A CONTAGION MODEL OF A FISHERY AND ITS USE IN ANALYZING CATCH-EFFORT DATA

W.J. Reed & C.M. Simons

DMS-697-IR July 1994

A Contagion Model of a Fishery and its Use in Analyzing Catch-Effort Data.

William J. Reed[1]
and
Clement M. Simons[2]

Department of Mathematics and Statistics
University of Victoria
PO Box 3045
Victoria B.C.
Canada V8W 2Y2.

June 1994

---

[1] Research supported by NSERC grant OGP 7252.

[2] The research described in this article formed part of the author's M.Sc. thesis at the University of Victoria.

Abstract

A stochastic model of a fishery which recognizes a contagion effect in the fishing process (*i.e.* the fact that the events of individual fish being caught are not independent) is developed and is used to establish how the probabality distribution of the aggregated annual catch of the fishery depends upon the aggregated annual fishing effort. From this relationship, and the assumption of deterministic, non-linear population dynamics, it is possible to determine analytically a likelihood function of the various model parameters (catchability coefficient, two parameters of the Ricker stock-dynamic model and a contagion parameter). Maximum likelihood estimates and likelihood ratio confidence intervals for these model parameters can then be obtained numerically. The method is applied to two sets of catch-effort data. The performance of the method is assessed by comparing predictions of catch and catch-per-unit-effort with observed values for reserved data which were not used for model fitting. Also confidence estimates of the maximum sustainable yield are obtained and the credibility of such estimates discussed.

Keywords: Fishery dynamics, contagion, catch-effort data, maximum likelihood estimation.

1 Introduction.

Mathematical modelling has been used in a number of ways in problems relating to fishery exploitation and fisheries management. Broadly speaking the applications can be grouped into three broad classes, which could be labelled "descriptive", "prescriptive" and "statistical".

Descriptive models have been very useful in developing an understanding of the factors affecting fishery exploitation, in particular in developing an understanding of the reasons behind the non-sustainable depletion of so many fisheries and other potentially very productive renewable resources. For example the "Tragedy of the Commons" phenomenon is now probably best known from G. Hardin's 1968 *Science* article in which he coined the phrase and used as a metaphor for the mechanism underlying human population growth, the situation of independent herdsmen exploiting a common pasture, and through each pursuing his own rational self interest and increasing his herd, jointly bringing about the destruction of the commons and the impoverishment of all. However this common property phenomenon had been identified over a decade earlier by the economist H. Scott Gordon [1954] using a very simple mathematical model, as the primary cause of poverty among fishermen and of the depletion of fish stocks[1].

The major objective of prescriptive (or normative) models is the determi-

---

[1] While even earlier M. Graham in *The Fish Gate* in 1943 had identified the freedom of individual fishermen to expand their fishing effort in response to falling catch per unit effort as the ultimate cause of rent dissipation.

nation of "optimal" patterns of exploitation. The earliest models of this type were based on equilibrium methods, with the usual objective of maximizing the sustainable yield (MSY). Later dynamic models using more sophisticated methods of optimal control, both deterministic and stochastic, usually with the objective of maximizing the expected net present value of economic returns, were employed (see *e.g.* Clark [1990] and references therein). In fact most of the models of this type are a long way from being operational models for the actual management of fisheries. Their purpose rather is an analytic understanding of what would constitute a "socially optimal" pattern of resource exploitation in an ideal world (that is to say in a world not limited by informational deficiencies and other practical complications). In this respect some prescriptive models can be very close to descriptive models, and in fact can sometimes serve a descriptive function. For example Clark (*ibid.*) showed in the early 1970s, that even under sole ownership, the extinction of a renewable biological resource could be economically "optimal" for suitably high discount rates, and suggested that this phenomenon could have been behind the reduction of blue whale and other baleen whale stocks almost to the point of extinction, in the years leading up to the IWC ban on the harvesting of blue whales in 1965 and of other species in subsequent years.

There have been attempts to incorporate informational deficiences into prescriptive management models. The simplest involve recognition of uncertainty with respect to the size of the fish stock; the more comlex also recognize uncertainty in the population dynamics *etc.* In a series of papers C. Walters, R. Hilborn and their co-workers (see Hilborn & Walters [1992, p.514] for references) have advocated the adoption of methods of adaptive control and the importance of probing strategies to gain greater knowledge of the biological system being managed.

In spite of this impressive work, for the most part the actual management of fisheries has been based on ideas which, from the mathemetical point of view, are far simpler than the optimal control methodologies employed in many prescriptive models. The reason for this is that there is so much complexity and statistical uncertainty in the data which form the informational base upon which management decisions have to be made. The statistical problems in themselves are usually of such a formidable nature that to combine statistical and control aspects is beyond the current realm of possibility. Thus usually the statistical and control aspects of management are treated separately.

Statistical analysis is the third area in which mathematical modelling is employed in fisheries. Usually in order to analyze fisheries data it is necessary to have a parametric statistical model for the way in which the data were generated. One approach is to use models which are simple from a statistical point of view, but which, by and large, ignore biological factors. There are some merits, but many limitations, to this approach. An alternative approach is to specify mathematical models which are more realistic from the biological and fisheries point of view. From plausible assumptions concerning the fish population and the fishery one seeks to obtain parametric relationships between *observable* quantities, and then to estimate parameters using established paradigms of statistical inference. However because of the inherent non-linearity of most biological systems, and because of the time-series nature of most fisheries data the resulting statistical problems are often very difficult.

In this paper a particular problem in analyzing fishery data will be discussed, and the mathematical modelling aspects developed. The resulting statistical difficulties will become apparent. The problem is that of analyzing aggregated catch-effort data, with the objective of obtaining information about the fish population dynamics and about the fishing process. The work builds upon a previous approach to the problem (Reed [1986]).

In the next section the problem of analyzing catch-effort data is discussed and a brief survey of proposed methods for addressing the problem presented. In Section 3 a model for the fishing process which recognizes a contagion effect in the fishing process (*i.e* the fact that the events of individual fish being caught are not independent) is developed and a way in which maximum likelihood estimates of the model parameters can be obtained is discussed. In Section 4 the model and methods are applied to two data sets, and the performance of the method discussed. The paper closes with a brief discussion in Section 5.

2 Catch Effort Data and Analysis.

The easiest data to obtain for many fisheries and sometimes the only data available for management, are the time series of annual aggregated catches and annual aggregated fishing effort. Such data have been analyzed by a variety of different methods. On the one hand there are methods which for pragmatic purposes are based on statistically straightforward techniques, ignoring by and large the complexities of the dynamics of the fishery. Such methods could be termed *black box* methods (Reed[1986]), since the essential idea is to develop a "black box" which processes the input series (efforts) to produce an output series which closely mimics the observed output series (catches). The earliest methods along these lines involved the linear regression of catch (or its logarithm) on some measure of average effort over the previous few years (Gulland[1961], Fox[1970]). A more sophisticated method along these lines was developed by Mendelssohn[1980] using Box-Jenkins ARIMA techniques. Roff [1983] used an autoregressive model, which involved the simple regression of catch in year t on the product of effort in year t and catch-per-unit-effort in year $t-1$, and demonstrated the usefulness of this model in predicting catch and catch-per-unit-effort. He argued however that great caution should be used in any attempt to use catch-effort data to determine equilibrium catches.

Opposed to the black box models are what have been called *surplus production models* or *dynamic based models* (Ludwig, Walters and Cooke [1988]), which attempt to model in a biologicallly plausible way, the dynamics of the exploited population and sometimes as well the dynamics of the "population" of fishermen who are exploiting it. Early models along these lines are those of Schaefer [1954], Pella & Tomlinson [1969] and Schnute [1977]. In all of these papers the continuous-time logistic growth model or some simple variant of it was used to model the fish population dynamics. More sophisticated population dynamic models which included age structure and a flexible parametric class of stock-recruitment functions were employed by Deriso [1980] and Schnute [1985]. Reed [1986] included a stochastic model relating catch to effort, arguing that much of the observed variation in catches could be explained by the year-to-year variation in effort, and that stochasticity in the catch-effort relationship could be more important than environmentally or biologically induced randomness. Ludwig, Walters and Cooke [1988] included randomness from both sources, but were forced to use approximate methods to obtain parameter estimates. Berck & Johns [1991] assumed that fishing effort changed from year to year in response to the profits generated in the previous season. They included noise in this relationship along with noise in the fish population dynamics, and used Kalman filter methods to estimate the model parameters. More recently Simons [1993] has used Kalman filter methods in a model in which effort is regarded as an exogenous variable, an assumption appropriate to a regulated, as opposed to an open-access, fishery.

Of course any model is a simplification of reality, and it is impossible to include all aspects of the problem in a single model. However one can to some extent examine the importance of a particular potential component of a model by comparing the results of an analysis which includes it with those of one from which it is omitted. In this paper the phenomenon of *contagion* (or lack of independence) in the catching process will be examined in the context of the model of Reed [1986]. For many species of fish, especially those that form schools, the assumption that whether or not a given fish is caught at a given time is independent of whether or not other fish are caught at that time, is clearly an oversimplification. One would expect that the fact that one fish is caught would increase the probability that other fish are caught. This is a contagion effect and a simple mathematical model for contagion in the catching process is presented in the following section. This model is then used to determine the likelihood function for the model parameters and to obtain parameter estimates by the method of maximum likelihood.

3 A Model for the Fishing Process and the Derivation of the Likelihood Function.

To begin we shall ignore natural mortality throughout the fishing season and assume that a fish, at any time during the fishing season, can be in one of two states: the state (0) if it has not yet been caught; and the state (1) if it has been caught. Let $p_0(t)$ and $p_1(t)$ denote the probabilities of a given fish being in states 0 and 1 respectively, t time units after the start of the fishing season. Suppose that there are N fishing boats which could be exploiting the fishery, and let $I_j(t)$ be an indicator variable which assumes a value 1 if boat j is fishing at time t, and a vaue 0 if it is not fishing at that time (for $j = 1, \dots, N$). Let the probability that a given fish is captured by boat j in the time interval $[t, t + dt]$, given that it has evaded capture up until that time, be given by $\lambda_j I_j(t) dt + o(dt)$, where λ_j is a constant reflecting the effectiveness of the jth. boat, $(j = 1, \dots, N)$. From these simple assumptions it follows that $p_0(t)$ satisfies:

$$p_0(t + dt) = [1 - \sum_{j=1}^{N} \lambda_j \, I_j(t) \, dt] \, p_0(t) + o(dt). \tag{1}$$

Subtracting $p_0(t)$ from both sides, dividing by dt and passing to the limit gives the following ordinary differential equation for $p_0(t)$:

$$p_0'(t) = -[\sum_{j=1}^{N} \lambda_j \, I_j(t)] \, p_0(t), \tag{2}$$

with initial condition $p_0(0) = 1$.

This is easily solved to give the probability that a fish has evaded capture through the whole season (of duration T) as :

$$p_0(T) = exp\{-\sum_{j=1}^{N} \lambda_j \int_0^T I_j(t)\, dt\} = exp\{-\sum_{j=1}^{N} \lambda_j T_j\} \tag{3}$$

where T_j is the time the jth. boat spends fishing throughout the season *i.e.* the *fishing effort* exerted by boat j during the season. The aggregated effort of the whole fleet over the season is $E_. = \sum_{j=1}^{N} T_j$. In terms of the aggregated effort the probability of survival can be written

$$p_0(T) = exp\{-\bar{q}E_.\} \tag{4}$$

where

$$\bar{q} = \frac{\sum_{j=1}^{N} \lambda_j T_j}{\sum_{j=1}^{N} T_j} \tag{5}$$

is the *average catchability* of the fleet. It is a weighted average of the catchability coefficients λ_j of the individual boats (with weights proportional to the times spent fishing).

Since we are ignoring natural mortality it follows that the probability of a fish being captured during the whole season is

$$p_1(T) = 1 - exp\{-\bar{q}E_.\}. \tag{6}$$

If the further assumption were made that individual fish either were captured or survived *independently of one another* then it would follow that the number of fish caught during the season would follow a *binomial* $(B(X, p_1(T)))$

distribution. This is essentially the model used in Reed [1986][2]. However the assumption of independent captures is probably not a very reasonable one, especially for schooling species of fish. Thus ideally we would like to have a model for captures in which the fact that one fish has been caught increases the probability of others being caught.

A model for a *contagion effect* of this type that has been used in the literature is that based upon *Polyá's Urn scheme* (see *e.g* Feller [1968,Chap. V.2] for examples). Polyá's Urn scheme was originally described in terms of balls in an urn. In its multivariate form (Johnson and Kotz [1977]) it assumes that initially there are $b_0, b_1, \ldots, b_k$ balls of colours $0, 1, \ldots, k$, respectively, in an urn. Balls are drawn at random with replacement. Furthermore, if a ball of colour i is drawn, c additional balls of that colour are put into the urn before the next draw. This of course increases the probability of drawing a ball of the same colour on the next draw - a contagion effect. It is possible to write down the joint probability mass function of the numbers of balls of various colours after n drawings. However the expression is very complicated. Of more practical use is the limiting (continuous) distribution of the *proportions* $y_0, y_1, \ldots, y_k$ of balls of colours $0, 1, \ldots, k$ as the number of drawings n goes to infinity. It can be shown (Johnson and Kotz [1977]) that the limiting distribution of these proportions is the **k-variate Dirichlet**

---

[2] However Reed went on to approximate the binomial distribution by a root-normal distribution, which had the desired property, exhibited by the data, that the variance of catch was proportional to the mean catch

distribution with parameters $b_0/c, b_1/c, \ldots, b_k/c$.

The distribution has joint p.d.f

$$f(y_0, y_1, \ldots, y_k) = \frac{\Gamma(1/\rho)}{\prod_{i=1}^{k} \Gamma(\theta_i/\rho)} y_0^{(\theta_0/\rho-1)} y_1^{(\theta_1/\rho-1)} \cdots y_k^{(\theta_k/\rho-1)} \tag{7}$$

on the simplex $y_0 + y_1 + \cdots + y_k = 1,\ \ 0 \le y_i \le 1$, where $\Gamma(a)$ is the usual gamma function defined by

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt. \tag{8}$$

In equation (??) the parameter ρ is defined as

$$\rho = \frac{c}{\sum_{i=1}^{k} b_i}, \tag{9}$$

in other words ρ is the number of new balls added at each draw expressed as a fraction of the initial number of balls in the urn. It reflects the degree of contagion. With ρ near zero, the probability of a ball of a given colour being drawn is only slightly increased following the drawing of a ball of the same colour (there would be little contagion present); but with ρ larger the probability would be increased to a larger extent (and there would be a greater degree of contagion present). The parameters θ_r $(= \frac{b_r}{\sum_{i=1}^{k} b_i})$ represent the *unconditional* probabilities that the nth. ball drawn is of colour r $(r = 0, \ldots, k)$; they are also *expected* limiting proportions of balls of colours $r = 0, \ldots, k$

Thus the Dirichlet distribution can thus be thought of as representing the probability distribution of the proportions of objects falling into different categories, when there is dependence between the objects. The unconditional probability of an object falling in category r is θ_r, but given that one object is in category r increases the probability that there will be others in that category (a contagion effect)[3]. When the *contagion parameter* ρ is large the presence of one object in category r will significantly increase the probability of others being in that category; when ρ is small the effect on the probability will be also be small.

For the fishery model ignoring natural mortality during the fishing season, there are only two catagories, (0 = escaped and 1 = captured) with respective unconditional probabilities:

$$\theta_0 = p_0(T) = exp\{-\bar{q}E_.\} \tag{10}$$

and

$$\theta_1 = p_1(T) = 1 - exp\{-\bar{q}E_.\} \tag{11}$$

Since it is reasonable to assume that the event of one fish being caught increases the probability of others being caught, the Dirichlet distribution with $k = 2$, can be used as a statistical model for the probability density for the proportions y_0 and y_1 escaping and being captured by the fishery. Thus the probability density for the proportion y_1 of the initial population which is captured by the fishery (with the remaining proportion y_0 $(= 1 - y_1)$

---

[3] This model was used by Reed [1994] in a different context for estimating historical forest fire hazards, where the assumption that individual stands burn or do not burn independently of one another is clearly not reasonable.

escaping the fishery) is:

$$f(y_1) = \frac{\Gamma(1/\rho) y_1^{(\theta_1/\rho - 1)} (1 - y_1)^{(\theta_0/\rho - 1)}}{\Gamma(\theta_1/\rho)\Gamma(\theta_0/\rho)} \tag{12}$$

This is a **Beta distribution** with parameters θ_0/ρ and θ_1/ρ. We now consider how this model can be fitted to data by the method of **maximum likelihood.**

Suppose $\{(C_i, E_i)\}_{i=1,\ldots,n}$ is the observed bivariate time series of annual aggregated catches and efforts. The probabilities of a fish being captured and escaping in year i are (from (11) and (10)) $exp\{-qE_i\}$ and $1 - exp\{-qE_i\}$, while the proportion captured in year i is $\frac{C_i}{X_i}$, where X_i is the size of the population at the beginning of the fishing season in year i. Substituting these into (12) and taking the product for $i = 1, \ldots, n$ gives the probability (density) of observing such data, or in other words the **likelihood function** for the parameters[4] q and ρ. The **log-likelihood** is then

$$\begin{aligned} & n \log \Gamma(1/\rho) - \sum_{i=1}^{n} \log \Gamma(e^{-qE_i}/\rho) - \sum_{i=1}^{n} \log \Gamma(\frac{1 - e^{-qE_i}}{\rho}) \\ & + \sum_{i=1}^{n} (\frac{e^{-qE_i}}{\rho} - 1) \log(\frac{X_i - C_i}{X_i}) + \sum_{t=1}^{n} (\frac{1 - e^{-qE_i}}{\rho} - 1) \log(\frac{C_i}{X_i}) \end{aligned} \tag{13}$$

We cannot simply maximize the log-likelihood over the parameters q and ρ because the log-likelihood involves the X_i, a sequence of unobserved variables. One possibility is to identify a surrogate variable such as catch per unit effort for the unobserved X_i, and use an approach along the lines of the Kalman

---

[4] Here we assume the same fixed "average" catchability q in every season.

filter, in which the X_i are predicted. This idea is pursued in Simons [1993] (without the contagion effect). Here we shall adopt a different approach and assume a *deterministic* model for the stock dynamics, during the closed season. Specifically we assume that the "returning stock" X_{i+1} in year $i+1$ is related to the "escapement" from the fishery $S_i = X_i - C_i$ in year i by a relationship of the form

$$X_{i+1} = F(S_i; \Theta) \tag{14}$$

where $F(\cdot;\Theta)$ is a parametrically specified function with a vector, Θ, of parameters. With this assumption the log-likelihood can be expressed as a function of ρ, q, Θ and the initial stock size, X_1. It can then be maximized (numerically) over these parameters, subject to the constraints $C_i \leq X_i$, for $i = 1, 2 \cdots, n$. Details of this procedure are given in the next section.

Before proceding however we note that in the above development natural mortality has been ignored. For fisheries with a relatively short fishing season, such an omission is likely to be of little consequence. For fisheries with a longer fishing season however it could be more serious. In principle one could include three states: 0 = escaped and 1 = captured, as above, with an additional state, 2, corresponding to natural death. Using unconditional probabilities θ_0, θ_1 and θ_2, one could then use a Dirichlet distribution with $k = 3$, for the proportions in the three classes. However such a model would assume similar contagion effects for all three classes, and in this respect it would probably be inadequate, in that natural mortality would, in reality, probably exhibit much less contagion than capture by the fishery. In consequence we shall not pursue this avenue further. Rather we shall assume that a *fixed proportion, m*, of fish suffer natural mortality throughout the season. If natural deaths occur independendtly then the variance in the proportion dying naturally should be inversely proportional to the number of fish initially in the population. Thus the variance in the proportion dying naturally should be small, and there should be no great loss in assuming m constant. The probabilities θ_0 and θ_1 in (10) and (11) should now be regarded as conditional probabilities, given that the fish does not die naturally. Also in (13) the proportions surviving and captured should be expressed as proportions of fish which do not die naturally; thus $\frac{(X_i-C_i)}{X_i}$ should be replaced by $\frac{(X_i-mX_i-C_i)}{X_i-mX_i}$ and $\frac{C_i}{X_i}$ by $\frac{C_i}{X_i-mX_i}$; this is equivalent to replacing X_i by $(1-m)X_i$ for $i = 1, \ldots, n$. This effect can be absorbed entirely in the stock dynamic model (14). In other words the log-likelihood function (??) remains valid even when deterministic natural mortality is included in the model. However if fishing continues year round then the above model is not appropriate, since it assumes that growth and recruitment *etc.* are occurring during the closed season (via (14)) and not during the fishing season. Thus the model is for a seasonal fishery even if the assumption of a short season is no longer necessary.

4 Application to Two Data Sets.

For a given time series $\{(C_i, E_i)\}_{i=1,\ldots,n}$ of catches and efforts and for given values of the parameters q, ρ, Θ and X_1, one can calculate numerically the log-likelihood (??), by iterating the stock dynamic equation (??). To find maximum likelihood (ML) estimates of the parameters one can then maximize (??) numerically over the paramaters. To do this one requires a numerical maximization scheme which does not require explicit expressions for the partial derivatives. Such a method is the simplex algorithm of Nelder & Mead [1965], and it proved succesful in accomplishing the numerical maximizations required.

The method was applied to two sets of catch-effort data: (1) North Atlantic redfish *(Sebastes marinus)* in NAFO division 3L(Power[1992]) and (2) Pacific rocksole *(Lepidopsetta bilineata)* in area 5C off the coast of British Columbia. (Fargo[1991]).

These data are plotted in Figures 1 and 2. It is apparent from the plots that overall catch increases with effort, and furthermore that the variance in catch increases with effort and with catch size. From the properties of the Beta distribution it is easily verified from (??) that for the Dirichlet contagion model the expected catch and variance in catch are:

$$\mathrm{E}(C_i) = (1 - e^{-qE_i})X_i \tag{15}$$

and

$$V(C_i) = \frac{\rho}{1+\rho}(1 - e^{-qE_i})e^{-qE_i}X_i^2 \tag{16}$$

Both of these quantities increase with effort E_i (at least over the range $e^{-qE_i} < \frac{1}{2}$ *i.e* provided less than fifty percent of the stock is harvested), and thus appear compatible with the data plotted in Figures 1 and 2.

A parametric stock dynamic function F (in equation ??) is required before the procedure can be implemented . The form used was the **Ricker model** (Ricker[1954]) which has been widely used in the fisheries and ecology literature. The parameterization used was

$$X_{i+1} = F(S_i; \Theta) = S_i e^{\beta(1-\frac{S_i}{X_e})} \tag{17}$$

where $\Theta = (\beta, X_e)'$; the parameter X_e represents the unfished equilibrium stock level and β is a parameter related to the intrinsic growth rate.

Estimates of the model parameters were made using catch-effort data for the first 29 years of each series. Data for later years were reserved for comparison with the one step-ahead predictions from the fitted model. Point (ML) estimates of the model parameters were obtained by maximizing the log-likelhood function. Approximate confidence intervals for the parameters were obtained using the likelihood ratio procedure (see *e.g.* Cox and Hinkley [1974, p. 322.] for general theory and Reed [1986] for an application to a similar problem). However parameter estimates can be correlated and so some care must be taken in interpreting these confidence intervals. Joint confidence regions for sets of parameters could also have been found using the likelihood ratio method, although we do not report such results here. ML point estimates and approximate 95% confidence intervals are given in Table 1.

Using the ML estimates of the various parameters and the estimates $\hat{X}_i$ of the stock sizes one can obtain (from (??)) "back estimates" (or fitted values) $\hat{C}_i = (1 - e^{-\hat{q}E_i})\hat{X}_i$ of the historical catches. Fig.?? shows the observed and back-estimated (fitted) catch for the redfish data. The solid lines join the actual catch, dotted lines the estimated catches. The agreement is very good. It should be noted that the closeness of the fit was not the criterion used to estimate the parameters. Fig.?? displays the analogous results for rocksole. For neither data set did an examination of the residuals $(C_i - \hat{C}_i)$ provide any evidence of model lack of fit.

One can obtain ML estimates of the maximum sustainable yield (MSY) by using the invariance property of ML estimates. For the Ricker model, the MSY is given by

$$S_0\, e^{\beta(1-\frac{S_0}{X_e})} - S_0 \tag{18}$$

where S_0 is a solution to

$$\frac{dF}{dS} = (1 - \beta S/X_e)e^{\beta(1-S/X_e)} = 1. \tag{19}$$

The effort level corresponding to MSY is obtained from (??) as $E_0 = \frac{\beta}{q}\frac{(1-S_0)}{X_e}$. Note that because of the assumed randomness in the catching process the effort E_0 will produce a catch which is a random variable with *expected* value equal to MSY. Substituting in the ML estimates of q, β and X_e provides ML point estimates of the MSY and the corresponding effort level.

For the redfish data the ML estimate of MSY is 80.8 thousand tons while that of the corresponding effort E_0 is 87.3 thousands of hours. Using the method discussed in Appendix 3 of Reed [1986], based on the observed information matrix, the standard error of the MSY estimate was estimated as 12.73 thousand tons. Thus an approximate confidence interval using a two-standard-error band for the MSY is 61.8 - 112.8 thousand tons. (In comparison the method of Reed [1986] which ignores contagion effects provides an estimate of MSY of 137.3 with a standard error of 736.3 (thousand tons). A confidence interval based on this standard error is so wide as to be meaningless.) For the period covered by the data, the greatest annual catches were of the order of 35 thousand tons, which is considerably less than the estimated MSY. One possibility is that the stock has been underexploited in the past. However an examination of the "fitted values" $\hat{X}_i$ of the biomass and the corresponding estimated escapement $\hat{S}_i = \hat{X}_i - C_i$ reveals that most of the points $(\hat{S}_i, \hat{X}_{i+1})$ lie close together and near the estimated equilibrium point $(\hat{X}_e, \hat{X}_e)$, which furthermore is on the *downward sloping* part of the Ricker curve. This means that the mode of the fitted Ricker curve lies below $\hat{X}_e$, and in consequence implies that the MSY would be achieved by allowing an escapement considerably smaller than that permitted in the past. Thus the estimation of the MSY involves extrapolating on the Ricker curve well outside of the range of fitted values. In consequence the estimation is critically dependent on the parametric form assumed in the Ricker model, which cannot be empirically verified outside of the relatively narrow range of the fitted values. In consequence not too much faith should be put in the estimate of MSY obtained here. To get a better idea of the potential MSY, more "probing" of the stock dynamics is required, through varying the catch over a much wider range (see Hilborn & Walters [1992, p.514]). Only by varying the escapement over a much wider range, can a better idea of the potential biological productivity of the resource be obtained.

For the rocksole data the ML estimate of MSY is 580.7, with a standard error of 10.26 (hundreds of tons). Thus an approximate 95% confidence interval for the MSY is 560.2 - 601.2 (hundreds of tons). (In comparison the ML estimate from the no-contagion model is, in the same units, 428.9 with a standard error of 39.90.) Over the period of the observations the observed catch exceeded the estimated MSY once during the 1950s, for the whole of the period 1965-70 and once again in the late 1970s. It therefore appears to provide a more credible estimate than does the corresponding estimate for the redfish. Indeed this is borne out by a plot of the fitted values $(\hat{S}_i, \hat{X}_{i+1})$, which all lie to the left of the estimated equilibrium point and on the *ascending arm* of the fitted Ricker curve. Thus estimating the MSY for rocksole involves considerably less extrapolation beyond the range of the fitted values than does the corresponding estimation of MSY for redfish. In consequence more (but still something less than complete) confidence can be placed in the estimate.

To assess the performance of the model in predicting future catches, estimates of the expected catch and the standard deviation in catch in year $T+1$ given data up to year T and the effort E_{T+1} in year $T+1$ were made using (*c.f.* (??),(??)) as predictors

$$\hat{E}(C_{T+1}) = (1 - e^{-\hat{q}E_{T+1}})\hat{X}_{T+1} \tag{20}$$

and the estimated variance of the catch in year $T+1$ is given by

$$\frac{\hat{\rho}}{1+\hat{\rho}}(1 - e^{-\hat{q}E_{T+1}})e^{-\hat{q}E_{T+1}}\hat{X}^2_{T+1} \tag{21}$$

where $\hat{q}$, $\hat{\rho}$ and $\hat{X}_{T+1}$ are maximum likelihood estimates using the data up to year T.

Since the sampling error in the parameter estimates is not taken into account in these formulas, the actual variance of the prediction of the catch could be much larger than that given by (??). Table **??** gives the estimates of the mean and standard deviation of the catch in year $T+1$ given effort in year $T+1$ and catch-effort data up until year T for the redfish data for $T = 30, \ldots, 33$, both for the contagion model and for the corresponding model without contagion. (*c.f.* Reed [1986]). Corresponding results for rocksole for $T = 30, \ldots, 34$, are given in Table ??.

For the redfish data the predictions are very good, with the observed catch lying within one estimated standard deviation of the estimated expected catch, for all years except the last, 1991. In this case the difference slightly exceeds two standard deviations; however bearing in mind that the sampling error in the parameter estimates has been ignored, the prediction seems quite adequate. The *prediction error sum of squares* for the catch predictions[5]

$$S = \sum_{T=30}^{n^*} (C_{T+1} - \hat{C}_{T+1})^2 \tag{22}$$

is 72.48, while for the no-contagion model it is 94.47. The prediction error sum of squares for catch per unit effort

$$H = \sum_{t=30}^{n^*} (\frac{C_t}{E_t} - \frac{\hat{C}_t}{E_t})^2 \tag{23}$$

is 0.35 for the contagion model and 0.44 for the no-contagion model. Thus the inclusion of the contagion effect seems to improve the predictive ability of the model for the redfish fishery.

The prediction results are less impressive for the rocksole fishery. The model with contagion still does better than the no-contagion model (S=16,012 compared with 33,538; and H=.0779 compared with .1077)[6]. However for all years, except 1988, the estimated expected catch is larger than the observed catch by between 22% and 62%. For the year 1989 it is smaller by 31%.

---

[5] where $\hat{C}_{T+1}$ is $\hat{E}(C_{T+1})$ in (??) and n^* is 33 for redfish and 34 for rocksole.

[6] Since the scales of catch and effort are not comparable for the two fisheries a comparison of S and H between the two fisheries has no meaning.

Furthermore the observed catch deviates from the estimated expected catch by up to almost seven estimated standard deviations above (in 1989) and twelve estimated standard deviations below (in 1988). However as noted above, sampling error in the parameter estimates is not included in the calculation of standard deviations, and so these results may not be as bad as they at first appear. The fact that the confidence intervals for the stock dynamic parameters (β, X_e and X_1) indicate considerably greater relative sampling error for the rocksole than for the redfish parameters suggests that sampling error plays a more important role for the rocksole predictions, and that the true standard deviations of the catches, when sampling error is included, may be considerably larger than those reported in Table ??.

5 Concluding Remarks.

In this paper a method of analyzing fishery catch-effort data has been presented and applied. Like the method of Reed [1986] the method discussed here recognizes the intrinsic uncertainty in the process of catching fish, while ignoring randomness in the stock dynamics. While undoubtedly randomness from both sources is present in the real world, incorporating it in a model which permits estimation of model parameters leads to very formidable statistical problems, which to date have not been satisfactorily resolved. The justification for ignoring the randomness in the population dynamics is based on the hypothesis that randomness from this source is of smaller magnitude, and consequently of less importance, than randomness in the fishing process. As Reed [1986] demonstrates with elementary arguments, one would expect that the variance in annual catch would increase more or less linearly with the size of the annual aggregated effort. A glance at Figures 1 and 2 will confirm that this situation prevails for the data sets analyzed herein. Thus much of the observed variation in annual catch can be explained by the variation in annual effort. The method of Reed [1986] and the method of this paper formalize this insight by using simple stochastic models for the fishing process, which enable the probability of observing the given data to be expressed in terms of the model parameters (*i.e* enable the computation of a likelihood function). The way in which the model in this paper differs from the earlier one is that it explicitly recognizes a contagion effect in the process of catching fish, that is to say it recognizes that the events of individual fish being caught (or escaping the fishery) are *not* statistically independent. The fact that one fish is caught will have the effect of increasing the probability that other fish are caught. This phenomenon will be especially true for schooling fish, and for fisheries where much of the fishing effort is expended in searching for fish. The contagion model employed in the paper introduces one new parameter which can be interpreted as a measure of the degree of contagion.

The modelling aspect of the problem involves developing a model for the fishery which includes the contagion effect, and for which the probability of observing the given catch series can be derived. This is accomplished in Section 3. In Section 4 the model is fitted to data series for two fisheries and maximum likelihood estimates (and likelihood ratio confidence intervals) of the model parameters are obtained. These ML estimates are used to obtain ML estimates of the maximum sustainable yield (MSY) and approximate confidence intervals for the MSY. In addition the estimates are used to asssess the predictive power of the model, by comparing model catch predictions with observed catches for the last few years of data in each series which were reserved for this purpose, and not used in model fitting. The results are compared throughout with those of the no-contagion model of Reed [1986]. It turns out that effectively in all respects the contagion model performs better than the model without contagion.

For one data set (rocksole) the estimate of MSY seems very plausible, but for the other (redfish) the estimated MSY exceeds all past observed catches. The reason for this discrepancy is discussed in Section 4, and hinges on the fact that for the redfish data the estimation of MSY involves extrapolation along a Ricker curve well outside the range of fitted values. In this case it would seem that the recommendations of Hilborn & Walters [*op.cit.*] concerning the advisability of a probing strategy, in which catches are varied experimentally to a high degree in order to learn more about the population dynamics, would be in order. The degree of extrapolation in the estimation of MSY for the rocksole fishery is much less, and the estimates much more plausible

In contrast, the catch predictions for redfish seem very good, all falling well within the range of the estimated variability predicted by the model, while those for the rocksole fishery are less impressive. The reason behind the difference in predictive performance for the two fisheries probably lies partly in the fact that sampling error in parameter estimates is ignored in the estimation of the standard deviation of predicted catches and the fact that, relatively speaking, sampling error is larger in the parameter estimates for the rocksole fishery than for the redfish fishery. Thus neglecting sampling error has relatively more serious consequences for the rocksole fishery. Incorporating analytically the sampling error of the parameter estimates in the calculation of the standard deviation of catch predictions is very difficult, because the predictions depend on the parameter values in a complicated fashion involving the multiple iteration of the non-linear Ricker population-dynamic model.

The results of the applications to the two data series examined would suggest that overall the inclusion of the contagion effect adds to the effectiveness of this method of analyzing catch-effort data. However it would be too much to claim that the new method provides a completely satisfactory answer to the challenge of using catch-effort data as a basis for fishery management. The method seems to do a better job of explaining the dependence of catch upon effort than does the corresponding model without contagion. However because of the presumed non-linearity of the stock dynamics and the relatively narrow range of observed catches and efforts for many fisheries, it is difficult to predict equilibrium behaviour outside of the observed range, and in consequence in many cases to predict maximum sustainable yield. The difficulty here lies more with the limited nature of the data, rather than with the method, and it is shared by all other methods of analyzing catch-effort data (see Roff [1983]). In fact it may be too much to expect that effective management can be based on such a narrow informational base. It may well be the case that additional information (for example in the form of independent estimates of stock abundance or in details of the age structure of the catch *etc.*) may be required in order to be able to effectively understand both the dynamics of the fish population, and those of the fishery dependent upon it.

References

[1] Berck, P., and G. Jones. 1991. Estimating structural resource models when stock is uncertain: Theory and its application to Pacific halibut. in *Stochastic Models and Option values: Application to Resource, Environment and Investment problems.* D. Lund and B. Øksendal (eds.). North Holland.

[2] Clark, C.W. 1990. *Mathematical Bioeconomics. The Optimal Management of Renewable Resources.* Second Edition. Wiley, New York, NY.

[3] Cox, D.R., and D.V. Hinkley. 1974. *Theoretical Statistics.* Chapman and Hall. London.

[4] Deriso, R.B. 1980. Harvesting strategies and parameter estimation for an age-structural model. *Can.J.Fish.Aquat.Sci.* **37**:268-282.

[5] Draper, N.R., and H.S. Smith. 1981. *Applied Regression Analysis.* Second Edition. John Wiley & Sons. New York NY.

[6] Fargo, J., and B.M. Leaman. 1991. Groundfish stock assessment for the west coast of Canada in 1990 and recommended yield options for 1991. *Can. Tech. Report Fish. Aquat. Sci.* **1778**.

[7] Feller, W. 1968. *An Introduction to Probability Theory and its Applications.* John Wiley & Sons. New York NY.

[8] Fox, W.W. 1970. An exponential yield model for optimizing exploited fish population. *Trans. Am. Fish. Soc.* **99**:80-88.

[9] Gordon, H.S. 1954. The economic theory of the common property resource: the fishery. *J. Political Economy* **80**:1031-1038.

[10] Graham M. 1943 *The Fish Gate.* Faber and Faber, London.

[11] Gulland, J.A. 1961. Fishing and the stocks of fish at Iceland. *U.K. Minist. Agric. Fish. Food Fish. Invest.(ser.2)* 23(4).

[12] Hardin, G. 1968. The tragedy of the commons. *Science* **162**:1243-1247

[13] Hilborn, R., and C.J. Walters. 1992. *Quantitative Fisheries and Stock Assessment: Choice, Dynamics and Uncertainty.* Chapman and Hall.New York NY.

[14] Johnson, N.L., and S. Kotz. 1977. *Urn Models and Their Application.* J. Wiley & Sons. New York.

[15] Ludwig, D., C.J. Walters and J. Cooke. 1988. Comparisms of two models and two estimations methods for catch-effort data. *Nat. Res. Model.* **2**:457-498.

[16] Mendelssohn, R. 1980. Using Box-Jenkins models to forecast fishery dynamics: Identification estimation and checking. *Fish. Bull.* **78**:887-896.

[17] Nelder, J.A., and R. Mead. 1965. A simplex method for function minimization. *Comput. J.* **5**:308-313.

[18] Pella, J.J., and P. K. Tomlinson. 1969. A generalized stock production model. *Bull. Inter-Am. Trop. Tuna Comm.* **13**:419-496.

[19] Power, D. 1992. An assessment of Redfish in NAFO division 3LN. *NAFO Sc. Doc.* 92/80.

[20] Reed, W.J. 1986. Analyzing catch-effort data allowing for randomness in the catching process. *Can. J. Fish. Aquat. Sc.* **43**:174-186.

[21] Reed, W.J. 1994. Estimating the natural fire hazard from the age-class distribution of old-growth forest. *For. Sci.* **40**:104-119

[22] Roff, D.A. 1983. Analysis of catch/effort data: A comparism of three methods. *Can. J. Fish. Aquat. Sc.* **40**:1496-1506

[23] Schaefer, M.B. 1954. Some aspect of the dynamics of the populations important to the management of the commercial marine fisheries. *Bull. Inter-Am. Trop. Tuna Comm.* **1**:27-56.

[24] Schnute, J. 1977. Improved estimates from the Schaefer production models: theoretical consideration. *J. Fish. Res. Board Can.* **34**:583-603.

[25] Schnute, J. 1985. A general theory for analysis of catch and effort data. *Can. J. Fish. Aquat. Sc.* **42**:414-429.

[26] Simons, C. 1993. Some methods of analysis of fishery catch/effort data. M. Sc. thesis. Department of Mathematics and Statistics, University of Victoria.

parameter	Redfish data	Rocksole data
$\hat{q}$	0.0098 [0.006 – 0.034]	0.0036 [0.002 – 0.006]
$\hat{X}_1$	131.35 [110.3 – 159.2]	144.03 [100.6 – 178.4]
$\hat{X}_e$	148.67 [132.6 – 180.7]	163.88 [101.2 – 188.5]
$\hat{\beta}$	1.43 [0.6 – 1.87]	0.1369 [0.06 – 0.46]
$\hat{\rho}$	0.003 [0.0008 – 0.005]	0.004 [0.001 – 0.008]

Table 1: Maximum likelihood estimates and approximate 95% confidence limits (in square brackets) for the model parameters for redfish data and rocksole data.

Year	Effort	observed catch	estimated expected catch	estimated expected catch (Reed[1986])
1988	20.205	26.267	27.354(4.458)	26.623(3.013)
1989	12.267	19.847	16.902(3.524)	16.190(2.310)
1990	17.409	17.707	21.600(4.071)	23.425(2.831)
1991	14.171	11.642	18.532(3.027)	18.597(2.638)

Table 2: One-step-ahead estimates of the expected catch and standard deviation(in brackets) of the upcoming catch for the years 1988 – 91 for redfish in NAFO division 3l. Units of catch are thousand of metric tons. Units of effort are thousand of hours.

Year	Effort	observed catch	estimated expected catch	estimated expected catch (Reed[1986])
1985	92	28	45.57(6.29)	41.742(2.36)
1986	497	86	113.20(6.60)	222.99(5.37)
1987	496	209	262.77(15.21)	178.28(6.25)
1988	545	189	113.17(6.24)	196.58(6.49)
1989	1476	406	532.54(18.68)	522.54(11.04)

Table 3: One-step-ahead estimates of the expected catch and standard deviation(in brackets) of the upcoming catch for the years 1985 – 89 for rock sole in area 5C.Catches are in tons and effort are in hours.

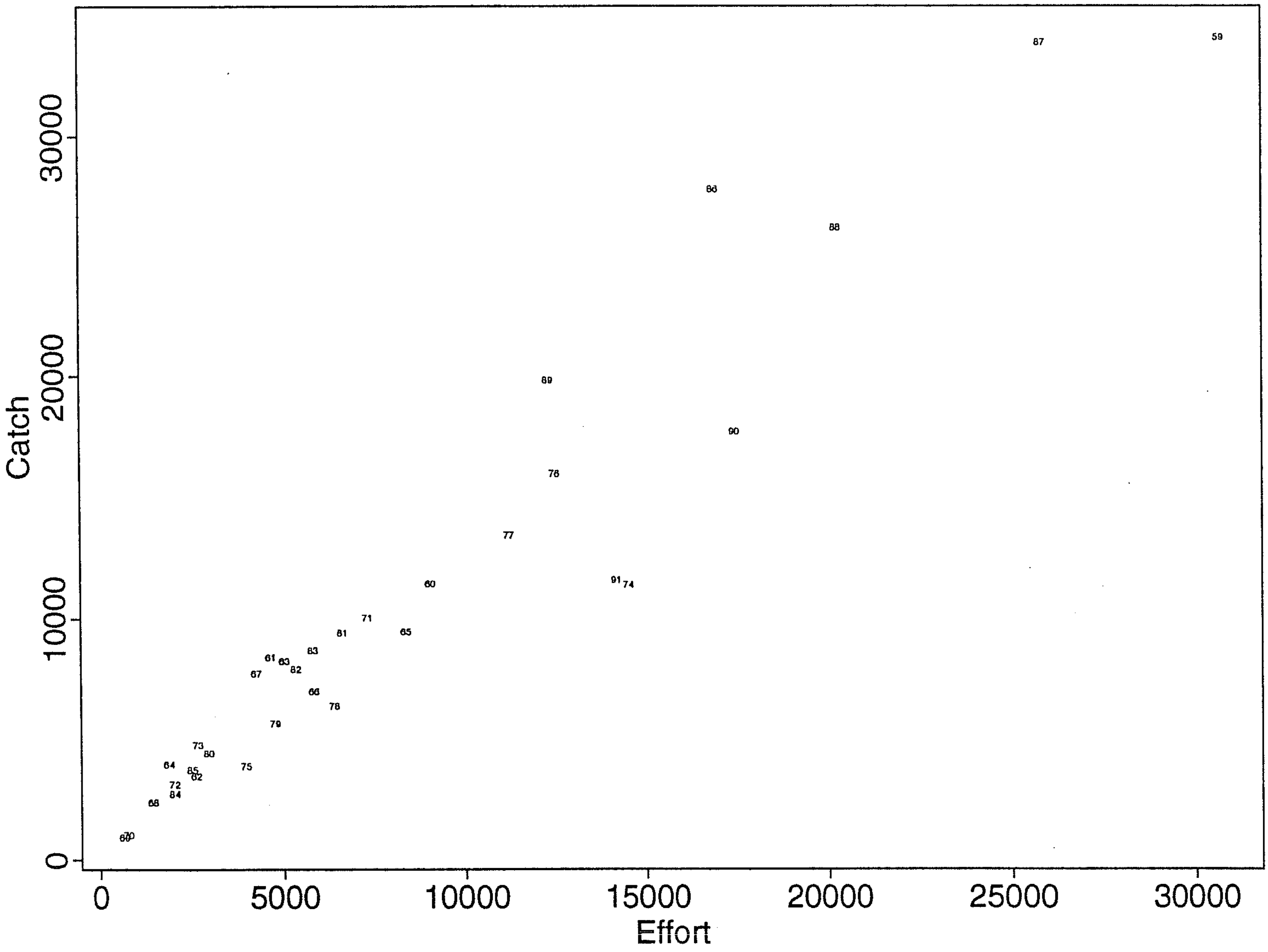

Figure 1: Total annual catch versus total annual effort for redfish in NAFO division 3L for the years 1959 – 1991. Units of effort are hours and units of catch are metric tons

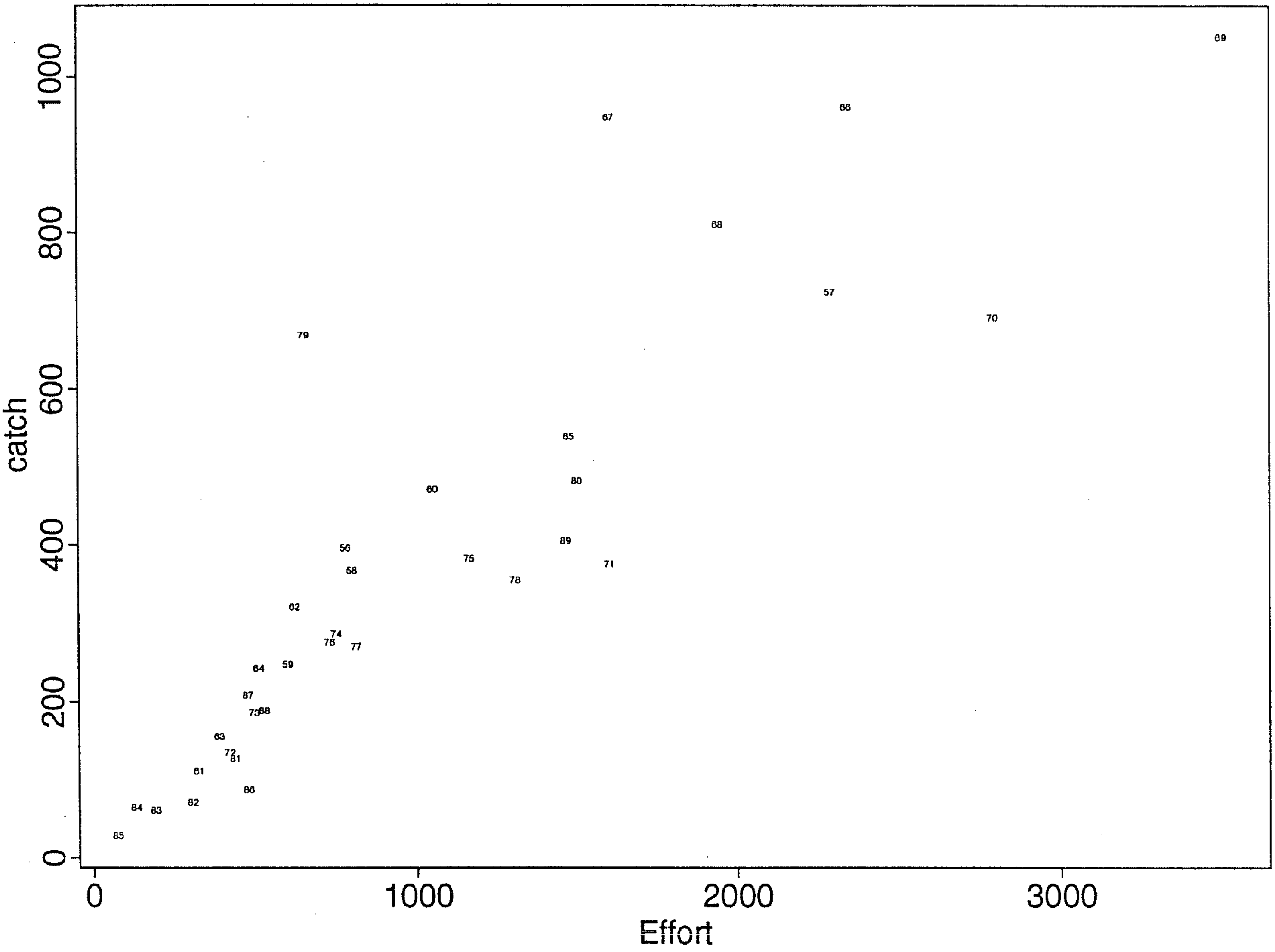

Figure 2: Total annual catch versus total annual effort for rock sole in area 5C off the coast of British Columbia for the years 1956 – 1989. Units of effort are hours and units of catch are metric tons

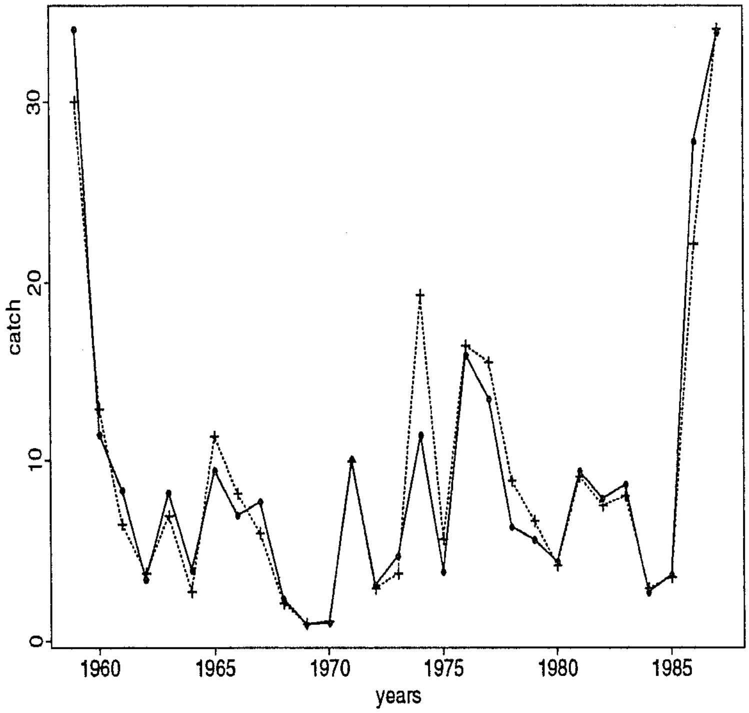

Figure 3: Annual redfish catch $C_i(10^2t)$ for the years 1959 – 1987 (dots (.) joined by solid line segments) and fitted annual catch $\hat{C}_i$ computed from the maximum likelihood estimates of the model ((+) joined by dotted line segments).

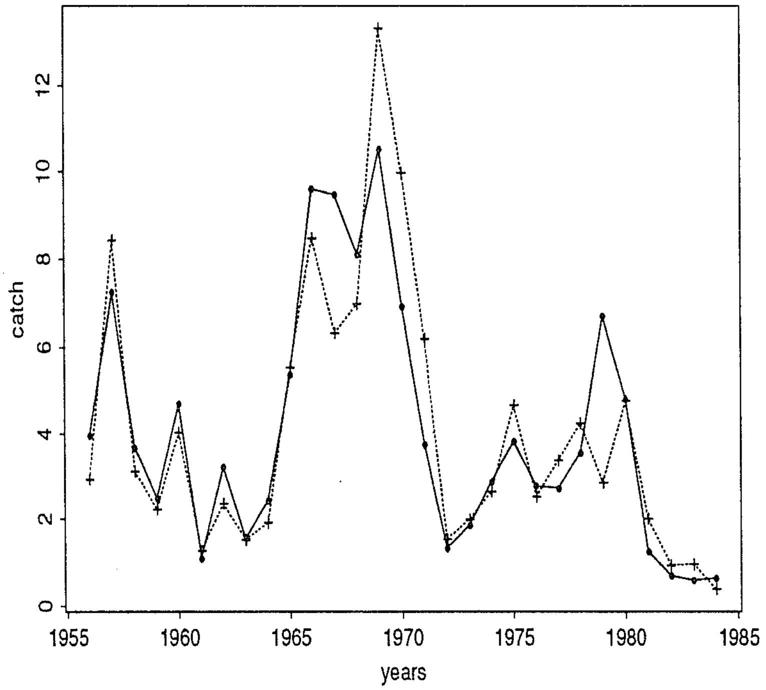

Figure 4: Annual rocksole catch $C_i(10^2 t)$ for the years 1956 – 1985 (dots (.) joined by solid line segements) and fitted annual catch $\hat{C}_i$ computed from the maximum likelihood estimates of the model ((+) joined by dotted line segments).