USING MARKOV DECISION MODELS AND RELATED TECHNIQUES FOR PURPOSES OTHER THAN SIMPLE OPTIMIZATION: ANALYZING THE CONSEQUENCES OF POLICY ALTERNATIVES ON THE MANAGEMENT OF SALMON RUNS

ROY MENDELSSOHN[1]

ABSTRACT

The mathematics of Markov decision processes and related techniques are used to analyze a model relevant to salmon management. It is shown that the choice of grid can have a significant effect on the results obtained. Optimal policies that maximize total expected discounted return may be too variable. Smoothing costs are included to trade off long-run total return against the smoothness of the year-to-year fluctuations in the allowed harvest. Simpler, approximate policies that have a smoothing effect are also found. Preliminary analysis suggests the results are robust against misspecification of the parameters of the model. Concepts such as maximum sustainable yield would seem to impute a very high smoothing cost and are probably not practical for fish populations with a significant degree of randomness.

The history of most managed natural populations is one of sizable, nondeterministic variations in the dynamics of the population. This observed variation tends to have two sources: The first source is actual randomness in the system, such as that due to environmental variability, which will exist no matter how accurate our models become; and the second source is the inaccurate or incomplete specification of the transition probabilities themselves. Standard production models (Schaefer 1954; Pella and Tomlinson 1969; Fox 1970, 1971, 1975) assume deterministic dynamics, as do most recent bioeconomic analyses, as in Clark (1976) or Anderson (1977). For randomly varying populations, at best only extremely low harvests may be sustainable year to year, and it is not difficult to develop realistic scenarios where policies that are sustainable in a deterministic model would cause possible depletion in a stochastic model.

In this paper, the latest tools from stochastic optimization, particularly in the area of Markov decision problems (MDP's) are used to analyze a model relevant to salmon management. The viewpoint taken is that of the analyst, who must analyze trade offs and provide a decision maker with as few policies as possible that contain the maximum amount of information, rather than that of the decision maker, who ultimately decides if a particular concern or trade off is worthwhile. The salmon model is used as an example—the goal is to gain insight into managing randomly varying populations.

Ricker (1958) appears to be the first to examine the effects of variability on management. He used intuition and simulation to arrive at policies that are of the same general form as many of the policies to be discussed in this paper. However, Ricker presented no systematic way of developing optimal policies and made the incorrect assumption that the long-run stochastic behavior will have a mean equal to the deterministic equilibrium yield, with noise around this mean.

Reed (1974) derived qualitative properties of optimal policies if the random variable has a mean of 1, if it affects the population dynamics in a multiplicative manner, and if it has costs when the system is shut down (no harvesting) and then started up again (resumption of harvesting). Reed's results are not relevant to the model discussed in this paper, since he assumed the deterministic population model is concave, while the models examined in what follows are pseudoconcave. A more complete treatment of one dimensional stochastic growth models can be found in Mendelssohn and Sobel (in press).

Walters (1975) and Walters and Hilborn (1976, 1978) discussed a variety of topics as the concerns

[1]Southwest Fisheries Center Honolulu Laboratory, National Marine Fisheries Service, NOAA, Honolulu, HI 96812.

Manuscript accepted September 1979.
FISHERY BULLETIN: VOL. 78, NO. 1, 1980.

of this paper. Some of the techniques they discussed, particularly the filtering techniques (Walters and Hilborn 1978), are only appropriate if the model has an additive error term. While a Ricker spawner-recruit curve can be transformed to an additive model, many models do not have this feature.

I am presenting what I feel is an improved way to smooth out the fluctuations in the year-to-year harvests as compared with the method suggested in Walters (1975) and show that the Bayesian (adaptive) model discussed in Walters and Hilborn (1976) has an optimal policy with a very simple form that can be readily calculated.

Moreover, a rigorous approach is taken to define the model on a grid and the effects of the grid choice. None of the papers cited deal with this important question; new results are presented which show that the most serious effect of the grid is on the estimates of the long-run (ergodic) probabilities of the population dynamics when following a given policy. Particularly the tail properties of the ergodic distribution, i.e., the long-run probability of low harvest or low population sizes, are misestimated. This is a new finding even in the MDP literature, and has numerical implications, particularly when calculating the trade off between the mean harvest of a given policy and the long-run probability of undesirable events when following that policy.

THE MODEL

The models to be analyzed were developed by Mathews (1967) to describe the spawner-recruit relationships of sockeye salmon, *Oncorhynchus nerka*, populations in two rivers that run into Bristol Bay, Alaska. Oceanographic and other factors affect the number of recruits to a degree where the relationships can be modeled by the random equations:

Wood River:

$$x_{t+1} = \exp(d)\ (4.077y_t)\ \exp(-0.800y_t)$$
$$d \approx N(0, 0.2098) \qquad (1.1a)$$

Branch River:

$$x_{t+1} = \exp(d)\ (4.554y_t)\ \exp(-1.845y_t)$$
$$d \approx N(0, 0.3352) \qquad (1.1b)$$

where y_t is the number of spawners in period t, x_{t+1} is the (random) number of recruits in period $t+1$, and $d \approx N(a, b)$ denotes that d is a normally distributed random variable, with mean a and variance b.

For deterministic versions of Equation (1.1), the primary objective of management is MSY (maximum sustainable yield), which is equivalent to the largest per period growth of the deterministic model. The stochastic equivalent of this criterion is to maximize the average per period harvest, or gain optimality. Mathematically, letting E be the expectation operator, this is

$$\max \lim_{T \to \infty} \left[\frac{1}{T} E \sum_{t=1}^{T} (x_t - y_t) \right]. \qquad (1.2a)$$

However, for many decision making situations, total expected discounted harvest may be a preferable criterion, since a discount factor can represent a measure of risk or uncertainty about the system, over and above the variability due to the random variable d. More formally, if α is a discount factor $0 \leq \alpha < 1$, the problem is to:

$$\text{maximize } E \left[\sum_{t=1}^{\infty} \alpha^{t-1}\, p(x_t - y_t) \right] \qquad (1.2b)$$

subject to $0 \leq y_t \leq x_t$; and Equation (1.1)

where p is a weighting factor, which could be 1 or could represent the average weight of the salmon harvested.

All the results in this paper are for expected discounted return with $\alpha = 0.97$. For $\alpha = 1$, Equation (1.2a) must be used, since Equation (1.2b) is infinite for most policies. The choice of $\alpha = 0.97$ is arbitrary, though numerical runs for α ranging from 0.95 to 1.00 produced no significant changes in the results. When actually implementing a model, a careful choice of α must be made, and the sensitivity of the results to changes in the value of α should be tested. It should be mentioned that $\alpha = 1$ is just as much a discount factor as any other value and implies certain temporal preferences and attitudes towards risk that may not adequately reflect the decision maker's preferences.

The shortcomings of Equation (1.1a) or (1.1b) should also be noted, such as no account is taken of ocean harvesting of the salmon, particularly by a foreign nation. This just reinforces the idea that the purpose of this analysis is not optimization per se, but rather to provide the decision maker with added insight and reasonable first choices.

Defining the Model on a Discrete Grid

In order to make Equation (1.2) amenable to numerical methods, it is necessary to define both the state space and the action space on a discrete grid, and then to redefine the transition probabilities, etc., on this grid. Several authors (Fox 1973; Bertsekas 1976; Hinderer 1978; Waldmann 1978; Whitt 1978; Larraneta[2]) have suggested techniques to reduce MDP's to a grid and give bounds on the error due to the approximation. I have shown elsewhere (Mendelssohn[3]) that grid choice can have a significant effect on the analysis. An optimal policy and the value of an optimal policy may not be greatly affected by the choice of grid, but the estimated probabilistic behavior of the population dynamics is affected significantly by the choice of grid.

A first effort then is to find an adequate grid for the problem, a grid fine enough for both the desired accuracy and for realistic approximations of observed population sizes and coarse enough for computational efficiency. Increased computational efficiency makes it reasonable to solve many variations of a given model, which allows for a more thorough exploration of the management questions of interest and their sensitivity to key assumptions.

Several different grids were tried for Equation (1.2) for both the Branch and Wood Rivers.

To define Equation (1.2) on a given grid, suppose a grid of k points has been chosen on which to discretize the problem and assume, as is reasonable for this problem, that the reduced action space (how many spawners to leave) is equivalent to the state space (how many recruits are observed at the beginning of the period). From Equation (1.1), letting R_1 and R_2 represent the parameters of the Ricker equation

$$P(x_{t+1} \leqslant \omega | y_t) = P[(e^d)R_1 y_t \exp(-R_2 y_t) \leqslant \omega]$$

$$= P(d \leqslant \ln\omega - \ln a) \qquad (2.1)$$

where $a = [R_1 y_t \exp(-R_2 y_t)]$. Let Φ be the standard normal integral for a random variable $\bar{d} = d/\sigma$, and let x_i, x_{i+1} be any two adjacent points on the grid. Then:

$$P(d \leqslant \ln x_i - \ln a) = \Phi\left(\frac{\ln x_i - \ln a}{\sigma}\right)$$

$$P(d \leqslant \ln x_{i+1} - \ln a) = \Phi\left(\frac{\ln x_{i+1} - \ln a}{\sigma}\right)$$

so that one method of defining the transition probabilities on a grid is:

$$P(x_{t+1} = x_{i+1} | y_t) = \Phi\left(\frac{\ln x_{i+1} - \ln a}{\sigma}\right) - \Phi\left(\frac{\ln x_i - \ln a}{\sigma}\right).$$

The discrete probability when the action is y_t is equal to the total probability of going to any state in the interval (x_i, x_{i+1}).

If zero is included as a state, the procedure needs to be modified slightly. Suppose the probability of going to x_1 is known for each decision y. Then an arbitrary fraction of this probability is assigned as going to the zero state. In this paper, one-half of the probability in the interval $(0, x_1)$ is assigned to the zero state. The results have been found not to be sensitive to the value of the fraction; this is because zero is an absorbing state. Either there exists a policy that never reaches $(0, x_1)$ and hence never reaches zero, or else with probability one the population goes to zero in finite time. Hence, it is the size of $(0, x_1)$ that most influences the results, not the fraction of this total that is assigned to going to the absorbing state.

Adding an absorbing state is sensible if the absorbing state is thought of as all states at low enough population levels such that it would take years for the fishery to recover again, if it recovers at all. Without the absorbing state, the models in Equation (1.1a, b) will always recover in fairly short order. Since fisheries can be depleted, the

[2]Larraneta, J. C. 1978. Approaches to approximate Markov decision processes. Paper presented at Joint National ORSA/TIMS Meeting, Nov. 13-15, 1978, Los Ang., Calif.

[3]Mendelssohn, R. 1978. The effects of grid size and approximation techniques on the solutions of Markov decision problems. SWFC Admin. Rep. 20H, 15 p. Southwest Fish. Cent. Honolulu Lab., Natl. Mar. Fish. Serv. NOAA, Honolulu, HI 96812.

inclusion of an absorbing state would seem to be a more realistic assumption. It is included in what follows.

A coarser grid implies, in a sense, less information about the state of the system. As the interval $(0, x_1)$ becomes large, our information has decreased about the true state of the population and this increased uncertainty is reflected in increased risk of absorbtion. Similarly, a finer grid implies more exact information—a grid should not be used which is finer than the precision of the estimate of the population size.

Optimal policies for grids of 16, 26, 51, 101, and 501 equally spaced points (including zero) for both rivers are shown in Table 1. The optimal equilibrium population for the equivalent deterministic models are shown also. All numbers are in units of millions of fish.

The optimal policies are all of the base stock variety, i.e. it is optimal to harvest to a fixed number of spawners, or else not to harvest at all. If the 501-point grid is taken as the standard, it can be seen that each coarser grid has as its base stock size the grid point closest to the base stock size for the 501-point grid.

Figure 1 gives the long-run (ergodic) cumulative distribution of being in any state when following an optimal policy on grids of 16, 26, 51, and 101 points. Grid size can be seen to play a crucial part in estimating the probabilistic behavior of the population. For the Wood River, extinction with probability one is predicted on grids of 16 and 26 points, while the probability is zero on grids of 51 and 101 points, so long as zero is not the initial state. Similar but not identical results are valid for the Branch River. It should be emphasized that for $\alpha = 1$, i.e., when the objective is given by Equation (1.2a), the estimated average per period harvest of any policy depends entirely on the ergodic distribution that arises from that policy. Therefore, this variation in estimated long-run behavior due to changes in grid size is nontrivial.

Probability one of extinction occurs because for a finite state, irreducible Markov chain with an absorbing state, the absorbing state is reached in

TABLE 1.—Optimal policies for the different grid sizes.

Wood River	Grid size	Branch River
y_t = min (x_t, 0.9333)	16	y_t = min (x_t, 0.3333)
y_t = min (x_t, 0.840)	26	y_t = min (x_t, 0.4000)
y_t = min (x_t, 0.700)	51	y_t = min (x_t, 0.3000)
y_t = min (x_t, 0.770)	101	y_t = min (x_t, 0.3500)
y_t = min (x_t, 0.742)	501	y_t = min (x_t, 0.3500)
Equilibrium stock 0.735	Deterministic	Equilibrium stock 0.345

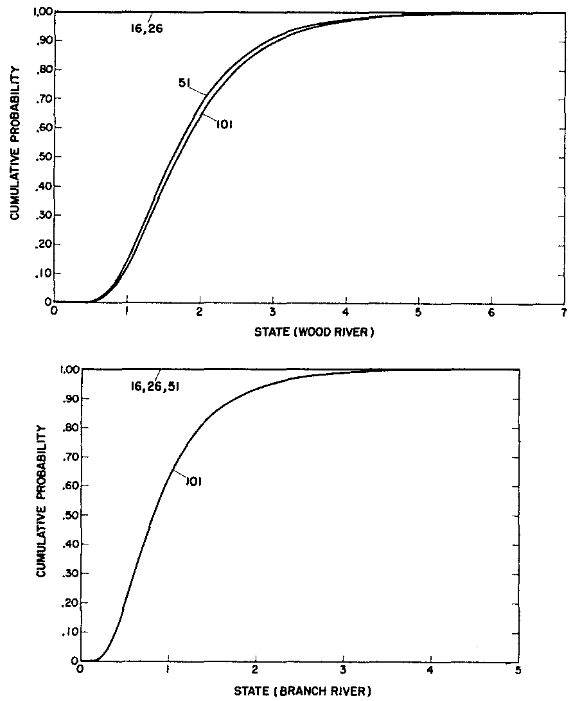

FIGURE 1.—Ergodic cumulative distributions for optimal harvesting strategies on grid sizes of 16, 26, 51, and 101 points for the Wood River and the Branch River.

finite time with probability one. However, for the larger grids, there exist policies that are reducible, in the sense that if the chain does not start in the interval $(0, x_1)$, it will never enter that interval. Since $P[x_t \epsilon (0, x_1)] = 0$, and a fraction of this probability has been assigned to the zero state, then $P(x_t = 0) = 0$. When using the smaller grids that induce Markov chains that are irreducible, the estimated time till absorption varies greatly also. For example, for the Branch River, if $P(x_1 = 0) = 0$ and $P(x_1 = \omega) = 1/(N-1)$, where ω is a grid point and N is the number of states, then a 16-point grid predicts absorption with probability one after 2,000 iterations, the 26-point grid predicts only a 76% chance of absorbtion after 2,000 iterations, and the 51-point grid predicts only a 17% chance of absorption.

When maximizing total expected discounted return, the discounted mean return depends on the values of these intermediate probability distributions, so that coarser grids can be expected to underestimate the long-run value of the harvest.

Finally, for the Wood River, note that the 51- and 101-point grids have similar long-run behavior. These results suggest that in order to find good policies, it is only necessary to use a grid size of 26 to 51 points for the problems under consideration. However, to analyze the long-run (probabilistic) behavior of a given policy, it is necessary to use a grid containing no fewer than 100 points.

It should be reemphasized that the reason for considering a coarser grid is that a smaller problem size allows for many problems to be solved at a small cost. This is desirable to obtain insight into the sensitivity of the problem. However, it is possible to solve quite large problems, making use of a variety of methods to accelerate computations (see for example Porteus 1971; Hastings and van Nunen 1977). For example, the 501-point grid for the Branch River used 1.80 s of CPU (central processing unit) time to perform the optimization. Computations, when smoothing costs are included (see Policy Analysis section), have 2,601 states. These used about 5 to 6 min of CPU time to perform the computations, but at a cost of about \$20. Our experience is that it is possible to obtain reasonable estimates using coarse grids and that this suffices for initial policy investigation. However, it is worthwhile to reanalyze the final two or three problems of greatest interest on a finer grid.

POLICY ANALYSIS

For the Wood River, the optimal policy for Equation (1.2) is given by

$$y_t = \text{minimum}\ (0.770, x_t)$$

and it produces a mean per period harvest of 1.14758, and a standard deviation in the harvest of 0.8963. The median harvest is 0.91, and no harvest occurs roughly 4.3% of the time. A harvest of 25% or less of the mean harvest occurs roughly 15% of the time, while a harvest greater than the mean harvest occurs approximately 38% of the time.

Similarly, for the Branch River, an optimal policy for Equation (1.2) is given by

$$y_t = \text{minimum}\ (0.300, x_t)$$

and it produces a mean per period harvest of 0.6622, and a standard deviation in the harvest of 0.6120. The median harvest is roughly 0.500; there is a 3.9% chance of no harvest. A harvest of 25% of the mean harvest or less occurs roughly 14.5% of the time, and a harvest greater than the mean harvest occurs approximately 61% of the time.

While these policies are similar in form to policies that are optimal for a deterministic version of Equation (1.2), they differ greatly in the year-to-year dynamics. There are two ways of finding the optimal deterministic policy. The first way is to assume a general model of the form:

$$x_{t+1} = R_1 y_t \exp\ (-R_2 y_t)$$

The second method is to assume a general model of the form:

$$x_{t+1} = E \exp\ (d)\, R_1\, y_t \exp(-R_2\, y_t)$$

where as before, R_1 and R_2 are the parameters of the Ricker equation. The second method is preferable since it uses all the information available. As d is a normal random variable with mean zero and variance σ^2, it is easy to show that exp(d) is a lognormal random variable with expectation exp ($\frac{1}{2}\, \sigma^2$). Solving for the optimum sustained yield (OSY) population size for each river gives:

	Wood River	*Branch River*
x_{OSY}	0.735	0.345
OSY	1.11346	0.63804

Both OSY values are lower than the mean per period harvests in the stochastic models, but the variation is too high to allow this amount to be harvested each year. However, the x_{OSY} level is a good estimate of the base stock size, and it is known a priori from Mendelssohn and Sobel (in press) that a base stock policy is optimal.

In the deterministic model, once x_{OSY} is reached, both the population size and the harvest size are maintained at steady, equilibrium levels. An optimal policy for the stochastic model, however, produces large fluctuations in both and may allow no harvesting 1 yr out of 25 in the long run. For many fisheries, these "boom and bust" conditions may not be acceptable. Many people, especially those with interest or mortgage payments, as are many fishermen, are concerned about smoothness of income received as well as the total amount received. The final decision on the acceptable amount of fluctuation is, of course, up to the decision maker with appropriate input.

There are several methods available to try to find a balance between the smoothness of the random income stream and its total discounted expected value. Walters (1975) and Walters and Hilborn (1978) suggested fixing a given mean harvest

u, and then finding a policy that minimizes $\lim_{T\to\infty} E \frac{1}{T} \sum_{t=1}^{T} (z_t - u)^2$. This methodology depends on the values of u chosen. It also determines the policy that minimizes the approximate long-run variance for a given long-run mean harvest. This is not equivalent to reducing the size of the year-to-year fluctuations.

A second method is to include "smoothing costs" into the one-period return. This approach has been studied analytically by Mendelssohn.[4] Let γ be the cost of a unit decrease in the harvest from year to year, and let ϵ be the cost of a unit increase in the harvest from year to year.

If z was harvested last year, then net revenues this year, for any harvest z_t, are decreased by

$$\begin{cases} \gamma(z - z_t) & \text{if } z > z_t \\ 0 & z = z_t \quad . \\ \epsilon(z_t - z) & z < z_t \end{cases}$$

Amended to Equation (1.2), this would imply a one-period net benefit of

$$p(x_t - y_t) - \gamma[z - (x_t - y_t)]^+ - \epsilon[(x_t - y_t) - z]^+$$

where $(a)^+$ denotes the positive part of a. An alternate form is to let $e = (\gamma - \epsilon)/2$ and $c = (\gamma + \epsilon)/2$. Then the one-period return is:

$$p(x_t - y_t) + e(x_t - y_t) - c\,|\,(x_t - y_t) - z\,|\, - ez.$$

One advantage to the smoothing cost approach over other approaches is that p, e, and c can be normalized so as to be interpreted as relative prices. That is, the normalized values $p = 1$, e/p, and c/p can be interpreted as the value of having the between period harvest "smoothed" by one unit relative to the value of one unit of additional harvest. Actual relative values are often difficult to determine. But by parameterizing on e and c, it is possible to present a decision maker not only a range of possible "optimal" policies and their consequences, but also some feeling for the relative trade off between total income and the smoothness of the received income stream.

For the Wood and Branch Rivers, two sets of computations were performed. The first set assumes that $\gamma = \epsilon$, i.e., there is an equal concern for increases in allowable harvest as well as for decreases. This is equivalent to $e = 0.0$ and $c = \gamma$ (or equivalently ϵ). The motivation for this cost structure is that fishermen typically resist any decrease in the allowed harvest, hence $\gamma > 0$. However, allowing increases in the harvest size often signals fishermen to gear up and invest in equipment, thereby making it even more difficult to decrease the allowable harvest later on. Therefore this cost should be equal to a cost due to a decrease in the harvest.

As a counterbalance to this, a second set of computations were performed with $\gamma > 0$ but $\epsilon = 0$, i.e., a cost only if the harvest is decreased. This is equivalent to $c = e = \gamma/2$.

For the first set of computations, with $e = 0.0$ and $p = 1.0$, values of c of 0.25, 0.50, 0.75, 1.00, 1.25, 1.50, 1.75, and 2.00 were used. These are equivalent to relative values of ⅛, ¼, ⅜, ½, ⅝, ¾, ⅞, and 1. For the second set of runs, with $c = e$, and $p = 1.0$, values of 0.25, 0.50, 0.75, 1.00 and 1.25 were used. These are equivalent to a ratio of γ/p equal to ¼, ½, ¾, 1, 1¼. The results are summarized in Figure 2(a)-(m) and Figure 3(a)-(m), which show an optimal policy for each river for each of these cases. All computations were performed on 26-point grids.

The figures are read as follows. Suppose z was harvested last year and x is the observed population size this period. Find the point (x, z) on the graph and follow the arrow in that zone to the appropriate boundary as indicated. Then read off the z value of this point; this is the optimal amount to harvest during this period.

For example, if $c = 0.50$, $e = 0.00$, $x_t = 0.84$, and the harvest last period was 0.28, Figure 2(b) shows that the optimal policy for the Wood River is to harvest 0.28 this period. Note that the dashed line is the equivalent base stock harvest with no smoothing costs.

While the policies in Figures 2 and 3 are optimal for the given relative values of p, e, and c, they are complex in nature and would be difficult for a layperson to understand. Practical management often implies determining simpler, good but suboptimal policies that achieve the same objectives. These policies are often more desirable since they

[4]Mendelssohn, R. 1976. Harvesting with smoothing costs. SWFC Admin. Rep. 9H, 26 p. Southwest Fish. Cent. Honolulu Lab., Natl. Mar. Fish. Serv., NOAA, Honolulu, HI 96812.

are easier to implement and easier to explain the rationale to the public.

As an example of suboptimal, approximate policies, the following nine modified base stock policies were examined:

Wood River

1) Base stock policy, base stock size = 0.84.
2) Policy of base stock size of 0.56 till 2.52, then a base stock size of 0.84.
3) Policy of base stock size of 0.56 till 1.40, then a base stock size of 0.84.
4) Harvest 0 till 0.28, harvest 0.28 till 0.84, a base stock size of 0.56 till 2.52, then a base stock size of 0.84.

Branch River

5) Base stock policy, base stock size of 0.40.
6) Base stock size of 0.4 till 1.6, then a base stock size of 0.6.

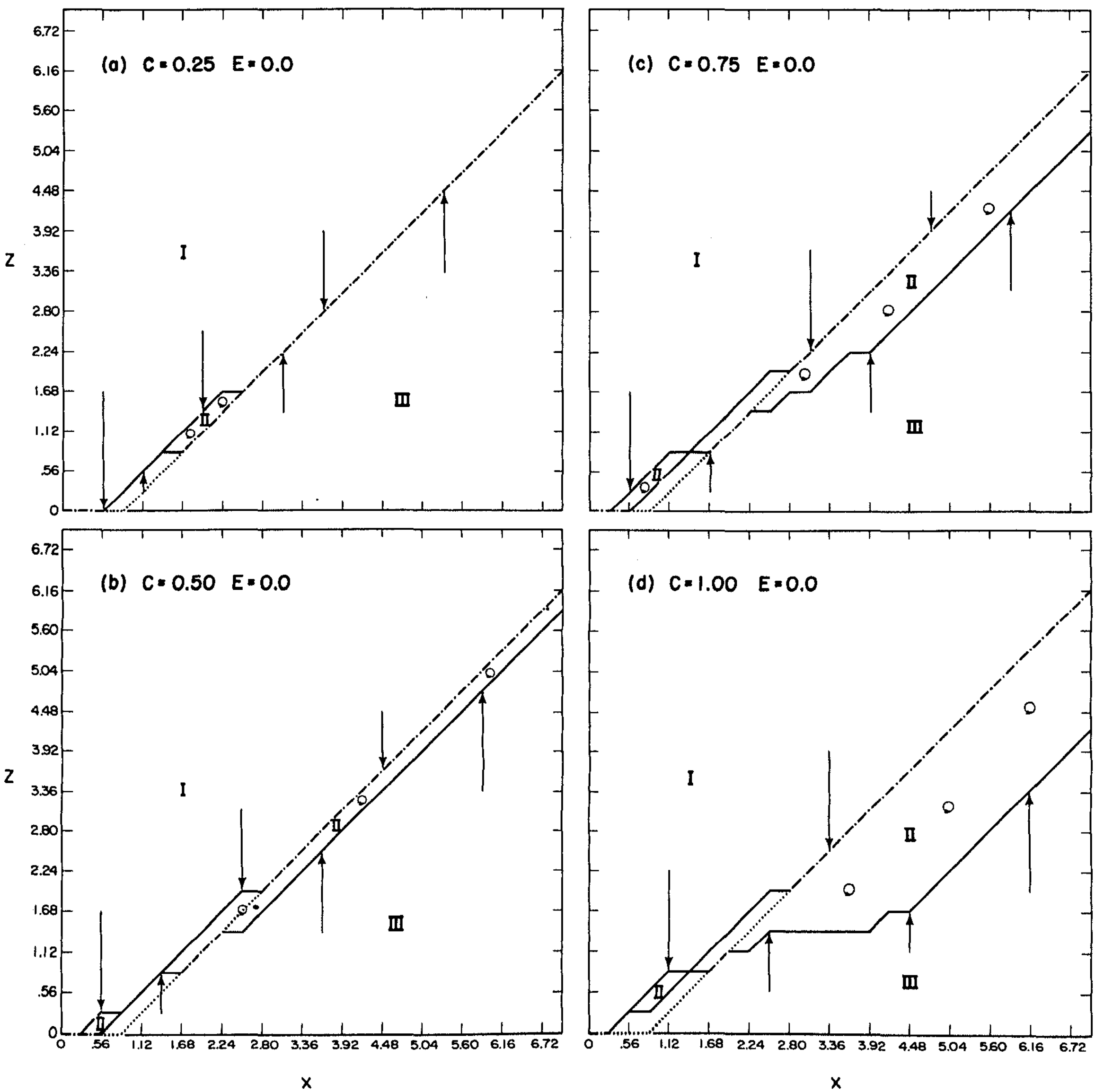

FIGURE 2(a-m).—Optimal policy functions for the Wood River for various assumptions about the relative value of smoothing costs. (See text for details).

7) Base stock size of 0.2 till 0.6, then a base stock size of 0.4.
8) Base stock size of 0.2 till 1.0, then a base stock size of 0.4.
9) Base stock size of 0.2 till 0.4, base stock size of 0.4 till 1.2, base stock size of 0.6 after that.

These nine approximate policies were devised by examining the functions that define the three regions in Figures 2 and 3. These approximate the boundaries of the three regions where the smoothing costs are one-fourth to one-half the per unit value of the harvest. The mean per period harvest, variance, standard deviation, median per period harvest, etc. for these nine policies are given in Table 2.

Policies 3 and 4 for the Wood River and 8 and 9 for the Branch River demonstrate how these approximate policies tend toward smoothing policies. For example, policy 4 has the same median harvest as the optimal base stock harvest, almost never closes the fishery, significantly de-

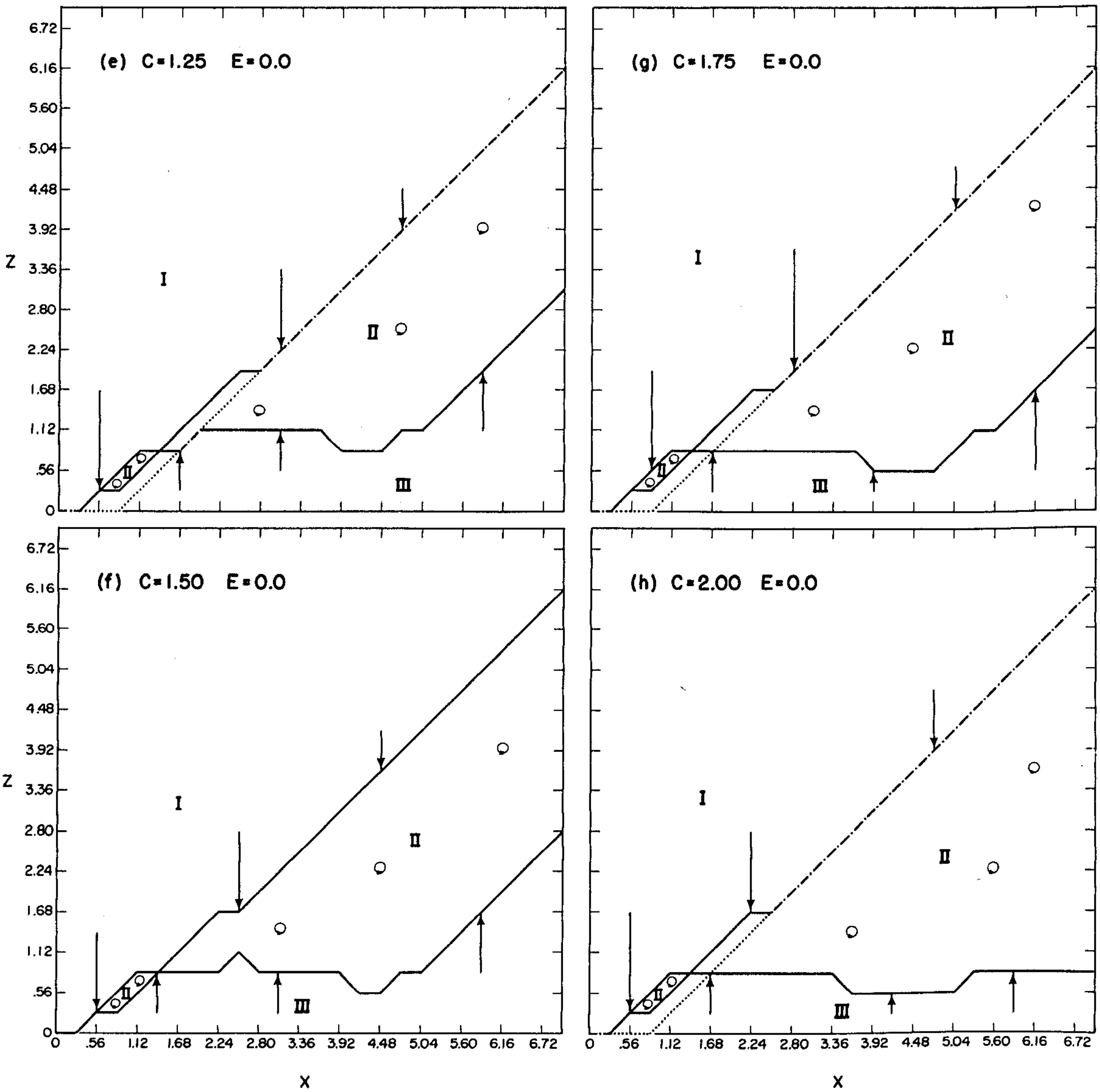

FIGURE 2.—Continued.

creases the percent of time there are low catches, and only reduces the mean per period harvest by 33,800 fish. In order to achieve a smoother catch, "potlatch" harvests from time to time have been sacrificed.

When looked at closely, these policies are actually very intuitive and represent an interesting variant of a base stock policy. These policies replace a single base stock size by a dual base stock size policy. The first base stock size is lower than the original one, while the second base stock size is greater than or equal to the original base stock size. This means that there are fewer states where there is no harvesting, but also lowers the likelihood of the really big harvests. The mean per period harvest tends to be very sensitive to these big harvests, while the median is not, particularly since the very large harvests are not too frequent.

It is curious that the population dynamics are so sensitive to such fine tuning, for the difference between policy 1 and policy 3, say, is quite marginal. It would be an interesting area of future

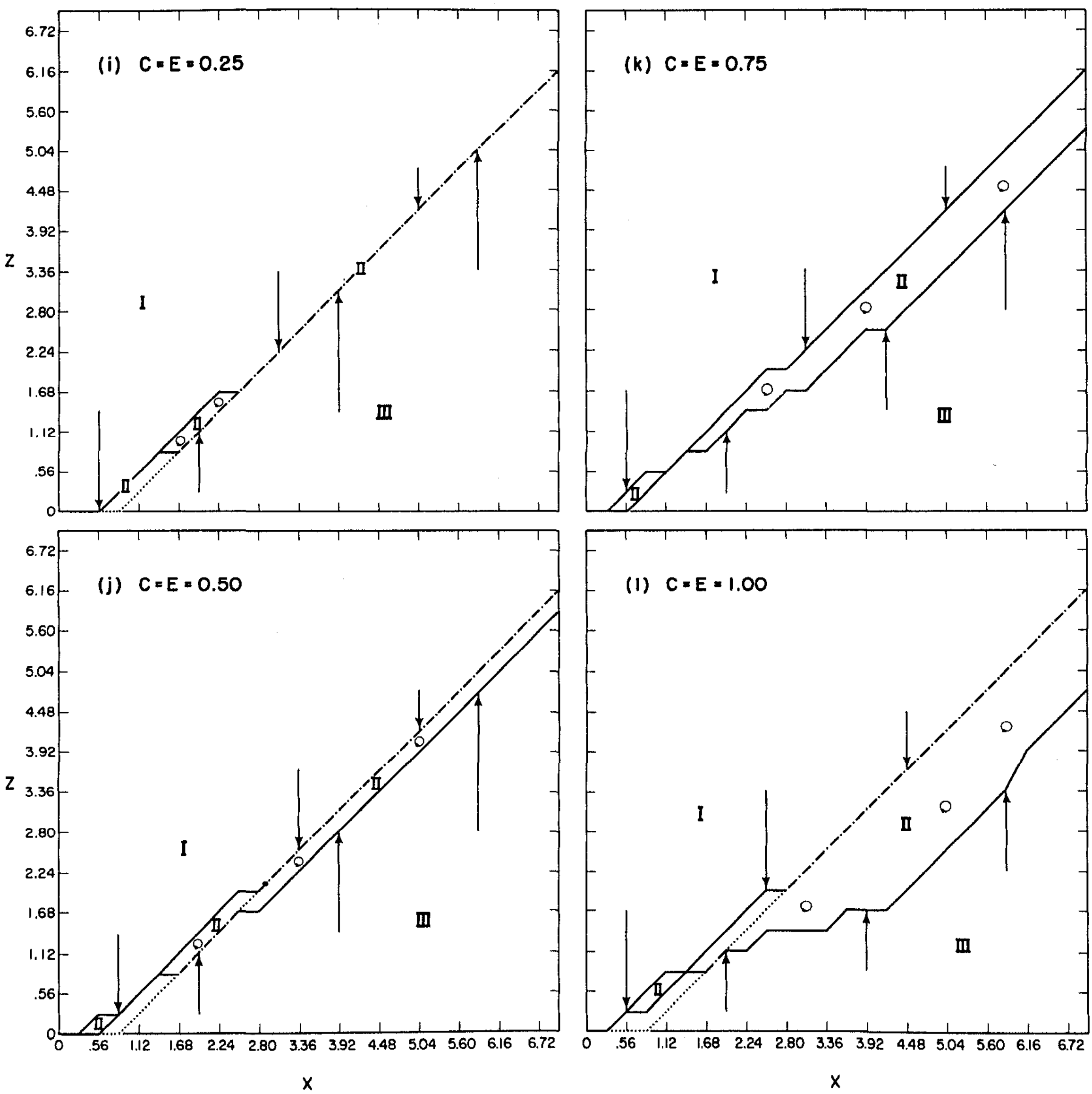

FIGURE 2.—Continued.

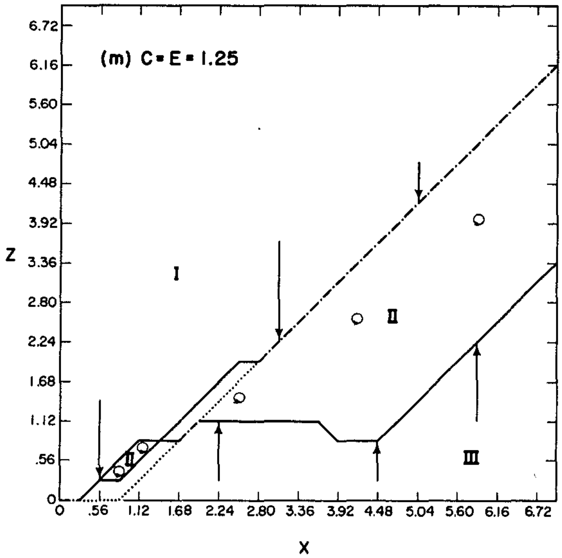

FIGURE 2.—Continued.

research to determine guidelines for when fine tuning would be expected to produce such "trimming" of the tails of the ergodic (long-run probability) distribution.

Including smoothing costs also tells us a great deal about traditional concepts of fisheries management, such as MSY. It is clear from Figures 2 and 3 that anything close to an MSY policy is optimal only if the smoothing costs exceed the per unit value of the harvest. As whole systems of laws for regulating fisheries have been constructed around the idea of smooth, constant harvests, it is clear that this imputes lower average catches, and a significant preference for constancy of the harvest over total amount harvested.

The analysis has assumed that Equation (1.1) or similar equations are available, and that the parameter estimates are accurate (in this case, estimates of R_1, R_2, and σ^2). In the latter case, management measures would seem more reasonable if they were known to be robust against misspecifying the parameters. This involves knowing how an optimal policy and total expected value would vary if the true underlying parameter values differ from those specified, and also how the estimate of the long-run probability distribution differs from the true one.

Walters and Hilborn (1976) have examined a similar question of trying to solve the Bayes model of this problem, i.e., where there is an original prior probability given to each value of the parameter, and this probability is updated each period using Bayes theorem and the observed values during the period. However, they could not obtain a solution, and Walters and Hilborn (1978) raised questions as to the validity of some of their numerical approximations.

Fortunately, qualitative results are possible for this particular class of Bayes problems. Let Θ be the parameter (or vector of parameters) under consideration. Let $q_0(\Theta)$ be the initial prior distribution on Θ, and let $q_n(\Theta)$ be the updated prior distribution after n period has elapsed. Let Ω be the set of all possible prior distributions. Then it is proven in van Hee (1977a) that if the state of the system is expanded to (x_t, q_t), the resulting optimization problem is Markovian. Following arguments similar to those in Scarf (1959) and van Hee (1977a) it follows that an optimal Bayes policy takes the form:

For each element $q \in \Omega$, there is an $x(q)$ such that:

$$\text{do not harvest if } x_t \leq x(q)$$

$$\text{harvest } x_t - x(q) \text{ if } x_t > x(q).$$

For example, if σ^2 in the distribution of d is itself a random variable, then each possible probability distribution of σ^2 yields a possibly unique base stock size policy.

TABLE 2.—Vital statistics for the nine policies approximating the smoothing cost policies for Wood and Branch Rivers.

River	Policy	Mean per period harvest	Variance of per period harvest	Standard deviation	% time no catch	% time less than 25% of mean	% time greater than mean catch	Median catch	Relative value: smoothing/price
Wood	1	1.1357	0.8468	0.9202	5.6	16.8	39	0.98	0/1
	2	1.0993	0.5460	0.7389	1.7	10.7	39.8	0.98	1/8
	3	1.1203	0.6506	0.8066	1.1	7.7	43.2	0.91	1/4
	4	1.1019	0.5758	0.7588	0.02	10.47	40	0.98	1/2
Branch	5	0.6528	0.3982	0.6310	9.2	21.8	40	0.500	0/1
	6	0.6290 ·	02532	0.5032	9.1	21.5	37.2	0.500	1/4
	7	0.6272	0.3077	0.5547	1.2	27.7	31.3	0.400	1/2
	8	0.5920	0.2202	0.4693	1.9	35.7	26.3	0.500	3/8
	9	0.5995	0.3038	0.5512	0.72	22.83	39.3	0.500	3/4

Van Hee (1977a) defined a set of policies that he terms Bayes equivalent policies. For problems such as the salmon models under discussion, a Bayes equivalent policy would be found as follows:

1) At the start of the period, the prior probability distribution is $q(\Theta)$.

2) The expected transition function (expectation with respect to Θ) is calculated, i.e.,

$$p(d, q) = \int p(d \mid \Theta) q(d\Theta) \tag{4.1}$$

where $p(\cdot \mid \cdot)$ describes the dependence of the random variable d on Θ.

3) $p(d, q)$ is used to solve a non-Bayesian Markov decision process, with $p(d, q)$ as the transition function.

4)The optimal policy from step 3 above is used for one period.

5) $q(\Theta)$ is updated using Bayes theorem and the observations from the last period, and the updated $q(\cdot)$ is used in step 1 at the next time period.

It is worth noting that a Bayes eqivalent policy

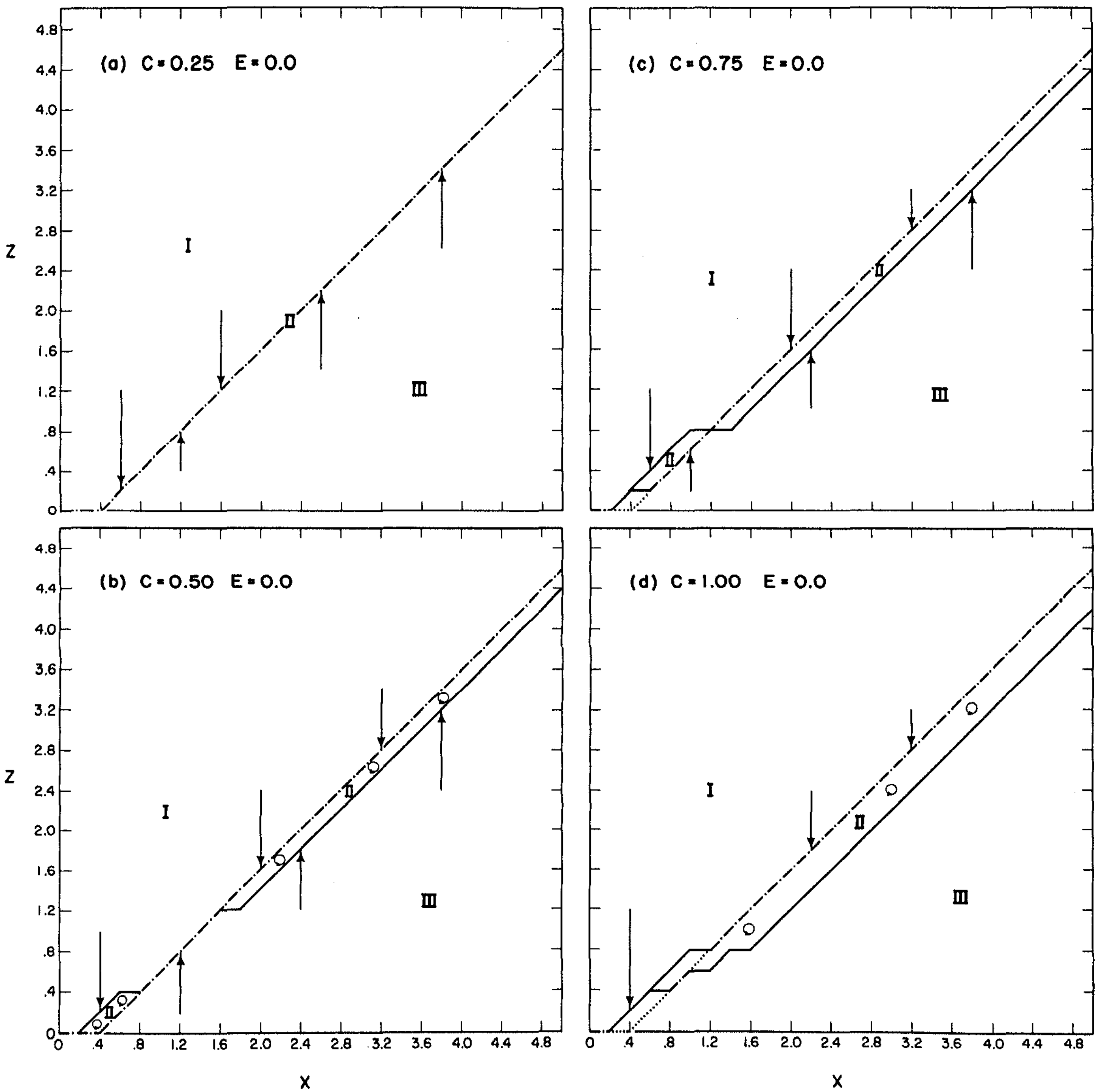

FIGURE 3(a-m).—Optimal policy functions for the Branch River for various assumptions about the relative value of smoothing costs. (See text for details.)

is adaptive, as the prior distribution is updated each period. Moreover, it is not the same as fixing Θ at its estimated value, and using a fixed value of Θ in step 3. The difference can be seen in the integral in Equation (4.1). The reason for considering Bayes equivalent policies is that van Hee (1977a, theorem 3.1) proved that for the models under discussion, when the objective is given by Equation (1.2a) or (1.2b), then the Bayes equivalent policy is optimal for the full Bayes model. For example, in Walters and Hilborn (1976), the parameter Θ is a scalar, i.e., R_2 in our notation. Their problem, for which an optimal policy was not found, can be solved by following a policy outlined in the five steps above.

Many models will not have the necessary structure for a Bayes equivalent policy to be optimal for the full Bayes model, and unlike salmon management, estimates of the population size may not be available every year. A legitimate question is: suppose the present best estimate of Θ were to be used from hereafter. What would be the loss in

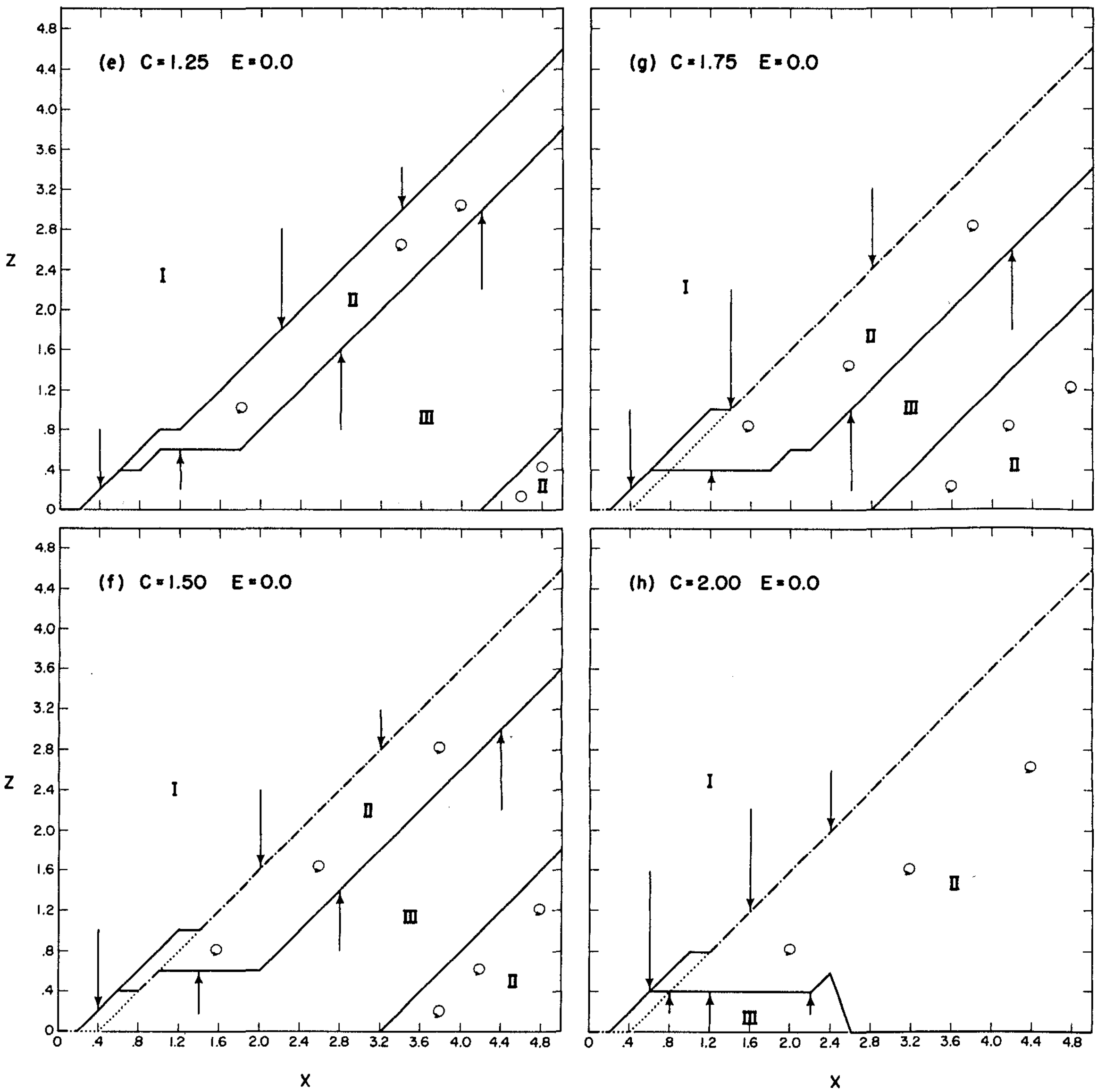

FIGURE 3.—Continued.

expected value? Van Hee (1977b) gave bounds on this expected loss that are easy to compute. To obtain a feel for these bounds, both σ^2 and R_2 are assumed to be random variables. For the Wood River, R_2 could take on the values -0.6, -0.8 and -1.0, and for the Branch River R_2 could take on the values -1.5, -1.85, and -2.00. For the Wood River, σ^2 could assume the values of 0.35, 0.45, and 0.55, and for the Branch River σ^2 could assume the values 0.48, 0.58, and 0.68. Three probability distributions were used as the present prior probability of the parameter values. These were (⅓, ⅓, ⅓,), (¼, ½, ¼), (⅛, ¾, ⅛). The results of the optimization using the parameters at each fixed value (which are needed to calculate the bounds) are given in Table 3. Table 4 gives the bounds on the expected loss of value from using the present estimates of the parameters as in Equation (1.1).

Table 3 suggests that as σ^2 varies for fixed values of R_1, R_2, the mean per period harvest varies little, but the variance of the long-term harvest size distribution increases significantly. As R_2

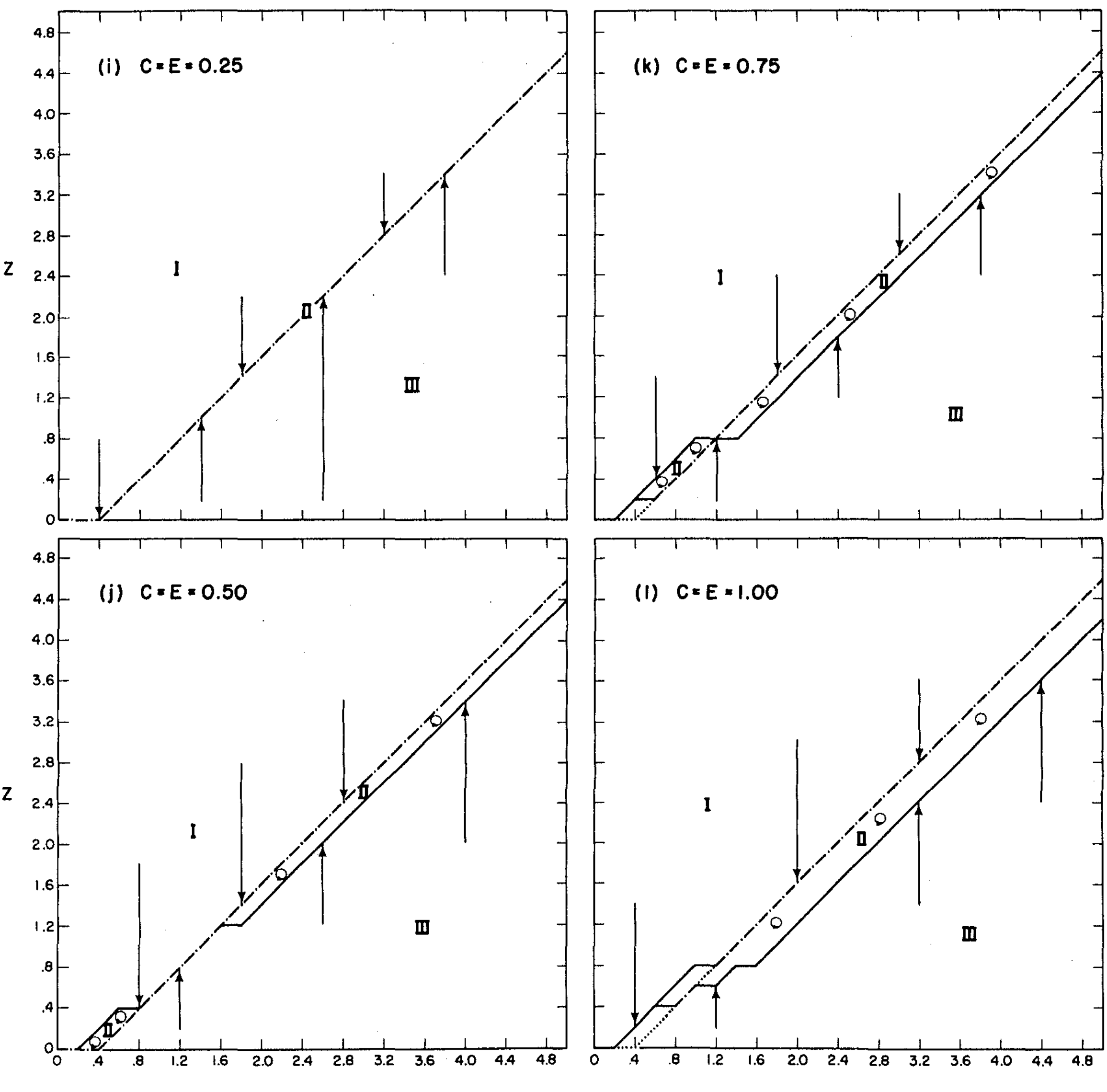

FIGURE 3.—Continued.

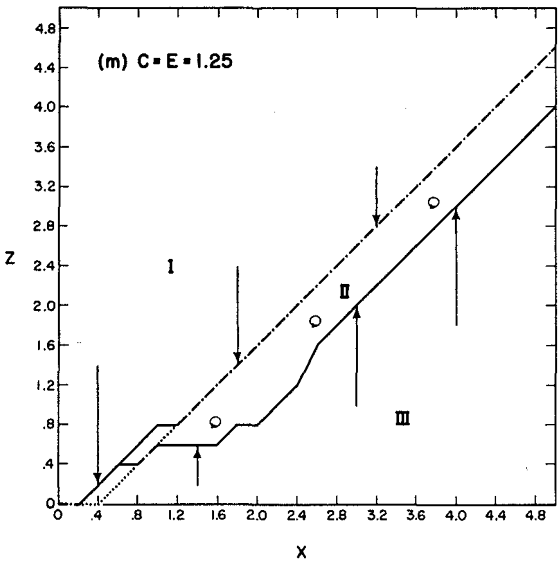

FIGURE 3.—Continued.

varies for fixed values of R_1, σ^2, both the mean and the variance vary significantly. Table 4 reinforces this impression to a degree. If the mean per period harvest does not vary significantly with changes in the value of σ^2, it might be expected that the present estimate of σ^2 will suffice. This is borne out by Table 4, where the bounds on the maximum expected total loss is <0.01, which is <1% of the optimal Bayes expected value.

Some significant expected loss in value when R_2 varies is seen, but the loss is less than might be expected from Table 3. The values in Table 4 when R_2 varies are all <4% of the true value. These results suggest that if Equation (1.1) is the correct form of the model, and the present parameter estimates have relatively small variance, then little is gained in expected value if the more complicated policy is used. The same may not be true if the population size is unobserved.

All of these results suggest a model that is fairly robust to our lack of understanding of nature. A possible explanation for this can be made from the discussion on the effect of grid size. As long as there is some cutoff population size below which no harvesting is allowed, and this cutoff assures that the absorbing state cannot be reached with probability one, then our management can only damage the stocks to a degree.

All of the policies examined in this paper have such a minimum cutoff. The rest of the policy will determine the relative mean and variance of the harvest, and techniques are presented to examine these features in detail. Uncertainty about the values of the parameters will affect the total return, but present estimates often can give a satisfactory approximation. The truly risk adverse decision maker can use present estimates of the parameters that are weighted to be on the cautious side.

SUMMARY

Uncertainty in fisheries management can be faced head on. Techniques exist that allow us to gain much insight on managing randomly varying populations. Optimization procedures allow us to reduce our attention to the few best policies, and to analyze their properties, rather than to pick policies ad hoc that meet no special criteria.

Optimization under uncertainty can also lead to a reconsideration of what is valued in managing a

TABLE 3.—Trials with varied parameters.

River	Value of R_2	Value of σ	Optimal policy	Mean per period harvest	Variance	% time no harvest
Wood	-0.800	0.35	min (x_t, 0.7)	1.0680	0.390976	0.79
Wood	-0.800	0.55	min (x_t, 0.77)	1.2267	1.2422	7.8
Wood	-0.600	0.458	min (x_t, 0.980)	1.5108	1.3136	3.8
Wood	-1.000	0.458	min (x_t, 0.560)	0.9225	0.4839	3.29
Branch	-1.845	0.48	min (x_t, 0.35)	0.6122	0.2253	3.54
Branch	-1.845	0.68	min (x_t, 0.35)	—	—	—
Branch	-1.500	0.579	min (x_t, 0.40)	1.989	0.5254	5.82
Branch	-2.000	0.579	min (x_t, 0.30)	0.9075	0.3068	5.82

TABLE 4.—Largest possible deviation in value of the approximate policy compared with the true Bayes policy.

Probability distribution	When R_2 is uncertain			When σ is uncertain		
	1/3, 1/3, 1/3	1/4, 1/2, 1/4	1/8, 3/4, 1/8	1/3, 1/3, 1/3	1/4, 1/2, 1/4	1/8, 3/4, 1/8
Wood River	1.4	0.51	0.5	0.04	0.03	0.03
Branch River	1.04	0.47	0.38	0.01	≤0.01	≤0.01

fishery—in the examples considered, some consistency in the amount harvested is a desirable alternative to high year-to-year fluctuations in the harvest size. But this reduced the average per period catch. Only in extreme situations, where the cost of smoothing out the catch is greater than the unit value of the catch, does any policy resembling MSY become optimal.

Finally, it is possible to obtain an understanding of how robust the management measures are to misspecifications of the underlying model. This is important, since the model is only a guide to our decision making, not the answer. In the models considered, the "best" policies are robust in view of this uncertainty.

A question not examined is the assumption that the population size is observed at the start of each period. This too is usually costly, and inexact. Recently, I and E. J. Sondik developed an efficient algorithm that addresses the relative merits of different sampling intervals for obtaining population estimates.[5] Together, all of these techniques allow for an integrated, realistic approach to management under uncertainty.

ACKNOWLEDGMENTS

Debra Chow provided invaluable assistance in programming the computer runs. David Stoutemeyer of the University of Hawaii gave much useful advice on maximizing the efficiency of the optimization algorithm used. Lee Anderson, George Fishman, and Adi Ben-Israel gave important comments for improving an earlier version of this paper. The paper also benefited greatly from the comments of one of the referees and from C. Walters who tempered some remarks in a previous version.

LITERATURE CITED

ANDERSON, L. G.
1977. The economics of fisheries management. The Johns Hopkins Univ. Press, Baltimore, 214 p.

BERTSEKAS, D. P.
1976. Dynamic programming and stochastic control. Acad. Press, N.Y., 396 p.

CLARK, C. W.
1976. Mathematical bioeconomics: The optimal management of renewable resources. John Wiley and Sons, N.Y., 352 p.

FOX, B. L.
1973. Discretizing dynamic programs. J. Optim. Theory Appl. 11:228-234.

FOX, W. W., JR.
1970. An exponential surplus-yield model for optimizing exploited fish populations. Trans. Am. Fish. Soc. 99:80-88.
1971. Random variability and parameter estimation for the generalized production model. Fish. Bull., U.S. 69:569-580.
1975. Fitting the generalized stock production model by least-squares and equilibrium approximation. Fish. Bull., U.S. 73:23-37.

HASTINGS, N. A. J., AND J. A. E. E. VAN NUNEN.
1977. The action elimination algorithm for Markov decision processes. *In* H. C. Tijms and J. Wessels (editors), Markov decision theory, p. 161-170. Proceedings of the Advanced Seminar on Markov Decision Theory held at Amsterdam, the Netherlands, September 13-17, 1976. Mathematical Centre Tract 93. Mathematisch Centrum, Amsterdam.

HINDERER, K.
1978. On approximate solution of finite-stage dynamic program. Universität Karlsruhe, Fukultät für Mathematik, Bericht 8, 42 p.

MATHEWS, S. B.
1967. The economic consequences of forecasting sockeye salmon (*Oncorhynchus nerka*, Walbaum), runs to Bristol Bay, Alaska: A computer simulation study of the potential benefits to a salmon canning industry from accurate forecasts of the runs. Ph.D. Thesis, Univ. Washington, Seattle, 225 p.

MENDELSSOHN, R., AND M. J. SOBEL.
In press. Capital accumulation and the optimization of renewable resource models. J. Econ. Theory.

PELLA, J. J., AND P. K. TOMLINSON.
1969. A generalized stock production model. [In Engl. and Span.] Inter-Am. Trop. Tuna Comm., Bull. 13:419-496.

PORTEUS, E. L.
1971. Some bounds for discounted sequential decision processes. Manage. Sci. 18:7-11.

REED, W. J.
1974. A stochastic model for the economic management of a renewable animal resource. Math. Biosci. 22:313-337.

RICKER, W. E.
1958. Maximum sustained yields from fluctuating environments and mixed stocks. J. Fish. Res. Board Can. 15:991-1006.

SCARF, H.
1959. Bayes solutions of the statistical inventory problem. Ann. Math. Stat. 30:490-508.

SCHAEFER, M. B.
1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Inter-Am. Trop. Tuna Comm., Bull. 1:25-56.

VAN HEE, K. M.
1977a. Adaptive control of specially structured Markov chains. *In* M. Schäl (editor), Dynamische optimierung, p. 99-116. Bonner Mathematische Schriften 98, Bonn.
1977b. Approximations in Bayesian controlled Markov chains. *In* H. C. Tijms and J. Wessels (editors), Markov

[5]Mendelssohn, R., and E. J. Sondik. 1979. The cost of information seeking in the optimal management of random renewable resources. SWFC Admin. Rep. H-79-12, 15 p. Southwest Fish. Cent. Honolulu Lab., Natl. Mar. Fish. Serv., NOAA, Honolulu, HI 96812.

decision theory, p. 171-182. Proceedings of the Advanced Seminar on Markov Decision Theory held at Amsterdam, the Netherlands, September 13-17, 1976. Mathematical Centre Tract 93. Mathematisch Centrum, Amsterdam.

WALDMANN, K.-H.
1978. On approximation of dynamic programs. Preprint No. 439, Fachberech Mathematik, Technische Hochschule Darmstadt, 17 p.

WALTERS, C. J.
1975. Optimal harvest strategies for salmon in relation to environmental variability and uncertain production parameters. J. Fish. Res. Board Can. 32:1777-1784.

WALTERS, C. J., AND R. HILBORN.
1976. Adaptive control of fishing systems. J. Fish. Res. Board. Can. 33:145-159.
1978. Ecological optimization and adaptive management. Annu. Rev. Ecol. Syst. 9:157-188.

WHITT, W.
1978. Approximations of dynamic programs, I. Math. Oper. Res. 3:231-243.