differentiate eqn 25 with respect to x to determine $\hat{h}$. As it turns out, we must differentiate four times:

$$\frac{\mathrm{d}H_1}{\mathrm{d}x} = -u\lambda_1 + v(1+\lambda_2) = 0, \tag{26}$$

$$\frac{\mathrm{d}^2H_1}{\mathrm{d}x^2} = -2v\lambda_1 - u[1 + 2\lambda_2 - u(1+3\lambda_2)] + 2hu(1+\lambda_2) = 0, \tag{27}$$

$$\frac{\mathrm{d}^3H_1}{\mathrm{d}x^3} = (4-5u)u\lambda_1 + v[-1 - 4\lambda_2 + 2u(1+5\lambda_2)] + h[-2u\lambda_1 + 2v(1+\lambda_2)] = 0, \tag{28}$$

$$\frac{\mathrm{d}^4H_1}{\mathrm{d}x^4} = u(1 - 20v\lambda_1 + 8\lambda_2) - 3u^2(1+9\lambda_2) + u^3(2+20\lambda_2) + 2v[4\lambda_1 + v(1+5\lambda_2)] + hu[-5 - 8\lambda_2 + u(7+15\lambda_2)] = 0. \tag{29}$$

As $\lambda_2 = -1$, and we are assuming $u > 0$, the first derivative (eqn 26) implies $\lambda_1 = 0$. The second derivative (eqn 27) then gives $u = 1/2$. Substituting these values into the third derivative (eqn 28) gives $v = 0$. Finally, the fourth derivative (eqn 29) reduces to

$$-\frac{1}{2}h + 1/4 = 0. \tag{30}$$

©2003 Blackwell Publishing Ltd/CNRS

Thus we set $h = 1/2$ to complete the determination of the singular arc; setting $\hat{h} = 1/2$ whenever $H_1[\boldsymbol{\lambda}(x), \hat{\mathbf{u}}(x)] = 0$ then completes the solution (eqn 24).

The number of derivatives that are required to find $\hat{h}$ determines the *order* of the singular arc (Lewis 1980). In general, it can be shown that an even number of derivatives are required (Robbins 1967). The *intrinsic order* of the singular arc is that integer q such that $(\mathrm{d}^{2q}/\mathrm{d}x^{2q})H_1$ is the lowest order total derivative of H_1 in which h appears explicitly. The *local order* of a singular arc is the least integer p such that

$$\frac{\partial}{\partial h}\frac{\mathrm{d}^{2p}}{\mathrm{d}x^{2p}}H_1\bigg|_{\lambda(x),\hat{\mathbf{u}}(x),\hat{h}(x)} \neq 0. \tag{31}$$

From eqns (27) and (30) we see that the singular arc for problem (9–12) has intrinsic order $q = 1$ and local order $p = 2$.

There are two reasons that the order of the singular arc is important. First, it has been shown that the inequality variously known as Kelley's condition (Zelikin & Borisov 1994) or the generalized Legendre–Clebsch condtion (Lewis 1980),

$$(-1)^p\frac{\partial}{\partial h}\frac{\mathrm{d}^{2p}}{\mathrm{d}x^{2p}}H_1\bigg|_{\lambda(x),\hat{\mathbf{u}}(x),\hat{h}(x)} \leq 0 \tag{32}$$

is necessary if the singular arc of local order p is to provide a maximum (Kelley *et al.* 1967). For the singular arc of problem (9–12), the derivative in (eqn 32) equals $-1/2$ (c.f. eqn 30) and the inequality is satisfied. Secondly, Zelikin & Borisov (1994) show that if a singular arc of second local order is to join a non-singular arc, and if the Hamiltonian system can be reduced to the semicanonical form

$$\frac{\mathrm{d}z_1}{\mathrm{d}x} = z_2 + f_1(\mathbf{z}, h), \tag{33}$$

$$\frac{\mathrm{d}z_2}{\mathrm{d}x} = z_3 + f_2(\mathbf{z}, h), \tag{34}$$

$$\frac{\mathrm{d}z_3}{\mathrm{d}x} = z_4 + f_3(\mathbf{z}, h), \tag{35}$$

$$\frac{\mathrm{d}z_4}{\mathrm{d}x} = a(\mathbf{z}) + hb(\mathbf{z}) \tag{36}$$

by a change of variables $(\boldsymbol{\lambda}, \mathbf{u}) \to \mathbf{z}$, and if the f_i satisfy

$$\lim_{\kappa\to+0}\frac{f_i(\kappa^4 z_1, \kappa^3 z_2, \kappa^2 z_3, \kappa z_4, h)}{\kappa^{4-i}} = 0, \tag{37}$$

then in the neighbourhood of the singular arc the optimal non-singular arc is chattering. 'Chattering' means that the optimal control exhibits an infinite number of switches over a finite length arc.

Setting

$$z_1 = 1 + \lambda_2,\ z_2 = -\lambda_1,\ z_3 = (2u - 1)\lambda_2,\ z_4 = 2v\lambda_2 \tag{38}$$

reduces the Hamiltonian system (eqns 17–20) to the semi-cannonical form (eqns 33–36) with

$$f_1 = 0, \quad f_2 = hz_1, \quad f_3 = \frac{z_2 z_3}{z_1 - 1}, \tag{39}$$

$$a = \frac{z_3^2 + 2z_2 z_4 - z_1(z_1 - 2) - 1}{2(z_1 - 1)}, \tag{40}$$

and

$$b = z_1 + z_3 - 1. \tag{41}$$

The functions f_i satisfy (eqn 37), and thus the singular arc is entered and exited via non-singular chattering arcs.

The distance between switches in h (from 0 to $h_{\max}$ or vice-versa) decays rapidly as one approaches the singular arc, and the sum of the lengths of both the non-singular arcs, call it ℓ^*, is finite. Numerical approximations suggest that $\ell^* \approx 7.1$. For habitats shorter than ℓ^* one cannot reach the singular arc in an optimal fashion.

Editor, M. Pascual
Manuscript received 1 April 2003
First decision made 12 May 2003
Manuscript accepted 2 June 2003

©2003 Blackwell Publishing Ltd/CNRS