Northwest and Alaska Fisheries Center

National Marine Fisheries Service

U.S. DEPARTMENT OF COMMERCE

NWAFC PROCESSED REPORT 88-21

A Review and Comparison of Age-Structured Stock Assessment Models from Theoretical and Applied Points of View

September 1988

A REVIEW AND COMPARISON OF AGE-STRUCTURED STOCK ASSESSMENT MODELS FROM THEORETICAL AND APPLIED POINTS OF VIEW

by

Bernard A. Megrey

U.S. Department of Commerce
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Northwest and Alaska Fisheries Center
7600 Sand Point Way NE
Seattle, WA 98115-0070

September 1988

PREFACE

This report is an updated and extensively revised version of "Review and Comparison of Three Methods of Cohort Analysis" a 1983 NWAFC Processed Report, Number 83-12 .

ABSTRACT

The development of age-structured stock assessment methodology (commonly referred to as cohort analysis or VPA) is reviewed. A comprehensive review of the literature on age-structured stock assessment methods is presented in an attempt to place the development of the methodology in a historical perspective. The review describes reasons why stock assessment models play an important role in fisheries management and traces the historical origins of the different mathematical models, their chronological development, and various levels of complexity.

Both catch and catch-per-unit-effort models are examined with special emphasis being placed on models incorporating the separability assumption. Specifically the methods of Derzhavin (1922), Fry (1949, 1957), Gulland (1965), Murphy (1965), Pope (1972), Doubleday (1976), Paloheimo (1961, 1980), Pope and Shepherd (1982), Fournier and Archibald (1982), Dupont (1983) and Deriso et al. (1985) are compared paying particular attention to similarities, differences, strengths, weaknesses, data requirements, underlying assumptions, sources of error, what parameters they estimate, mathematical formulation, and parameter estimation techniques. Useful extensions to the basic stock assessment procedures are also discussed.

1.0 INTRODUCTION

Stock assessment methods for age-structured animal populations were developed over sixty years ago to facilitate the analysis and interpretation of commercial catch statistics. Age-structured stock assessment methods are ideally suited to the rational management of fisheries resources because application of these powerful analytic techniques permit reconstruction of the population dynamics of exploited fish stocks and provide estimates of vital mortality rates and absolute population abundance. Today these population assessment methods serve as the primary basis for providing management advice in many world fisheries.

The main advantage of age-structured stock assessment models over more traditional approaches such as stock production (Graham 1935, 1939; Schaefer 1954, 1957) or dynamic pool (Beverton and Holt 1957) models are that they can be applied without knowledge of effective fishing effort, catchability or gear selectivity. Thus they do not suffer from many of the problems associated with using CPUE as an index of population abundance. Early success with these methods after their appearance in the early 1960's and the greater availability of aged catch data led to their widespread application. A heightened awareness of the utility of stock assessment models has, in the past few years, resulted in a dramatic proliferation of new methodology. As researchers embraced the methodology they naturally improved upon the original idea and in several cases parallel development has taken place. Figure 1 shows how the level of research activity in the area of age-structured stock assessment methods has increased over the decades since it was introduced. Today age-structured stock assessment methods constitute one of the principal research tools used by fisheries scientists to evaluate the status of exploited fish stocks.

In addition to widespread application, a number of important new developments have been proposed in the past few years. This has prompted Beamish (1986) to suggest that perhaps it was time for a review paper on the assumptions that go into age-structured stock assessment models and the consequences of not meeting those assumptions. In view of the profusion of new methods, especially since the newer techniques have been introduced only recently and in relatively quick succession, it seemed desirable to review old and new stock assessment methods and compare them to historic paradigms. Thus, the objective of this paper is to review and compare current and historic stock assessment methods. Because this paper is focused on review and comparison, it does not present any new theory.

At the foundation of each stock assessment method is one or more mathematical equations that symbolize a hypothesis about the way we believe dynamic changes in an exploited population take place. In subsequent discussion these mathematical expressions will be referred to as the model. Even though all age-structured stock assessment methods share the same theoretical underpinnings, the numerous methods proposed throughout the past sixty years are quite different with respect to mathematical formulation, which parameters they estimate, and solution techniques. Because of these differences, each model has its own strengths and weaknesses and different methods contain different sources of error. Moreover, application of more than one method to a common data set may not give identical results. Since few comparative studies have been carried out, it is not clear which method is best to use under a given set of circumstances.

In order to better state the model's assumptions and consequences each model is described in a standard format which includes a statement of the model using a consistent notation, the objective function, parameter estimation procedure, assumptions, advantages, and disadvantages. In particular, each model is described with respect to data requirements, underlying biological and technical assumptions (both implicit and explicit), sources of error, what parameters they estimate, mathematical formulation, and parameter estimation techniques. Also strengths, weaknesses, advantages, disadvantages, similarities, and differences of each method will be compared. The symbols and notation used to describe the models are defined in Table 3. Underlying each model are assumptions regarding the specific mathematical model, the input data required by the model, and any solution methods or statistical procedures used to estimate the parameters of the model. Assumptions, advantages, and disadvantages underlying each model are summarized in Tables 4, 5, and 6 respectively. A summary of the mathematical models, error models, parameter estimation methods, and parameter estimation sequence underlying each model are summarized in Table 7.

The methods reviewed in this paper include those of Derzhavin (1922), Fry (1949, 1957), Gulland (1965), Murphy (1965), Pope (1972), Doubleday (1976), Paloheimo (1961, 1980), Pope and Shepherd (1982), Fournier and Archibald (1982), Dupont (1983) and Deriso et al. (1985). Two methods not included in the review are those of Doubleday (1981) and Collie and Sissenwine (1983).

1.1 Historical Development

The historical roots of fisheries stock assessment methods can be traced back to the U.S.S.R. in the beginning of the 19th century (Ricker 1971). A. N. Derzhavin (1922) was perhaps the first to conceive of the idea of applying data describing the age structure of a population to catch statistics in order to calculate the contribution of each cohort to each years total catch. The theoretical basis for Derzhavin's approach was developed several years earlier by F. I. Baranov in his classic 1918 paper on the theory of the exploitation of fish stocks (Baranov 1918). Baranov did not contribute directly to Derzhavin's work although Ricker (1971, 1975) reports that Derzhavin's application to the Kura River stellate sturgeon (_Acipenser fulvescens_ Rafinesque) builds upon an earlier work by Tereshchenko (1917) in which Baranov's assistance is acknowledged. Ricker (1971) also points out that a computation similar to Derzhavin's appears in Baranov (1918, p. 100).

The idea behind Derzhavin's method is conceptually very simple. In an age-structured population, the population size of a cohort at the time it enters the exploitable phase can be approximated by simply summing the catches removed from that cohort during the years it contributes to the fishery. Thus summing the catches provides an estimate of the population that must have been alive in order to generate the catches we observe. As input data Derzhavin used estimates of mean age composition and catch data from many years. With his approach, and an assumption of negligible natural mortality and no long term trend in age composition, Derzhavin could estimate the absolute abundance of any age group in any year and a separate exploitation rate for each age group. Derzhavin's analysis provided only a minimum estimate of the total population present at any given time. Ricker (1971, 1975), refers to this quantity as the utilized stock after Voevodin (1938) since it does not include fish that die naturally. The ratio of fish caught to the utilized stock present at the start of the year is referred to by Ricker (1971, 1975) as the biostatistical rate of exploitation. As a consequence of the utilized stock being a minimum estimate, the biostatistical rate of exploitation is always greater than the true rate of exploitation, provided fishing and natural mortality are constant with age.

For many years Derzhavin's method went largely unknown outside of Russia. The first reference to Derzhavin's method in western fisheries literature was in a paper by Bajkov (1932) (often cited as Bajkov 1933) in which Derzhavin's method was applied to populations of whitefish (Coregonus clupeaformis) in three Manitoban lakes systems. Using notation first proposed by DeLury (1947) to state his assumptions, Fry (1949) refined

Derzhavin's method by sampling the age structure of the population annually rather than using averages, but retained the assumption of no natural mortality. Use of annual age composition estimates made possible a more accurate determination of the minimum population abundance. Fry (1949) applied his method to statistics from a sport fishery for lake trout (Salvelinus namaycush) in Lake Opeongo. Ricker (1975, p. 184) points out that, in addition to Fry, several Russian scientists (Boiko 1934; Monastyrsky 1935; Chugunov 1935) independently applied the same modification to Derzhavin's method. Fry called the minimum population estimate calculated by his method the virtual population. He defined the virtual population as "the sum of the fish, belonging to a given year class, present in the water at any given time that are destined to be captured in the fishery in that year and all subsequent years" (Fry 1957).

Soon after the appearance of Fry's paper, Beverton (1954) and later Beverton and Holt (1957) and Paloheimo (1958) proposed age-structured models that emphasized estimation of mortality rates given catch and effort data. The significant feature of these newer models over Derzhavin and Fry's method was that in addition to estimating fishing mortality rates as a product of fishing effort and catchability they also allowed an explicit provision for a nonzero natural mortality rate.

Taking his lead from Beverton and Holt, Murphy (1965) proposed a nonlinear catch ratio model also based on the Baranov catch equation (Baranov 1918). In Murphy's nonlinear sequential model fishing mortality is represented as a fraction of catch to the total stock rather than the product of fishing effort and the catchability coefficient. His method estimates fishing mortality rates and absolute population abundance given catch-at-age data, a known or assumed rate of natural mortality and one starting value of fishing mortality for the youngest fully exploited age group. Because of the nonlinear nature of the equations making up Murphy's model no simple algebraic expression for determining population size or fishing mortality rate could be derived. Rather a sequential computational scheme was used to link successive age groups within a cohort. At each step an iterative procedure was required to solve the equations. In contrast to Derzhavin and Fry, Murphy explicitly incorporated natural mortality in his procedure similar to the Beverton-Holt equations. Also in 1965 Gulland described a slightly different model (Gulland 1965) in which the Baranov catch equation is combined with Fry's concept of virtual population. Gulland used a solution method similar to Murphy's to estimate the parameters of his model, however Gulland used a backward solution to link successive age groups. The procedure was started by providing a starting guess of the fishing mortality rate for the oldest age of a cohort. Gulland called this value the terminal fishing mortality or "terminal F". Sequential computations are usually made at one year intervals but the method is flexible enough to accommodate shorter intervals which is useful for short-lived fishes. Burd and Valdivia (1970) use a monthly interval in their application of the nonlinear sequential model to a series of cohorts of Peruvian anchovies (Engraulis ringens).

Gulland and Murphy were not the first to use the nonlinear sequential model, however they were primarily responsible for popularizing the method. Other earlier papers describe solutions similar to Murphy and Gulland but these are based on different initial assumptions. In an often cited but not widely available report, Ricker (1948) described the mechanics of sequential computation in an application to stocks of halibut. Also Jones (1964) published an application of the nonlinear sequential model to North Sea whiting. Because Gulland and Murphy popularized their methods of stock assessment their papers are cited often. This is especially true of Gulland's 1965 paper which is a mimeographed appendix to an ICES working group report. A description of Gulland's method did not make its way into the formal fisheries literature until it appeared in an appendix to Garrod (1967). Alternative methods to estimate population parameters from catch-at-age and effort data were also proposed at about the same time by Allen (1966, 1969).

During this period convergence properties of the nonlinear sequential model were being investigated. The equations for any cohort could be solved forward in time (from the youngest age to the oldest) after Murphy or backward in time (from the oldest age to the youngest) after Gulland. Jones (1961) was the first to demonstrate the convergence/divergence properties of fishing mortality estimates derived from foreword or backward solutions.

He found that when the sequential procedure began with the oldest age and worked back progressively toward the youngest age, estimates of fishing mortality converged asymptotically to their true values. If the procedure started with the youngest age, fishing mortality estimates for older ages diverged unless the terminal F was very close to the true value. Tomlinson (1970) also confirmed the superiority of "backward" solution. Bishop (1959) and Ricker (1971) investigated other aspects of the parameters of the nonlinear sequential model.

In addition to the sensitivity of the results to the choice of terminal F values, there was one other problem with the backward solution. The method did not give a clear objective picture of the situation in the most recent (current) fishing year. Unfortunately this was exactly the information needed most urgently by the fisheries manager. Since the oldest age in each cohort represents the most recent catch observation, estimates of population size in the current year are only as good as the estimate or guess of the terminal fishing mortality.

The nonlinear sequential model was indisputably helpful and a definite improvement over Derzhavin's or Fry's method in terms of providing a more accurate estimate of absolute population abundance and vital population rates. Accuracy did not come without a cost. The complicated equations of the nonlinear sequential model required a computationally intensive iterative procedure to solve the equations. In the era before computer resources were widely available, solutions were carried out manually by constructing complicated work tables (Schumacher 1970) requiring numerous and tedious hand calculations. Without a computer program to deal with the complicated bookkeeping the procedure proved to be time-consuming and prone to miscalculation error, especially when there were long series of ages. Also the procedure would usually be repeated several times using a range of parameter values. Researchers, even until recently, are exploring which iterative method most efficiently solves the nonlinear equations (Abramson 1971; Doubleday 1975a; Gray 1977, 1979; Miller 1977; Mesnil 1978; Sims 1982a)

Pope (1972) resolved this problem by proposing a less complicated model which greatly simplified the computations needed to calculate a solution to the nonlinear sequential model. Pope's model is based on an approximation to the nonlinear sequential model in which the curve describing exponential decrease in population numbers with time is replaced with a step function. The approximation is based on the assumption that all fish caught in any age group are taken exactly half-way through the year. With Pope's approximation population abundance estimates and vital population rates could be calculated directly from catch data without any need for an iterative procedure, a computer, or reference to virtual populations. As in Gulland's model, Pope's model exhibited the "self correcting" property with respect to fishing mortality estimates. That is, a backward solution causes progressively smaller errors on estimates of fishing mortality and population abundance as the analysis works back along a cohort. As with the models of Gulland and Murphy, the results of cohort analysis are very sensitive to the choice of the terminal fishing mortality and the estimate of natural mortality. Investigations into the effects of systematic and random errors in the sequential computations became easier and much work was done in this area in the years after Pope introduced his method (Agger et al. 1973; Ulltang 1977; Rivard and Doubleday 1979; Sims 1982b, 1984; Rivard 1983; Sampson 1988). Furthermore, the simplicity of Pope's model made the nonlinear sequential model accessible to a wider audience of practicing fisheries biologists. One major problem with the sequential models was that final parameter values were critically dependent on the (often arbitrary) choice of the terminal fishing mortality value. There was no unique solution. In fact there are an infinite number of possible solutions. Pope (1977, table 1) shows how several solutions will satisfy the catch data equally well.

The stock assessment methods described up to this point were completely deterministic. Since the parameters estimated by the models predicted exactly the observed catch data, the models provide no measure of how well the parameters were estimated. In addition cohorts were not linked. That is, catch-at-age data was analyzed one cohort at a time. Parameter values estimated from one cohort were in no way related to other cohorts in the population. This situation changed with the introduction of the separability assumption. In contrast to the nonlinear sequential model where fishing mortality is represented as a fraction of catch to the total stock, the more restrictive separable model represents fishing mortality as the product of an age-specific and a year-specific coefficient. This is similar to Beverton-Holt's representation of fishing mortality except that the meaning behind the coefficients in the separable formulation are somewhat more general.

The idea behind the separability assumption is that in any one year fishing mortality can be described by two factors, a full-recruitment fishing mortality or exploitation pattern and a factor to account for the differential effect of the annual exploitation pattern on various age groups in the stock. The origin of the separability assumption is difficult to pin down. Doubleday (1976) is usually credited with introducing this concept however evidence of earlier thinking on this topic can be demonstrated. On closer examination we find that Doubleday references an earlier unpublished manuscript by Pope (1974) in which the separability assumption is first proposed. Pope's manuscript, which appeared in the literature a year after Doubleday's paper (Pope 1977), refers to an even earlier manuscript by Agger et al. (1971) and credits these researchers as the first to cast the stock assessment model into an optimization framework based on a separable fishing mortality assumption. Casting the stock assessment model into a framework of minimizing the squared difference between observed and predicted catch observations (Agger et al. 1971) dramatically reduced the number of parameters in the model, provided a means by which all model parameters could be statistically estimated in a simultaneous manner rather than sequentially, and provided a means of simultaneously linking data from several cohorts. Introduction of the separable formulation of fishing mortality was an important conceptual advance because it moved the study of stock assessment methodology into the realm of more generalized mathematical models and went a long way towards promoting statistical analysis of catch-at-age data.

Since Doubleday's paper, the separability assumption has become almost a standard feature of newer models (Paloheimo 1980; Pope and Shepherd 1982; Fournier and Archibald 1982; Dupont 1983; Deriso et al. 1985). Despite the advantages of the separability assumption, one problem that carried over from sequential models was the unfortunate fact that catch-at-age data alone do not contain enough information to estimate fishing mortality in the most recent fishing year with acceptable precision (Doubleday 1976; Pope 1977; Pope and Shepherd 1982). Also solutions to a separable model using catch-at-age data alone are not unique because stock abundance and fishing mortality parameters are highly negatively correlated. In order to overcome the difficulties associated with trying to simultaneously estimate stock size and fishing mortality from catch-at-age data, recent efforts have concentrated on developing methods to "calibrate" results from stock assessment models with effort data or some other fishery-independent data (Doubleday 1981; Collie and Sissenwine 1982; Laurec and Shepherd 1983; Mohn 1983; Pope and Shepherd 1984; Lewy

1985; Pope and Shepherd 1985). This approach is often referred to as "tuning VPAs" or sometimes as "integrated analysis". Pope and Shepherd (1985) present a comparison of the performance of ten different tuning methods. Fournier and Archibald (1982) and more recently Deriso et al. (1985) have proposed very generalized mathematical models using the separability assumption that allow incorporation of fishery-independent data directly into the simultaneous parameter estimation procedure. These later two methods will be described and compared in later sections.

1.2 Etymology and Naming Conventions

Up to this point in the discussion, reference to any particular method of stock assessment by commonly used names (i.e. VPA, cohort analysis etc.) has deliberately been avoided. The reason for this tactic stems from a lack of established conventions for associating a name or label with a particular method of stock assessment. In fact the quick proliferation of methods in recent years has served to exacerbate an already high level of confusion regarding the naming convention, which method is which, and how one method relates to another. This is unfortunate, especially if one is trying to determine exactly which method was used to generate results presented in a scientific report. Even though the various methods share many similarities, each is distinct with respect to the mathematical model, underlying assumptions, solution algorithm and what parameters the model estimates. In order to effectively discuss each method a means of referring to them in an unambiguous manner is needed. This section lays out the origins of the various naming conventions and proposes, at least for the purposes of this paper, an operational scheme for referring to each method.

Derzhavin did not name his method although Ricker (1971) has called it "Derzhavin's Biostatistical Method". Ricker's use of the term biostatistical method is based on the original use of this term to describe an analysis which combines catch statistics with other biological information such as age or size composition.

Fry called his method "Virtual Population Analysis" (VPA). Fry's choice of the title was based on the analogy with the virtual image of the physicist - "... although it is not the real population it is the only one that is seen." (Fry 1957). In physics, "virtual" quantities were used frequently in the analysis of physical systems and while unmeasurable themselves, they were tools that could predict the behavior of real objects.

Murphy did not give his model a name but it sometimes has been referred to as "Murphy's Equation" (Tomlinson 1970). I will retain this naming convention.

Gulland also did not name his model but it is often incorrectly referred to as VPA and/or cohort analysis. In order to distinguish between Gulland's model and these other two approaches I will refer to Gulland's model as the "Sequential Population Assessment" or "Sequential Population Analysis" (SPA) model.

In this country the name VPA has commonly been used as more of a generic descriptive title rather than a reference to one specific method of stock assessment. This is especially true for the SPA model, apparently because Gulland (1965) demonstrated that his formulation could be based on a table of virtual populations in the sense of Fry (1949) (i.e. the sum of the fish present in the population that would ultimately appear in the catch). Gulland recognized the confusion over the naming convention. In an attempt to clarify the distinction between his method and Fry's Gulland (1977, p. 87) wrote

> "This method of cohort analysis -- often misleadingly referred to as 'virtual population analysis' in North American studies, from the original derivation of Gulland, although the method is quite distinct from Fry's usage of virtual population -- is widely used in the North Atlantic".

Pope called his approximation to the SPA model "Cohort Analysis" (CA). Ricker (1975, p. 194) claims that the name cohort analysis had earlier been applied to SPA models but does not provide any references.

In recent history the two names VPA and cohort analysis are used more or less interchangeably in the fisheries literature even though the two names may refer to any one of three different but similar methods of stock assessment; VPA, SPA, and CA. The claim that confusion exists in the naming convention can best be demonstrated by examining how the two names are used in the fisheries literature. Often both names are used at the same time, apparently to avoid confusion (Ulltang 1977; Sims 1982b, 1984). Other times one name is preferred over the other (Aldenberg 1975; Garrod 1976), or the names are intentionally avoided altogether (Ricker 1971).

Because this paper deals with several stock assessment methods the following naming conventions will be adhered to in subsequent discussion. Age-structured stock assessment or age-structured stock analysis (ASA) techniques will be used as a generic name to refer to any analytic tool used by fisheries managers to estimate fishing mortality rates and absolute population abundance of commercially exploited fish stocks given catch-at-age data and possibly data independent of the commercial fishery. Naming conventions for a particular method of ASA will be as follows: as originally proposed by Ricker (1971, 1975) "Derzhavin's Biostatistical Method" (DBM) refers to the method of Derzhavin (1922); "VPA" (VPA) refers to Fry's modification of Derzhavin's method (Fry 1949, 1957); "Murphy's Equation" (ME) refers to the catch ratio model using a forward solution (Murphy 1965); "Sequential Population Assessment" (SPA) model refers to Gulland's model using a backward solution (Gulland 1965; Garrod 1967); "Cohort Analysis" (CA) refers to Pope's approximation to SPA (Pope 1972); "Doubleday's Model" (DO) refers to the separable model of Doubleday (1976); "Paloheimo's Model" (PLO) refers to the separable log CPUE model of Paloheimo (1980); "Pope and Shepherd's Model" (PS) refers to the separable model of Pope and Shepherd (1982); "Fournier and Archibald's Model" (FA) refers to the separable model of Fournier and Archibald (1982); "Dupont's Model" (DU) refers to the separable model of Dupont (1983); and "CAGEAN"

(CAGEAN) refers to the "separable Catch-AGE-ANalysis" model of Deriso et al. (1985).

1.3 The Need for ASA Methods

Early fisheries management relied primarily on theoretical advances made in the 1940's and 1950's by Ricker and Beverton and Holt. Management regulations developed from these advances in fishery science were based on the assumption that catch per unit of effort (CPUE) could be used as an index of relative abundance in the assessment of total mortality. Management policies such as mesh size regulations worked well during the era when fishing fleets were relatively unchanging with respect to their design, fishing patterns and efficiency. However, rapid changes in fishing technology, increases in fishing effort in the late 1960's and 1970's and declines in the production of numerous stocks caused the validity of an assumed relationship between CPUE and stock abundance to be questioned. In multispecies fisheries it was even more difficult to maintain a time series of consistent CPUE estimates because of the inability to separate directed effort from total effort or interannual changes in the availability of the target species.

As a result of these problems, concern began to be expressed as to the effectiveness of the regulatory mechanisms mentioned earlier. Because of shifts toward quota management the need developed to describe stock numbers in absolute numbers rather than by a relative index which had variable calibration between stock areas. Furthermore, in fisheries where partially recruited age classes contributed a significant part to the overall catch, estimates of fishing mortality on these groups critically needed to be included in management regulations.

In summary, problems associated with CPUE-based fisheries management models directly contributed to the development and application of theory to estimate fishing mortality and population abundance without reliance on CPUE.

2.0 BACKGROUND

2.1 Data Sources

In describing characteristics of fisheries data typically submitted to age-structured stock assessment methods, a hypothetical example will be used so that concrete examples can be presented.

In most situations catch-at-age data are not directly collected from the fishery. Rather these data are generated by combining some aggregate measure of total commercial catch with information derived from biological sampling. Total catch is usually tabulated in terms of either total biomass or total numbers. Biological information is collected by sampling a small portion of the overall catch. Data usually consists of length and weight measurements. Also a scale, otolith or fin ray is collected in the biological sample. These structures are used to determine an age composition estimate.

The usual data collection process proceeds as follows. In each year, an estimate of the total catch of fish taken by the fishery, $O(y)$, is tabulated. Part of the total catch is randomly sampled and aged. From the random sample, a fraction of the catch sample, $H(a,y)$, is observed to be of age a in year y. Let $C(y)$ be the true catch in year y and $G(a,y)$ be the true age composition. The true catch-at-age can be calculated from the product $[C(y)\ G(a,y)]$ but since $C(y)$ and $G(a,y)$ are unknown we can only estimate the catch-at-age by $C(a,y)=O(y)\ H(a,y)$ if catch is recorded in numbers, or $C(a,y)=[O(y)\ H(a,y)]/W(a)$ if the catch is recorded in biomass.

Information other than catch-at-age data can be incorporated into an age-structured stock assessment model. Probably the most common additional information available from the fishery is annual effort data. One value is collected each year and represents cumulative effort expended over the fishing season. Sometimes effort data can be stratified by gear types or vessel classes. If differences in gear/vessel efficiency exist, effort standardization will be required to convert nominal effort to effective effort. Other data sources independent of the commercial fishery that may provide information on stock dynamics include 1) estimates of population abundance such as CPUE indicies, 2) population biomass estimates from acoustic or research surveys and 3) information from biological sampling such as fecundity-at-age data and weight-at-age data.

2.1.1 Description of Catch-at-age Data. Table 1 shows a catch-at-age data set from a hypothetical fishery which operated for years 1977 through 1981. Shown in the borders of

Table 1 are absolute age (left) and year (top) and relative age (right) and year (bottom) indicies. Each entry in the catch-at-age matrix, C(a,y), represents the catch of one age group (a) in one year (y). Each row in the matrix represents the contribution of one age group from different cohorts to the annual catch over a series of years. The first exploited age (r) is age two and the oldest exploited age (s) is age nine. The total number of age groups vulnerable to the fishery (A) is eight (A=s-r+1). Each column in the matrix represents the aged catch from one fishing year. The number of years the fishery was prosecuted (Y) is five (Y=ly-fy+1). The total number of catch observations (n) collected over the 1977-1981 interval is forty (n=A*Y).

Several cohorts are represented in the catch-at-age matrix. Cohorts consist of individuals in the population that share the same birth year. Each diagonal in the catch-at-age matrix, moving from upper left to lower right, represents one cohort or year class. For example, the catch observations C(2,78), C(3,79), C(4,80) and C(5,81) are all from the 1976 year-class. The total number of cohorts in the matrix (K) is twelve (K=A+Y-1) representing year-classes 1968 through 1979. Each column in the catch-at-age matrix represents members of the population that share the same death year. Table 2 shows the hypothetical catch-at-age data organized by cohort. Note that each member of the population can be assigned to any one of k distinct cohorts. Also, each cohort consists of a different number of catch observations. For example, two cohorts consist of only one catch observation; the catch of nine year-olds in 1977 is the only catch observation from the 1968 year class and the catch of two year-olds in 1981 is the only catch observation from the 1979 year-class. Also the 1972 through 1975 year-classes each are made up of five catch observations. Note also that each cohort also has a unique combination of starting/ending age and year indicies.

The cohorts making up the catch-at-age matrix can be further classified as to whether they are completely fished out or not. A completely fished out cohort means that, relative to the current fishing year, no individual from that cohort will contribute to the catch in the next fishing year. For example, the 1970 year-class contributed to the fishery as seven year-olds in 1977, eight year-olds in 1978, and nine year-olds in 1979. In 1980 the 1970 year-class ceased to contribute to the catch, thus it was completely fished out in 1979. Of the eight cohorts contributing to the catch in 1981, the 1972 can be considered completely fished out while individuals from year-classes 1973 through 1979 will appear in the 1982 catch.

2.2 Basic Governing Equations of Age-structured Populations

In age-structured models events at the population level are studied by analyzing the fate of individual fish or similar age groups. This is in contrast to stock production models (Graham 1935, 1939; Schaefer 1954, 1957) where the stock is treated as a single entity. Age-structured models explicitly consider the separate biological processes that alter fish population abundance. These include mechanisms to describe the processes of reproduction, mortality losses due to natural causes and from man's fishing activity. Usually each process is monitored independently for each age group by treating each process as a stand-alone submodel.

ASA models do not provide an explicit means of modeling the birth rate although there are some exceptions which will discussed later. Most ASA models estimate the size of a year class at the time when members of the cohort first appear in the catch. This is often referred to as recruitment to the fishery. Using this approach avoids problems associated with estimating birth and juvenile mortality rates.

When describing the significant mortality factors impacting fish populations, the usual convention is to assume the total mortality rate (Z) is the sum of a natural mortality rate (M) and a rate due to fishing (F)

$$Z = F + M \qquad [1]$$

The natural mortality component of total mortality is the most difficult to quantify. Usually this mortality component is a "catch-all" mortality that includes all sources of mortality not related to directed fishing activity. Uncertainty regarding the exact relationship of natural mortality to an animals age, nutritional status, size etc. is usually dealt with by simply assuming that death from natural causes is constant for all individuals in the population. When fishing mortality is considered in the aggregate, it does not include any random components and does not vary by age or year.

These assumptions are clearly not biologically reasonable since from an intuitive standpoint mortality should vary with age because of age-specific variations in predation pressures, disease, and iter/intra-specific competition. Unfortunately data are rarely available to further refine the natural mortality assumption and as a first approximation a constant value is usually employed. The so-called separability assumption (Agger et al. 1971; Doubleday 1976; Pope 1977) permits a more realistic formulation of fishing mortality by providing an explicit accounting of factors that may cause F to vary by age and/or year. Fournier and Archibald (1982) provide a insightful justification for the separable fishing mortality formulation. They suggest that the level of fishing mortality depends on an interplay between fishermen and fish. No matter how many fish there are, the fishing mortality will be zero if fishermen do not fish. If fish are unavailable to the fishery or abundance is low, fishing mortality will also be zero regardless of what the fishermen do.

The separability assumption is usually written as

$$F(a,y) = s(a)\ f(y) \qquad [2]$$

which expresses fishing mortality as a product of two quantities, one representing an age-specific factor and one a year-specific factor. Equation [2] can also be written in a logarithmic form as

$$\ln[F(a,y)] = \ln[s(a)] + \ln[f(y)]$$
$$= v(a) + e(y) \qquad [3]$$

which expresses fishing mortality as a sum of two quantities rather than as a product. A nonlinear equation for fishing mortality can then be written as

$$F(a,y) = \exp\left[v(a) + e(y)\right] \qquad [4]$$
$$= \exp\left[\ln[s(a)] + \ln[f(y)]\right] \qquad [5]$$

The separability formulation has an inherent problem. There is an indeterminacy in the model since s(a) and f(y) affect the predicted catches through their sum (v(a) + e(y)). For example v(a) could increase and e(y) decrease by a constant without changing the value of ln[F(a,y)]. This will be discussed later.

The age-specific factor in the separability formulation is referred to as availability (Doubleday 1976), partial recruitment (Pope 1977), exploitation pattern (Pope and Shepherd 1982), relative level of fishing vulnerability (Fournier and Archibald 1982) or the selectivity coefficient (Deriso et al. 1985). The year-specific factor is referred to as the effective effort multiplier (Doubleday 1976), fully recruited fishing mortality (Pope 1977; Deriso et al. 1985) or fully exploited fishing mortality (Pope and Shepherd 1982).

In comparison to [1] a more realistic representation of total mortality would include a separable fishing mortality and a constant natural mortality rate factor. Total mortality can then be written

$$Z(a,y) = F(a,y) + M \qquad [6]$$

2.2.1 Exponential Survival Model. The exponential survival model has a distinguished pedigree, dating back to the mid 17th century. Indeed one of the first references to this equation was in a very early paper by Euler in 1740. Much of the mathematical theory on contemporary age-structured survival models was developed by A. J. Lotka (1925). Lotka's original formulation was a continuous-time model developed for application to human populations. In fisheries applications, however, the discrete-time version of the model is used since in fish populations generations are separate and the aging process is usually considered to take place in discrete steps of one year. The tracking of the population in discrete age intervals precludes the need to account for growth if average weight-at-age data is known.

The exponential survival model describes the change in numbers that take place in the population due to natural and fishing mortality factors. The theory behind the model (see Ricker 1975) results in the familiar recursive relationship

$$N(a+1,y+1) = N(a,y) \exp[-F(a,y)-M] \qquad [7]$$

which describes survival for two successive ages in a cohort. Equation [7] can be generalized to follow any one cohort over all years it is vulnerable to the fishery by explicitly providing indexes for age (a), year (y), and past age groups (if any) over which the cohort was exploited (i). Also it is assumed that fish recruit to the fishery at a fixed age (r). This is often referred to as "knife-edged" recruitment (Deriso 1980). The relationship between the time indicies indicate that an age a fish, alive in year y, was born in year y-a and recruited to the fishery at age r in year y-a+r. If R represents the original number of individuals making up a cohort when the cohort first entered the exploitable phase (i.e. $a \geq r$) then equation [7] can be written

$$N(a,y) = R(r,y-a+r) \exp\left[-\sum_{i=r}^{a-1} [F(i,y-a+i)+M] \right] \qquad [8]$$

The total number of individuals in the population in any year is the sum of all age groups

$$N(y) = \sum_{a=r}^{s} N(a,y) \qquad [9]$$

and the total biomass of the population is

$$B(y) = \sum_{a=r}^{s} N(a,y)\ W(a) \qquad [10]$$

2.2.2 Catch Equation. The catch equation of Baranov (1918) is basic in most approaches to solving fish population dynamics problems. The equation describes the relationship between the rate at which fish are caught and the numbers alive in the catchable population. The catch equation combines a differential equation model of the catch process and the exponential survival model. Once again the theory behind the catch equation is described in many population dynamics texts (Ricker 1975; Gulland 1977; Seber 1982). The model is usually written

$$C(a,y) = \frac{F(a,y)}{F(a,y)+M} N(a,y) \left[1-\exp[-F(a,y)-M]\right] \qquad [11]$$

The total catch in any year is the sum of the catch for all age groups in the population

$$C(y) = \sum_{a=r}^{s} C(a,y) \qquad [12]$$

and the total catch biomass is given by

$$CB(y) = \sum_{a=r}^{s} C(a,y)\ W(a) \qquad [13]$$

2.2.3 Availability. The term availability, first introduced by Widrig (1954), refers to the accessibility of the fish in the population to the fishing gear. In the fisheries literature availability is often confused or used interchangeably with the terms selectivity and/or catchability. Sometimes the term vulnerability is used instead of catchability. These terms represent three important concepts in fisheries science. Because the availability concept is being applied in some newer models they will be explained here.

Fournier and Doonan (1987) define availability as the proportion of individuals in an age group with a positive probability of being caught. Availability can alternatively be defined as that proportion of the unit stock present on the fishing grounds. Normally we assume that all fish are equally available to capture (i.e. availability=1). Conceivably a situation could arise where a cohort could consist of two components, one segment that was available to the fishery and a second segment that was unavailable. This separation could arise because of migratory behavior or because part of the cohort resides in habitats not reached by the fishing gear. The concept of selectivity refers to idea that separate age groups experience differential mortality once they come within the influence of the fishing gear. Selectivity differences arise because of age/size related behavioral characteristics of the exploited stock. They are a result of partial recruitment of a year class into the fishery and selectivity of the gear. The concept of catchability (vulnerability) relates to the probability that, once available, a fish will be caught by one unit of effort.

Equation [7] can easily be modified to incorporate an availability feature if this is appropriate to the fishery being analyzed. If we assume a separable fishing mortality and we let P represent the availability (assume P is constant for all age groups) then [7] can be modified to

$$N(a+1,y+1) = \Big[N(a,y)\ P\ \exp[-F(a,y)-M]\Big] + \Big[N(a,y)\ (1-P)\ \exp(-M)\Big] \qquad [14]$$

The first term on the right hand side represents the fished segment and the second term the unfished segment. Further, [11] can also be modified to give

$$C(a,y) = \frac{P\ F(a,y)}{F(a,y)+M}\ N(a,y)\ \{1-\exp[-F(a,y)-M]\} \qquad [15]$$

Most ASA models assume that the fished stock is homogeneous (i.e. availability for all age groups is assumed to be 1), however there are exceptions. Murphy (1965) was perhaps one of the first researchers to propose an ASA model that incorporated an availability feature. More recently, MacCall (1986) uses [14] and [15] as a starting point from which to derive a useful approximation similar to Pope's cohort analysis. Also Fournier and Doonan (1987) propose a model that incorporates age-dependent selectivity, age-dependent availability and catchability.

2.2.5 ASA Structural Models. Structural models refer to the exact mathematical expressions that result from combining equations [7] and [11] in various ways. The structural models used in the various ASA methods are the SPA model, log catch model, catch ratio model, log catch ratio model, and the log CPUE model. Each structural model will be discussed in more detail as the ASA methods described later. A summary of which model underlies each ASA model is provided in Table 7.

3.0 DESCRIPTION OF ASA METHODS - NO EFFORT DATA REQUIRED

3.1 Derzhavin's Biostatistical Method

The model of Derzhavin (1922) uses annual catch values and average age composition values to estimate the maximum exploitation rate and also the minimum population abundance.

3.1.1 The Model. Derzhavin's model describes the population at the start of a reference year as the sum of the catch in the reference year plus the catch in the following year diminished by the number of fish that were not in the population in the reference year. The summation proceeds until either the end of the catch time series is reached or the maximum number of age groups are accounted for. Derzhavin's model assumes that natural mortality is insignificant (i.e. $M=0$). If this assumption is true, then the catch must estimate the total population. If in reality $M>0$ but we assume that $M=0$, then we have a minimum population estimate. The expression for the minimum population at the start of any reference year (cf. Derzhavin, p. 15) is

$$N(y) = [1-H(0,y)]\, O(y) + [1-H(0,y)-H(1,y)]\, O(y+1) + [1-H(0,y)-H(1,y)-H(2,y)]\, O(y+2) + [1-H(0,y)-H(1,y)-H(2,y)-\ldots-H(s-1,y)]\, O(y+s-1) \qquad [16]$$

where $H(a,y)$ is the fraction of age a fish in the catch of year y (i.e. the estimated age composition) and $O(y)$ is the total estimated catch (numbers) in year y. These values are assumed to be constant for each age from year to year so that $H(a,y)=H(a,y+1)=H(a,y+2)$ etc. Also $H(0,y)$, the percentage of young of the year in the catch, is assumed to be equal to 0. An expression similar to [16] can be found in Baranov (1918, p. 100).

3.1.2 Objective Function. This method does not have an objective function.

3.1.3 Parameter Estimation Procedure. Minimum population estimates are calculated directly from [16]. The rate of fishing mortality, called the biostatistical rate of exploitation (Ricker 1971, 1975), is calculated as the ratio of the catch of fish of a given age in a given year to the utilized stock of that age at the start of the year. This can be calculated for the whole stock or for individual ages separately. Since the population estimate is a minimum, the calculated biostatistical rate of exploitation will always be greater than the "true" rate of exploitation.

Over one unit of time, [11] can be rearranged (ignoring indexes) to get

$$\frac{C}{N} = 1 - \exp(-F) \qquad [17]$$

since all mortality is attributed to fishing (i.e. M=0). From [17] the maximum rate of fishing mortality can be estimated from

$$F = \ln[1 - \frac{C}{N}] \qquad [18]$$

3.2 Virtual Population Analysis

The VPA model (Fry 1949, 1957) uses annual catch values and annual age composition values to estimate the maximum exploitation rate and the minimum absolute population abundance.

3.2.1 The Model. Fry's VPA model is almost identical to Derzhavin's. The only difference is that instead using one average age composition estimate for all years, the VPA model uses a separate age composition estimate for each year. The catch-at-age in any year is given by $C(a,y)=O(y)H(a,y)$, where $H(a,y)$ is not equal to $H(a,y+1)$. The virtual population for any age and year is calculated as (cf. Fry 1949, equation 1 in appendix, p. 66) is

$$V(a,y) = \sum_{i=1}^{\min(A,Y)} C(a+i-1,y+i-1) \qquad [19]$$

3.2.2 Objective Function. This method does not have an objective function.

3.2.3 Parameter Estimation Procedure. Estimates of the virtual population and the maximum rate of exploitation proceed as described in Section 3.1.3.

3.3 Gulland's SPA Model

Input data required by the model of Gulland (1965) (also see Garrod 1967) and estimated parameters are summarized in Table 8. Gulland used a backward solution to the nonlinear sequential equations (tabulating parameter estimates from the oldest age group back in time to the youngest recruiting age group in the cohort). Also the SPA model can be expressed in terms of the virtual population.

3.3.1 The Model. In the SPA model the catch equation of Baranov (1918) [11] and the exponential survival model [7] are combined together without reference to virtual populations to give

$$\frac{N(a+1,y+1)}{C(a,y)} = \frac{[F(a,y)+M]\ \exp[-F(a,y)-M]}{F(a,y)\ \{1-\exp[-F(a,y)-M]\}} \qquad [20]$$

For any cohort, Gulland's SPA model expresses the ratio of population abundance to catch as a nonlinear function of fishing mortality. There is an important relationship in time between C(a,y) and N(a+1,y+1) in that the ratio of these two quantities references the same point in time. C(a,y) represents those fish, age a, caught up until the end of the year y while the stock of fish of age a+1, N(a+1,y+1), represents the population exploited from the beginning of the year y+1. Note that this is the same population left from the end of previous year. Gulland (1965) defines this ratio as the catch during the year expressed as a proportion of the population at the end of year.

For each cohort, the total number of ages for which catch data are available, A, yields A equations similar to [20] that contain A+2 unknown parameters. As stated, the model has too many parameters. Gulland's solution to this problem was to provide an estimate of M and a terminal fishing mortality value. Providing values for two parameters allows explicit but non-unique solutions to the linked system of equations to be calculated.

3.3.2 Objective Function. Because equation [20] is nonlinear in F(a,y) an iterative procedure is required to solve the system of linked equations. First substitute $Z(a,y)=[F(a,y)+M]$ and let $S(a,y)=\exp[-Z(a,y)]$. Then in Gulland's model, the objective function to be minimized at each age step in the iterative procedure can be written

$$f[F(a,y)] = \frac{Z(a,y)\ S(a,y)}{F(a,y)\ [1-S(a,y)]} - \frac{N(a+1,y+1)}{C(a,y)} = 0 \qquad [21]$$

3.3.3 Parameter Estimation Procedure. The Newton-Raphson method is one of several methods that can be used to solve [21] (Doubleday 1975a; Miller 1977; Mesnil 1978; Sims 1982a). By iteratively correcting a trial value of F(a,y), the Newton-Raphson method can provide an value of F(a,y) that causes $f[F(a,y)]=0$. In the first iteration, the procedure usually starts out with a trial value of fishing mortality $F(a,y)_{(1)}$. The derivative of equation [21] with respect to F(a,y) is

$$\frac{d\, f[F(a,y)]}{d\, F(a,y)} = f'[F(a,y)] =$$

$$\frac{S(a,y)}{F(a,y)\ [1-S(a,y)]} \left[1 - \frac{Z(a,y)}{F(a,y)} - Z(a,y) - \frac{Z(a,y)\ S(a,y)}{[1-S(a,y)]} \right] \quad [22]$$

Equation [22] is evaluated at $F(a,y)=F(a,y)_{(1)}$ and estimates of F(a,y) are updated for the next iteration using the equation

$$F(a,y)_{(i+1)} = F(a,y)_{(i)} - \frac{f[F(a,y)]_{(i)}}{f'[F(a,y)]_{(i)}} \quad [23]$$

Iterations continue until $|F(a,y)_{(i+1)} - F(a,y)_{(i)}| \leq$ some stopping criteria.

If C(a,y), F(a,y), and M are known, then [7] and [11] can be manipulated together (iterated) in a backwards or hindcast mode to yield estimates of N(a,y) and F(a,y) for all past years of life of the cohort. Parameters are estimated separately at each stage of the procedure by applying a root finding algorithm to equation [21]. The procedure begins by supplying an estimate of terminal fishing mortality for the oldest age in a cohort, F(amax,y), and observed catch for the oldest age in the cohort, C(amax,y). Then [11] is used to solve for the terminal population, N(amax,y). Since C(amax-1,y-1), N(amax-1,y-1) and M are known, [20] is used to solve for F(amax-1,y-1). Finally [7] can be used to estimate N(amax-1,y-1) and the procedure is repeated for the next youngest age group until the youngest age in the cohort is done. A key point to remember about the solution procedure to this model is that final parameter values derived from the system of linked equations are not unique. If one starts with a different terminal fishing mortality value, one can obtain a totally different solution.

The mechanics of sequential computation of these two equations was described by Ricker (1948) and the method was popularized by Murphy (1965) and Gulland (1965). When estimating the terminal population, equation [11] can take two forms depending on whether the cohort is completely fished out or not. If the cohort is not completely fished out (i.e. there are survivors from the cohort that will show up in the catch one year older in the following year) then [11] holds. If the cohort is completely fished out (i.e. there are no survivors or the fish has passed beyond the exploited phase), then [11] is modified to

$$C(a,y) = \frac{F(a,y)}{F(a,y)+M} N(a,y) \qquad [24]$$

In the procedure described above, calculations are carried out without reference to Fry's virtual population. However in Gulland's original description of his method (Gulland 1965; Garrod 1967) he showed that the solution to equation [20] (cf. Garrod 1967, appendix B, equation A) could be related to the virtual population. Gulland showed that the ratio of the population at the end of the year to the catch during the year $(N(a+1,y+1)/C(a,y))$ could be expressed as a fraction of the apparent survival during the year (as estimated from virtual populations) and the exploitation rate applicable to the fish alive at the end of the year. In addition to the above references, Gulland's correction to the VPA is described in Cushing (1975, pg. 149-150).

Much work has been done on the behavior of this method under failure of the assumptions: 1) Results of the analysis are insensitive to errors in the estimated or assumed value of terminal F, but only when fishing accounts for about 50% or more of the total deaths (i.e. the ratio F/Z is in the range 0.5-1.0) (Jones 1981) or cumulative F over the life of a cohort is greater than Z (Pope 1972); 2) Errors in the estimates of N(a,y) and F(a,y) caused by random fluctuations in M (when M assumed constant) are likely to be small (6%) when M fluctuates moderately (Ulltang 1977; Pope 1979b), although this would tend to be more severe on older animals since they occur in relatively smaller numbers. Agger et al. (1973) considered the effect of uncertainty in the value of a presumed constant natural mortality rate. They estimated that the bias in F would be 25% if M is known with a mean error of 0.1; 3) Results are relatively insensitive to seasonal trends in M and F (Ulltang 1977); and 4) Effects of unevenly distributed catches (i.e. the intra-year frequency distribution of catches is not constant) on the relative error in estimates of N(a,y) are not severe (< 10%) unless the majority of the catch is taken at the beginning or end of the year and M is large (0.6) and/or F is high (0.6) and (Sims 1982b). These results also apply to Murphy's Equation and Pope's Cohort Analysis which are described below.

3.4 Murphy's Equation

Murphy's Equation (Murphy 1965) is similar to Gulland's SPA model. Input data required by the model and estimated parameters are summarized in Table 8. Murphy used a foreword solution to the nonlinear equations (tabulating parameter estimates from the youngest age group foreword in time to the oldest recruiting age group in the cohort).

3.4.1 The Model. Murphy's catch-ratio model consists of taking equation [11] and creating a ratio of catches from the same cohort in two successive seasons. The catch ratio model is density-independent thus it is only a function of mortalities. Since the model represents a ratio of catches, the ratio of original population abundances in succeeding years can be written

$$\frac{N(a',y')}{N(a,y)} = \frac{R(r,y'-a'+r)\ \exp\left[-\sum_{i=r}^{a'-1}[F(i,y'-a'+i)+M]\right]}{R(r,y-a+r)\ \exp\left[-\sum_{i=r}^{a-1}[F(i,y-a+i)+M]\right]} \qquad [25]$$

$$= \exp[-Z(a,y)]$$

where $y'=y+1$ and $a'=a+1$.

If we define (for ease of exposition)

$$U = \exp[-F(a,y)-M] = \exp[-Z(a,y)] \text{ and}$$

$$V = \exp[-F(a+1,y+1)-M] = \exp[-Z(a+1,y+1)]$$

then Murphy's equation (cf. Murphy 1965, equation 5) can be completely expressed as a function of the mortalities in year y and year y+1. It is written

$$\frac{C(a+1,y+1)}{C(a,y)} = \frac{U\ F(a+1,y+1)\ Z(a,y)\ (1-V)}{F(a,y)\ Z(a+1,y+1)\ (1-U)} \qquad [26]$$

Note that the catch ratio can also by calculated from the ratio C(a,y)/C(a+1,y+1).

For each cohort, the total number of ages for which catch data are available, A, yields A-1 equations similar to [26] that contain A+1 unknown parameters. Explicit solutions to the linked system of equations can be derived given the catch ratio, an estimate of the natural mortality rate, and an estimate of a terminal F. Once these values are provided, estimates of F(a+1,y+1) can be calculated by an iterative method.

3.4.2 Objective Function. Since equation [26] is nonlinear in F, an iterative procedure is required to solve the system of linked equations. The objective function to be minimized at each age step can be written

$$f[F(a,+1,y+1)] = \frac{U\ F(a+1,y+1)\ Z(a,y)\ (1-V)}{F(a,y)\ Z(a+1,y+1)\ (1-U)} - \frac{C(a+1,y+1)}{C(a,y)} = 0 \qquad [27]$$

3.4.3 Parameter Estimation Procedure. The parameter estimation procedure employed in Murphy's method is essentially the same as that used in Gulland's model except for slight differences in the manipulative procedures used to reach a solution. Where Gulland assumes M is known and guesses F(amax,y), Murphy assumes M is known and guesses F(amin,y). Once the first terminal F is provided, each method proceeds to estimate the remaining F's and then the population size in a similar manner. At each age, values of F are corrected by an iterative procedure until the objective function is equal to zero. Gulland's model uses a backward solution (working from the oldest age in the cohort towards the youngest age) while Murphy uses a forward solution (working from the youngest age to the oldest).

The derivative of equation [27] with respect to F(a+1,y+1) (substituting Z=F+M) is

$$\frac{d\ f[F(a+1,y+1)]}{d\ F(a+1,y+1)} = f'[F(a+1,y+1)] =$$

$$\frac{U\ Z(a,y)\ (1-V)}{Z(a+1,y+1)\ F(a,y)\ (1-V)} \left[1 - F(a+1,y+1)\ V - \frac{F(a+1,y+1)}{Z(a+1,y+1)} \right] \quad [28]$$

As in Gulland's SPA model, equation [11] can take two forms depending on if the oldest age in the cohort is completely fished out or not. See equation [24] in Section 3.3.3.

3.5 Cohort Analysis

The cohort analysis model (Pope 1972) is similar to Murphy's Equation and Gulland's SPA model except that Pope introduced an approximation for the exponential survival model. Input data required by the model and estimated parameters are summarized in Table 8.

3.5.1 The Model. Pope (1972) simplified the SPA model by introducing a discrete approximation to the continuous exponential survival model [7]

$$\exp(M/2) = \frac{[F(a,y)+M]\ \{1-\exp[-F(a,y)]\}}{F(a,y)\ \{1-\exp[-F(a,y)-M]\}} \quad [29]$$

Agger et al. (1973) point out that Pope (1971) based cohort analysis on the approximation

$$\frac{\sinh[F/2]}{\sinh[(F+M)/2]} \approx \frac{F}{F+M} \qquad [30]$$

Pope's approximation assumes that the entire catch is taken exactly midway through the year. Pope (1972) shows the approximation is usable at least up to values of M=0.3 and F=1.2. As pointed out by Jones (1981), these mortality limits actually relate to Mdt and Fdt, where dt is the length of the time interval used. Normally dt is one year. If M and F are larger than the limits specified by Pope, cohort analysis is still usable if the catch-at-age data is divided into intervals of less than one year.

Recently MacCall (1986) shows that

$$\frac{M}{1-\exp(-M)} \qquad [31]$$

is a slightly better approximation compared to exp(M/2). MacCall's new approximation becomes better as M get larger, thus extending the range of mortality rates under which the approximation proves useful. MacCall's approximation is also better when the assumption that all the catch is taken half way through the year (i.e. there is seasonality in the fisheries) is not valid.

Siddeek (1982) identified an error in equation 2.4 of Pope's 1972 paper and presents a correction. The error, acknowledged in Pope (1982), deals with the expression describing the error introduced into cohort analysis due to an incorrect choice of terminal fishing mortality.

3.5.2 Objective Function. This method does not have an objective function.

3.5.3 Parameter Estimation Procedure. With Pope's cohort analysis catch data are literally transformed into population abundance estimates. Using a backward solution, the terminal F and catch for the oldest age in the terminal year are used to calculate the terminal population using [11] or [24] depending on if the oldest age in the cohort is completely fished out or not. Then the remaining catches from the cohort and the equation (cf. Pope 1972, equation 1.2)

$$N(a,y) = N(a+1,y+1)\ \exp(M) + C(a,y)\ \exp(M/2) \qquad [32]$$

are used to calculate the remaining abundance levels for the cohort. Abundance estimates from [32] are used to calculate estimates of fishing mortality directly from the expression (cf. Pope 1972, equation 1.8)

$$F(a,y) = \ln\left[\frac{N(a,y)}{N(a+1,y+1)}\right] - M \qquad [33]$$

Parameters are estimated similarly to Murphy's equation and Gulland's SPA model. The only difference is that population abundance and fishing mortality estimates are calculated in reverse order compared to Murphy's equation or Gulland's SPA model.

3.6 Doubleday's Model

Input data required by the model and estimated parameters are summarized in Table 8. Age-specific selectivities are estimated for all ages but the oldest, and effective effort parameters are estimated for all years but the last.

3.6.1 The Model. The model of Doubleday (1976) is based on the log catch model and the log catch ratio model. Both incorporate a separable fishing mortality formulation. Random variation is added to the catch equation by assuming that the observed catch is distributed lognormally. The stochastic counterpart to [11] is

$$C(a,y)\ \exp[\epsilon_c(a,y)] = \frac{N(a,y)\ F(a,y)}{F(a,y)+M}\ \{1-\exp[-F(a,y)-M]\} \qquad [34]$$

where $\epsilon_c(a,y)$ is random variable distributed $N(0,\sigma_c^2)$ with constant variance and $\exp[\epsilon_c(a,y)]$ is a lognormally distributed random variable.

The model is derived by substituting [8] for N(a,y) into [11] to provide an equation that relates catch for any age and year to its earlier recruitment and an exploitation history.

$$C(a,y)\ \exp[\epsilon_c(a,y)] = \frac{F(a,y)}{F(a,y)+M}\ \{1-\exp[-F(a,y)-M]\}$$
$$R(r,y-a+r)\ \exp\left[-\sum_{i=r}^{a-1}[F(i,y-a+i)+M]\right] \qquad [35]$$

When the stochastic catch equation is combined with the exponential survival model, the resulting equation is usually linearized by taking the natural log transform. On the log scale the errors are additive and are distributed normally. Taking logs of both sides of [35] gives

$$\ln[C(a,y)] = \ln[R(r,y-a+r)] - (a-r)M + \ln[F(a,y)] -$$
$$\sum_{i=r}^{a-1}\left[F(i,y-a+i)\right] - \ln[F(a,y)+M] +$$
$$\ln[1-\exp(-F(a,y)-M)] + \epsilon_c(a,y) \qquad [36]$$

In the first of Doubleday's models, equations [3] and [4] are substituted for F(a,y) in [36] to give (cf. Doubleday 1976, equation 7)

$$\ln[C(a,y)] = \ln[R(r,y-a+r)] - (a-r)M + [v(a)+e(y)] - \sum_{i=r}^{a-1}\Big[\exp[v(i)+e(y-a+i)]\Big] - \ln\Big[\exp[v(a)+e(y)]+M\Big] + \ln\Big[1-\exp\Big[-\exp[v(a)+e(y)]-M\Big]\Big] + \epsilon_c(a,y) \qquad [37]$$

where $\epsilon_c(a,y)$ is defined earlier. Equation [37] is nearly linear in the range $0.01 \le F(a,y) \le 2.72$.

A second equation, the log catch ratio model, is also used. Parameter values estimated from the log catch ratio model are used to supply starting values required by the iterative procedure when parameters from the log catch model are being estimated. A catch ratio model is constructed from [35] similar to [26]. Equations [3] and [4] are substituted for F(a,y) in [35] and the natural log is taken to give (cf. Doubleday 1976, p. 73, no equation number given)

$$\ln\left[\frac{C(a,y)}{C(a+1,y+1)}\right] = [v(a)+e(y)] - \ln\Big[\exp[v(a)+e(y)]+M\Big] + \ln\Big[1-\exp\Big[-\exp[v(a)+e(y)]-M\Big]\Big] + \Big[\exp[v(a)+e(y)]+M\Big] - [v(a+1)+e(y+1)] + \ln\Big[\exp[v(a+1)+e(y+1)]+M\Big] - \ln\Big[1-\exp\Big[-\exp[v(a+1)+e(y+1)]-M\Big]\Big] + \epsilon_c(a,y) \qquad [38]$$

3.6.2 Objective Function. The objective function is the sum of squared differences between the observed log catch-at-age data and the log catch-at-age data predicted by [37]. The parameter estimation procedure seeks to minimize

$$\sum_{y=1}^{Y} \sum_{a=1}^{A} \left[\text{pred } \ln[C(a,y)] - \text{obs } \ln[C(a,y)] \right]^2 \qquad [39]$$

3.6.3 Parameter Estimation Procedure. Doubleday uses a process of iterative linear approximation and estimation (linearization) to obtain least squares estimates of the parameters. This is a rather crude form of more recent non-linear regression least squares algorithms. The linearization method along with other more recent nonlinear regression methods are described by Bard (1974, p. 96) and Draper and Smith (1981, p. 462-468).

The value of the separability assumption is especially evident during parameter estimation. The number of fishing mortality parameters that need to be estimated in a non-separable model is AY. In the separable model this is reduced to A+Y. The separability assumption also permits statistical estimation of the parameters of interest and simultaneous analysis of data from all cohorts. Also an objective measure of how well the values predicted by the model correspond to the observed data set (i.e. the residual sums of squares) is available.

One problem associated with estimating parameters using [37] is that parameters are estimated on the logarithmic scale which introduces some bias in the estimates. We are more interested in these values on the arithmetic scale and logarithmic means and variances can be converted. Generally if

$Y=\ln(X)$, where Y is a normally distributed random variable with mean $\mu(Y)$ and variance $\sigma^2(Y)$, then $\mu(X)=\exp\{\mu(Y)+[\sigma^2(Y)/2]\}$ and $\sigma^2(X)=[\exp\{\sigma^2(Y)+2\mu(Y)\}(\exp\{\sigma^2(Y)\}-1)]$ (Brownlee 1965). If parameters on the natural log scale are transformed back to the arithmetic scale by simply taking the antilog then they would be biased by the factor $\exp(\sigma^2(Y)/2)$. Bias could be especially significant in the case of population biomass estimates which are not estimated directly from the model. Computation of population biomass involves a sum of several abundance parameters, each of which may be biased. For instance if the model provides estimates of $\ln[N(a,y)]$ then population biomass (uncorrected for bias) is given by

$$B(y) = \sum_{a=r}^{s} \exp\{\ln[N(a,y)]\}\ W(a) \qquad [40]$$

Clearly B(y) will be biased downward if there are a large number of ages or the variance for each $\ln[N(a,y)]$ are large. If abundance parameter variances are small then any bias introduced by taking the logarithmic transform will be small. Beauchamp and Olson (1973) describe other methods of correcting for bias.

One other very serious problem with the separable model recognized by Doubleday (1976) and then later Pope (1977) is that catch-at-age information alone are not sufficient to reliably estimate stock abundance because fishing mortality and stock size are highly negatively correlated. Additional data are required to resolve the multicolinearity between population abundance and mortality parameters. This problem is discussed in detail by Shepherd and Nicholson (1986).

3.7 Pope and Shepherd's Model

Input data required by the log catch ratio model of Pope and Shepherd (1982) and estimated parameters are summarized in Table 8. In the following description age and year indicies are reported on a relative scale.

3.7.1 The Model. The model of Pope and Shepherd (1982) is based on the nonlinear log catch ratio model that incorporates the separability assumption and a random error term. It is derived by taking the natural log of the catch ratio model [26], substituting the separability assumption [2] for F(a,y) and taking logarithms. It is written

$$\begin{aligned}\ln\left[\frac{C(a+1,y+1)}{C(a,y)}\right] = \\ &\ln[s(a+1)] + \ln[f(y+1)] + \ln[s(a)f(y)+M] - \\ &s(a)f(y) - M + \ln\left[1-\exp\{-[s(a+1)f(y+1)]-M\}\right] - \\ &\ln[s(a)] - \ln[f(y)] - \ln[s(a+1)f(y+1)+M] - \\ &\ln\left[1-\exp\{-[s(a)f(y)]-M\}\right] + \epsilon_c(a,y) \qquad [41]\end{aligned}$$

where $\epsilon_c(a,y)$ is a random variable distributed $N(0,2\sigma_c^2)$ with constant variance. Note that terms in [41] have opposite signs when compared to [38] since the catch ratio is reversed.

Equation [41] is augmented by an additional equation

$$s(a') = 1.0 \qquad [42]$$

where a' is an intermediate reference age, usually the age of full recruitment. Equation [42] assures that all $s(a) \leq 1.0$.

3.7.2 Parameter Estimation Procedure. Pope and Shepherd use a sequential two stage least squares algorithm as their parameter estimation procedure. The first stage uses an iteratively re-weighted solution algorithm, somewhat analogous to a two-way analysis of variance, to estimate mortality parameters. The second stage estimates population numbers at age. Each stage utilizes a separate objective function.

3.7.2.1 Stage One Objective Function. The stage one objective function consists of two equations representing marginal totals; one for each year, summed over all age-classes within a year giving Y-1 annual residuals and one for each age-class, summed over all years within an age-class giving A-1 age-class residuals. Pope and Shepherd point out the similarity of this approach to a two-way analysis of variance. The two equations to be minimized,

$$Re(y) = \left[\sum_{a=1}^{A-1} \text{pred} \ln\left[\frac{C(a+1,y+1)}{C(a,y)}\right] - \sum_{a=1}^{A-1} \text{obs} \ln\left[\frac{C(a+1,y+1)}{C(a,y)}\right]\right] \quad [43]$$

(for y=1,...,Y-1) and

$$Re(a) = \left[\sum_{y=1}^{Y-1} \text{pred} \ln\left[\frac{C(a+1,y+1)}{C(a,y)}\right] - \sum_{y=1}^{Y-1} \text{obs} \ln\left[\frac{C(a+1,y+1)}{C(a,y)}\right]\right] \quad [44]$$

(for a=1,...,A-1), calculate the sum of differences between the observed log catch ratio and the log catch ratio predicted by [41] summed over the appropriate index (cf. Pope and Shepherd 1982, equations 5 and 6).

3.7.2.2 Stage One Parameter Estimation Procedure. Parameter estimates are determined by minimizing equations [43] and

[44] with the following procedure. First, all s(a) are initialized to be equal to 1.0 for a=1,...,A-1 and f(y) are initialized to be equal to f(Y) for y=1,...,Y-1. Next equation [41] is evaluated and the residuals Re(a) and Re(y) are calculated via the objective functions [43] and [44], where Re(a) is the residual for age summed over all years and Re(y) is the residual for year summed over all ages. At each iteration, updated parameter estimates are calculated by multiplying the old parameter estimate by the empirical weighting factors We(a) and We(y), where We(a) = exp[Re(a)/Y] and We(y) = exp[Re(y)/A]. The term Re(a)/Y can be considered an average (over Y years) residual for age a. Similarly, Re(y)/A can be considered the average (over A ages) residual for year y. The s(a) parameters are renormalized relative to the reference age (a') and the procedure is repeated again until the solution converges.

Weighting factors are determined by considering the change in the parameter required to eliminate the residual. The actual functional form is a result of three approximations and one assumption (see Pope and Shepherd 1982, Appendix 1).

3.7.2.3 Stage Two Objective Function. The parameter estimation algorithm of stage one cannot directly estimate population at age. Because the log catch ratio model is density-independent population abundance parameters are not involved. Estimates of the abundance of the youngest age in each cohort, R(1,y) and R(a,1), are derived by minimizing a second objective function. Note, however, that these abundance estimates are conditional upon the values of s(a) and f(y) obtained from evaluating the first objective function. Once the population numbers of the youngest age of each cohort are estimated then any N(a,y) can be estimated from the recurrence relationship [7].

Two separate objective functions are used in the stage two parameter estimation procedure. First define

$$E(a,y+a-1) = \frac{F(a,y+a-1)}{Z(a,y+a-1)} \left[1 - \exp[-Z(a,y+a-1)]\right] \exp\left[- \sum_{i=1}^{a-1} Z(i,y+i-1)\right] \qquad [45]$$

(cf. Pope and Shepherd 1982, equation 11). Then the first equation to be minimized is

$$\sum_{i=1}^{amax} \left[\ln[C(i,y+i-1)] - \ln[R(1,y)]E(i,y+i-1)\right]^2 \qquad [46]$$

where i is the i th age in the cohort and amax is the oldest age in the cohort. The second equation to be minimized is

$$\sum_{m=1}^{ymax} \left[\ln[C(a+m-1,m)] - \ln[R(a,1)]E(a+m-1,m)\right]^2 \qquad [47]$$

where m is the m th year in the cohort and ymax is the last year that the cohort is in the catch-at-age data matrix. Equations [46] and [47] have closed analytical solutions.

3.7.2.4 Stage Two Parameter Estimation Procedure. The objective functions [46] and [47] are minimized when the partial derivative of the function with respect to the parameters R(1,y) and R(a,1) are zero. These are closed solutions which can be expressed as

$$\ln[R(1,y)] = \frac{1}{amax} \sum_{i=1}^{amax} \left[\ln[C(i,y+i-1)] - \ln[E(i,y+i-1)] \right] \quad [48]$$

(cf. Pope and Shepherd 1982, equation 12) and

$$\ln[R(a,1)] = \frac{1}{ymax} \sum_{m=1}^{ymax} \left[\ln[C(y+m-1,m)] - \ln[E(y+m-1,m)] \right] \quad [49]$$

Estimates of N(a,y) for all succeeding ages and years are obtained with the recurrence relationship [7].

4.0 DESCRIPTION OF ASA METHODS - EFFORT DATA REQUIRED

All methods described up to now have not required effort as input data. The introduction of effort data into the ASA procedures creates entirely new problems. In contrast to previous methods where fishing mortality was expressed as a fraction of catch to the total stock, effort/catchability formulations assume that fishing mortality is proportional to effort or the fishing intensity exerted by the fishing gear (Beverton and Holt 1957). The constant of proportionality is often referred to as the catchability coefficient or the degree to which the fish are vulnerable to the gear. If catchability is considered a constant or average the fishing mortality model is written

$$F = q\ f \quad [50]$$

Equation [50] is similar to the separable fishing mortality formulation except that the meaning behind the coefficients are different.

Fishing mortality is seldom constant from year to year and often within any one year there is a significant age effect. To accommodate these features further structure can be added to [50] by considering the catchability and effort parameters as separate submodels. The level of complexity depends on what contingencies are appropriate to the fishery being analyzed. A fully developed fishing mortality model might have to account for an age effect such as age-specific selectivity of the gear, an age and/or year effect such as changes in catchability with age and/or time, a population density effect such as a density-dependent catchability coefficient (Fox 1974; Schaaf 1975; MacCall 1976; Ulltang 1976; Garrod 1977; Peterman and Steer 1981; Bannerot and Austin 1983) or a gear saturation effect (Bannerot and Austin 1983). To accommodate all of these features substitute

$$q = q(y)\ s(a)\ N(a,y)^{\Theta} f(y)^{\Phi} \qquad [51]$$

into [50] which gives the generalized fishing mortality model

$$F(a,y) = q(y)\ s(a)\ N(a,y)^{\Theta} f(y)^{1+\Phi} \qquad [52]$$

Equation [52] is indeed a complex fishing mortality model. Values of F could be easily computed if estimates of all parameters were at hand. In practice, lack of data necessitates a number of simplifying assumptions. Equation [52] can be reduced to a less complicated form if we are willing to hold constant the value of one or more parameters. For instance consider the following assumptions: all age groups in the fishery are fully recruited to the gear (all $s(a)=1.0$), there are no temporal trends in catchability (q is constant), and there are no significant density-dependent effects ($\Theta=0$) or gear saturation effects ($\Phi=0$). Equation [52] then would simplify to [50]. If we assume that $\Theta=0$, $\Phi=0$ and catchability exhibits no temporal trends (q is constant and equal to 1.0) but there exists a significant age effect then [52] simplifies to [2]. Finally we could assume that catchability does not exhibit any temporal, density-dependent, or gear saturation effects but there are some partially recruited age-classes in the fishery. In this case fishing mortality can be expressed as

$$F(a,y) = q\ s(a)\ f(y) \qquad [53]$$

Even though equation [52] is highly nonlinear, it is useful because it represents a very general expression from which numerous different expression of fishing mortality can result. Unfortunately it is extremely difficult to predict what effect these nonlinearities will have on the total estimation process. Paloheimo and Dickie (1964) and more recently Cooke (1985) provide a good discussion of the problems introduced by nonlinearities in the catch effort relationship.

4.1 Paloheimo's Model

The catch-per-unit effort model of Paloheimo (Paloheimo 1980) uses catch-at-age in numbers and annual effective effort data as input data. Paloheimo's model can take two forms. Paloheimo calls these the constant catchability model (hereafter referred to as the linear model) and the variable catchability model (hereafter referred to as the nonlinear model). The parameters estimated by these are summarized in Table 8.

4.1.1 The model. The log CPUE model relates the decline in CPUE of a cohort to age and cumulative effort. The model is based on the catch equation [11] and uses the separability assumption [53] to describe fishing mortality. Random variation is added to the catch equation by assuming that the observed catch is distributed lognormally. The model is written after taking the natural log of both sides

$$\ln\left[\frac{C(a,y)}{s(a)f(y)}\right] = \ln[qR(r,y-a+r)] - \ln[qs(a)f(y)+M] + \ln[1-\exp\{-qs(a)f(y)-M\}] - q\sum_{i=r}^{a-1}[s(i)f(y-a+i)] - (a-r)M + \epsilon_c(a,y) \qquad [54]$$

where $\epsilon_c(a,y)$ is random variable distributed $N(0,\sigma_c^2)$ with constant variance. Equation [54] can be simplified by using Paloheimo's approximation

$$\frac{1-\exp(-x)}{x} \approx \exp(-x/2) \qquad [55]$$

which is valid for small values of x (Paloheimo 1961). The approximation [55] can be substituted for the second and third terms on the RHS of equation [54] by letting $x=[-qs(a)f(y)-M]$. The result is the log CPUE equation

$$\ln\left[\frac{C(a,y)}{s(a)f(y)}\right] = \ln[qR(r,y-a+r)] - [2(a-r)-1]\frac{M}{2} - q\left[\frac{1}{2}[s(a)f(y)] + \sum_{i=r}^{a-1}[s(i)f(y-a+i)]\right] + \epsilon_c(a,y) \qquad [56]$$

which is valid for $a > r$. If $a=r$ then the summation term in

[56] is not evaluated. $\epsilon_c(a,y)$ is as described earlier.

4.1.1.1 Linear Model. If the assumption that all ages are fully recruited (i.e. all s(a)=1.0) is valid, then [50] can be substituted for F(a,y) in [54] instead of using the separable formulation. Equation [54] then reduces to an equation that is linear in all the parameters. It can be written (cf. Paloheimo 1980, equation 5)

$$\ln\left[\frac{C(a,y)}{f(y)}\right] = \ln[qR(r,y-a+r)] - [2(a-r)-1]\,\frac{M}{2}$$
$$q\left[\frac{1}{2}\,f(y) + \sum_{i=r}^{a-1} f(y-a-i)\right] + \epsilon_c(a,y) \qquad [57]$$

4.1.1.2 Nonlinear Model. The nonlinear model relaxes the assumption of average constant catchability for all ages. Paloheimo expresses age-specific catchability as deviations from an average

$$q(a) = q + d(a) \qquad [58]$$

where d(a) is an age-dependent correction term that can be positive or negative. In order to assure that the parameters are estimable one of the d(a)'s must be zero or alternatively $\Sigma\, d(a) = 0$. The nonlinear model is created by substituting [58] for q into [57] to give (cf. Paloheimo 1980, equation 6)

$$\ln\left[\frac{C(a,y)}{f(y)}\right] = \ln[qR(r,y-a+r)] - [2(a-r)-1]\,\frac{M}{2} +$$
$$\left[\frac{\ln[q+d(a)]}{q}\right] - \frac{1}{2}\,[q+d(a)]\,f(y) -$$
$$\sum_{i=r}^{a-1} [q+d(a-i)]\,f(y-a-i) + \epsilon_c(a,y) \qquad [59]$$

4.1.2 Objective function. The objective function is the sum of squared differences between the observed log catch-per-unit-effort and the log catch-per-unit-effort predicted by either [57] or [59]. The parameter estimation procedure seeks to minimize

$$\sum_{y=1}^{Y} \sum_{a=1}^{A} \left[\text{pred } \ln\left[\frac{C(a,y)}{f(y)}\right] - \text{obs } \ln\left[\frac{C(a,y)}{f(y)}\right]\right]^2 \qquad [60]$$

4.1.3 Parameter Estimation Procedure. Since equation [57] is linear in all parameters it can solved directly with standard multiple regression statistical techniques. Paloheimo estimates the parameters in the nonlinear equation by the linearization technique described in Section 3.6.3.

4.2 Fournier and Archibald's Model

The two goals behind the model of Fournier and Archibald (1982) are 1) to provide a flexible mathematical formulation and 2) to recognize that the processes the model is trying to describe and the information submitted to the model are subject to error. The model provides a systematic way to deal with the fact that different sources of data contain different types of error.

Data submitted and parameters estimated from the model are summarized in Table 8. The catchability submodel can include time-independent (constant), density-dependent or time-dependent catchability similar to equation [51] in Section 4.0. The reader can refer to Fournier and Archibald (1982) and Archibald et al. (1983a) for specifics.

Because this model is generalized it is very flexible with respect to the number of parameters it can estimate. The following description will only cover the main concepts of the model.

4.2.1 The model. The model is a combination of a stochastic model describing the catch data and a model describing the exploitation process. The stochastic model acknowledges that there are errors in estimating total catch and age composition percentages.

Given O(y) and H(a,y) and the assumptions 1) H(a,y) and O(y) are independent, 2) $O(y)=C(y)\exp[\epsilon_c(y)]$ where the $\epsilon_c(y)$ are independent random variables distributed $N(0,\sigma_c^2)$ with constant variance, and 3) there are no aging errors, then the likelihood function for the parameters C(y), G(a,y), and σ_c is given by (cf. Fournier and Archibald 1982, equation 1.0)

$$\prod_{a=1}^{A}\prod_{y=1}^{Y} G(a,y)^{H(a,y)} \prod_{y=1}^{Y} \frac{1}{\sqrt{2\pi}\sigma_c} \exp\left[-\frac{1}{2}\left[\frac{\ln[O(y)]-\ln[C(y)]}{\sigma_c}\right]^2\right] \qquad [61]$$

The exploitation process is described by a log catch ratio model where catch-at-age is expressed as a product of the total catch times the age composition estimate

$$C(a,y) = C(y)\ G(a,y) \qquad [62]$$

Substituting [62] for C(a,y) in [26] (note that the ratio of catches is reversed relative to catch-ratio models discussed earlier) and taking logs of both sides gives

$$\ln\left[\frac{C(a,y)\ G(a,y)}{C(a+1,y+1)\ G(a+1,y+1)}\right] =$$

$$= \ln[G(a,y)] + \ln[C(y)] - \ln[G(a+1,y+1)] - \ln[C(y+1)]$$

$$= \ln[F(a,y)] - \ln[F(a,y)+M] + \ln[F(a+1,y+1)+M] - \ln[F(a+1,y+1)] + \ln[1-\exp\{-F(a,y)-M\}] - \ln[1-\exp\{-F(a+1,y+1)-M\}] \quad [63]$$

A separable fishing mortality formulation is used. The age-dependent component is represented with a function rather than a fixed number of parameters. Representing age-specific selectivities with a curve requires a smaller number of parameters than the number of age groups. Fournier and Archibald (1982) use a nonlinear scaling of the age index which they call the VB parameterization. This is written (cf. Fournier and Archibald 1982, equation 2.2)

$$a(w) = \frac{-1 + 2\ (1 - w^{a-1})}{(1 - w^{A-1})} \quad [64]$$

for $0 < w < 1$. The relationship between fishing mortality and fishing effort is (cf. Fournier and Archibald 1982, equation 3.1)

$$\ln[F(a,y)] = b(w) + \ln[q(y)] + \ln[f(y)] + \epsilon_f(y) \quad [65]$$

where $b(w) = b_1\ a(w) + b_2\ a(w)^2$

Expressed in terms of the deviation between observed and predicted fishing mortality [65] is written

$$\epsilon_f(y) = \left[\left[\ln[F(a,y)]\right] - \left[b(w) + \ln[q(y)] + \ln[f(y)]\right]\right] \quad [66]$$

where $\epsilon_f(y)$ is the annual deviation between the predicted level of fishing mortality and the observed level. Assume that $\epsilon_f(y)$ is a random variable distributed $N(0,\sigma_f^2)$ with constant variance.

With fecundity-at-age data further structure can be added to the model by incorporating a spawner-recruit relationship. Recruitment is modeled by a Ricker spawner-recruit function with log-normal errors

$$R(r,y+r) = \alpha\ SP(y)\ \exp[-\beta SP(y)]\ \exp[\epsilon_{sr}(y)] \qquad [67]$$

Expressed in terms of the deviation between observed and predicted recruitment [67] can be written

$$\epsilon_{sr}(y) = \left[\left[\ln[R(r,y+r)]\right] - \left[\ln(\alpha) + \ln[SP(y)] - \beta\ SP(y)\right]\right] \qquad [68]$$

where SP(y) is the reproductive potential of the population in year y and $\epsilon_{sr}(y)$ are random variables distributed $N(0,\sigma_{sr}^2)$ with constant variance. Relative reproductive potential is given by (cf. Fournier and Archibald 1982, equation 5.1)

$$SP(y) = \sum_{a=r}^{s} fec(a)\ N(a,y) \qquad [69]$$

where fec(a) is the fecundity of age group a.

Aging error information can be incorporated into the model with a minimum of difficulty. Let p(a,i) be the probability that a fish from age class i is judged during age determination to be age a. Then the probability that a fish picked at random in year y will be classified as age a is given by (cf. Fournier and Archibald 1982, equation 6.1)

$$\sum_{i=1}^{A} p(a,i)G(i,y) \qquad [70]$$

4.2.2 Objective Function. The objective function to be minimized is obtained by taking the natural log of the likelihood function [61] which gives (cf. Fournier and Archibald 1982, equation 1.1) the log likelihood equation (ignoring the constant)

$$\sum_{a=1}^{A} \sum_{y=1}^{Y} H(a,y) \ln[G(a,y)] - \sum_{y=1}^{Y} \frac{1}{2} \left[\frac{\ln[O(y)]-\ln[C(y)]}{\sigma_c}\right]^2$$

$$- Y \ln[\sigma_c] \qquad [71]$$

If we define

$$\beta(a,y) = \ln\left[\frac{G(a,y)\ C(y)}{O(y)}\right] \qquad [72]$$

then

$$\left[\ln[O(y)] - \ln[C(y)]\right]^2 = \left[\ln\left[\sum_{a=1}^{A} \exp[\beta(a,y)]\right]^2\right] \qquad [73]$$

and

$$G(a,y) = \frac{\exp[\beta(a,y)]}{\left[\sum_{a=1}^{A} \exp[\beta(a,y)]\right]} \qquad [74]$$

By substituting [72] and [74] into [71] the log likelihood function is given by (cf. Fournier and Archibald 1982, equation 1.9)

$$\sum_{y=1}^{Y} \sum_{a=1}^{A} H(a,y) \left[\beta(a,y) - \ln\left[\sum_{a=1}^{A} \exp[\beta(a,y)]\right]\right]$$

$$- L_c \sum_{i=1}^{Y} \frac{1}{2} \ln\left[\sum_{a=1}^{A} \exp[\beta(a,y)]\right]^2 - m \ln[\sigma_c] \qquad [75]$$

where m is the greatest integer less than or equal to s/2 and $L_c = 1/\sigma_c^2$. L_c is considered a penalty weight that determines the penalty for deviating from the observed catch relationship.

When additional information is added to the model the objective function is augmented by additional terms. If effort data is available then the following term is added to [75] (cf. Fournier and Archibald 1982, equation 4.1)

$$- L_f \sum_{y=1}^{Y} \frac{1}{2} \epsilon_f(y)^2 \qquad [76]$$

where $L_f = 1/\sigma_f^2$. As before, L_f is the penalty weight for deviating from the observed effort-fishing mortality relationship.

If spawner-recruit information is available then the objective function is augmented by the term (cf. Fournier and Archibald 1982, equation 5.3)

$$- L_{sr} \sum_{y=1}^{Y-r} \frac{1}{2} \epsilon_{sr}(y)^2 \qquad [77]$$

where $L_{sr} = 1/\sigma_{sr}^2$ is the penalty weight for deviating from the spawner-recruit relationship.

If aging error information is included in the model then just the first term in [75] is changed (cf. Fournier and Archibald 1982, equation 6.3). The objective function for a full model that included catch-at-age, effort, fecundity-at-age and aging error data would be

$$\sum_{y=1}^{Y} \sum_{a=1}^{A} H(a,y) \ln\left[\sum_{i=1}^{A} p(a,i) \left[\frac{\exp[\beta(i,y)]}{\sum_{m=1}^{A} \exp[\beta(m,y)]} \right]\right]$$

$$- L_c \sum_{i=1}^{Y} \frac{1}{2} \ln\left[\sum_{a=1}^{A} \exp[\beta(a,y)]\right]^2$$

$$- L_f \sum_{y=1}^{Y} \frac{1}{2} \epsilon_f(y)^2 - L_{sr} \sum_{y=1}^{Y-r} \frac{1}{2} \epsilon_{sr}(y)^2 \qquad [78]$$

4.2.3 Parameter Estimation Procedure. A maximum likelihood solution algorithm is used to solve for the parameters of the model. Seber (1982, p. 4), Bard (1974, p. 61-71) and Norden (1972, 1973) provide a helpful review and discussion of this parameter estimation method. Briefly, a likelihood function is an expression that describes the joint probability distribution of the observations viewed as a function of the parameters. The maximum likelihood estimate of a parameter, Γ say, is that value of Γ for which the likelihood function attains its maximum value. The log likelihood function is frequently used because it is less complicated. Note that maximizing a likelihood function is the same as minimizing the negative of the log likelihood function. The normal equations for finding the maximum likelihood estimates are obtained by taking partial derivatives of the log-likelihood function with respect to the unknown parameters and setting the results to zero. These equations usually do not allow a simple solution so iterative computer algorithms must be used.

To initiate the parameter estimation procedure, a guess or estimate of each parameter in the model must be supplied to the maximum likelihood solution algorithm. These values are used as a starting point by the minimization procedure in its search for the final set of parameter values that minimize the objective function.

In Fournier and Archibald's implementation of this method, models of increasing complexity are solved in a stagewise fashion by exploiting this feature of the minimization procedure coupled with the flexible mathematical model. At each step the parameters are estimated simultaneously. For instance a simple model could be fit using just catch and effort data. Once final parameter values are obtained, these can then be used as starting estimates for another run where a more complicated model, say one that includes a spawner-recruit relationship and time-dependent catchability, is fit. By gradually adding data to the model, refining hypotheses, and relaxing assumptions a stagewise minimization can be carried out. Bard (1974, p. 72) calls the process of making use of prior information in a likelihood function Bayesian estimation. Applications of this method are presented in Archibald et al. (1983b) and Fournier and Archibald (1982).

4.3 Dupont's Model

Data requirements for the model of Dupont (1983) and estimated parameters are summarized in Table 8. The goal of Dupont's model is to relax as much as possible what he considers strong assumptions about natural mortality and the nature of the birth/death/capture process used in most other models. The notation used in describing the model generally follows Dupont (1983).

4.3.1 The model. The model assumes a competing risk model similar to the hazard regression models of Cox (1972) where the stochastic mortality processes include natural and fishing mortality. This approach is similar to Chapman (1961) who described a death mechanism using a simple Poisson stochastic process. Chapman (1961) provides definitions of distribution functions and a derivation of this approach. From the input data each member of the population can be assigned to one of k distinct cohorts. We assume that catch from cohort k, D_{kj}, is known for consecutive time intervals (t_j, t_{j+1}) $j=1,\ldots,Y-1$. Let $\pi_k(t)$ be the hazard function due to fishing and $\phi_k(t)$ be the hazard function for all other causes of death not associated with directed fishing effort. Finally define $(t_{j(k)}, t_{j(k)+1})$ to be the earliest time interval that members of the k th cohort are caught; N_k be the size of the k th cohort at time $t_{j(k)}$; N_k be the size of the k th cohort at time t; N(t) be the total population size at time t; $D_k = (D_{k,j(k)},\ldots,D_{k,Y-1})$ be the catch vector from the k th cohort; $D = (D_k,\ldots,D_K)$ be the total catch vector; d, d_k and d_{kj} are realizations of the random variables D, D_k and D_{kj} ; and $|d_k|$ is the total catch from the k th cohort. Each fish from the k th cohort alive at time $t > t_{j(k)}$ is subject to competing hazard functions. With these definitions the probability that a member of cohort k is alive at time t is given by

$$r_k(t) = \exp\left[- \int_{t_{j(k)}}^{t} \pi_k(\tau) + \phi_k(\tau) \, d\tau \right] \qquad [79]$$

then the probability that a fish from cohort k is caught in the j th catch interval is given by

$$p_{kj} = \int_{t_j}^{t_{j+1}} \pi_k(t) \, r_k(t) \, dt \qquad [80]$$

the probability that a fish from cohort k is caught in any interval is

$$p_k = \sum_{j=j(k)}^{Y-1} p_{kj} \qquad [81]$$

and the probability that a fish from cohort k is never caught is $1-p_k$.

Given the competing risk model, the likelihood equation for the deaths in each catch interval, from each cohort, is given (cf. Dupont 1983, equation 2)

$$\prod_{k=1}^{K} \frac{N_k!}{(N_k-|d_k|)! \prod_{j=j(k)}^{Y-1} d_{kj}!} (1-p_k)^{(N_k-|d_k|)} \prod_{j=j(k)}^{Y-1} p_{kj}^{(d_{kj})} \qquad [82]$$

The mortality functions are quite flexible. In the most traditional fisheries application let

$$\phi_k(t) = M \qquad [83]$$

and

$$\pi_k(t) = F(y) \qquad [84]$$

A separable formulation is also possible. This can be expressed (cf. Dupont 1983, Model I, p. 1026) with

$$\phi_k(t) = M \qquad [85]$$

and

$$\pi_k(t) = \pi\ s(a)\ f(y) \qquad [86]$$

The following two models relax the separability assumption somewhat by allowing the shape of the selectivity and effort curves to be fit to the data. In the first, selectivity is stratified by groups depending on a chosen annual partitioning of the data. This approach would be appropriate if the selectivity characteristics of the fishery changed over time. Let g index the number of groups (assume there are two), z(a,g) be an dummy variable, v(a,g)=ln[s(a,g)], and y(x) be the calendar year in which the selectivities changed. Then (cf. Dupont 1983, Selectivity Model, p. 1027)

$$\phi_k(t) = M \qquad [87]$$

and

$$\pi_k(t) = \pi\ f(y)\ \exp\left[\sum_{a=1}^{A} \sum_{g=1}^{G} z(a,g)\ v(a,g)\right] \qquad [88]$$

where

$$z(a,g) = \begin{cases} 1 \text{ if age} = a,\ g = 1 \text{ and } y < y(x) \\ 1 \text{ if age} = a,\ g = 2 \text{ and } y \geq y(x) \\ 0 \text{ otherwise} \end{cases}$$

The final mortality model assumes that selectivity values are know. If e(y)=ln[f(y)] then this can be written (cf. Dupont 1983, Effort Model, p. 1028)

$$\phi_k(t) = M \qquad [89]$$

and

$$\pi_k(t) = \pi\ s(a,g)\ \exp\left[\sum_{y=1}^{Y} z(y)\ e(y)\right] \qquad [90]$$

where

$$z(y) = \begin{cases} 1 \text{ if current year} = y \\ 0 \text{ otherwise} \end{cases}$$

Using the competing risk model, Dupont derives an expression for the estimated size of the k th cohort at time t_j (cf.Dupont 1983, equation 3). This is given as the number of fish known to be alive at that time plus the expected number of survivors who will not be caught. This can also be expressed as the expected number of fish alive at time t_j minus the expected catch after time t_j plus the actual catch after this time. Dupont also derives variance formulas for the parameters and an expression for the prediction of future catches.

4.3.2 Objective Function. The objective function to be minimized is

$$\sum_{k=1}^{K} \left[\left[\ln(N_k!) - \ln[(N_k-|d_k|)!] - \sum_{j=j(k)}^{Y-1} d_{kj}! + (N_k-|d_k|) \ln[1-p_k] \right] + \sum_{j=j(k)}^{Y-1} \left[d_{kj} \ln(p_{kj}) \right] \right] \quad [91]$$

Dupont (1983) gives a full derivation of the maximum likelihood equation in an appendix.

4.3.3 Parameter Estimation Procedure. Parameters of the model are estimated by maximum likelihood estimation. Initial starting values for the mortality parameters are supplied by applying regression methods to the data (Seber 1982, p. 329).

4.4 The CAGEAN Model

Data requirements for the Catch-Age Analysis (CAGEAN) model of Deriso et al. (1985) and estimated parameters are summarized in Table 8. Deriso et al. (1987) discuss extensions to the CAGEAN model that permit the incorporation of fishery independent data into the analysis and other nonstandard extensions.

4.4.1 The model. The CAGEAN model is based on the log catch equation that incorporates the separability assumption and the assumption that observed catch-at-age data differ from predicted values by a log-normal random variable after Doubleday (1976) (see equation [37] in Section 3.6.1).

As in the model of Fournier and Archibald (1982), the CAGEAN model can also incorporate additional information into an analysis. If effort data is available but we believe effort is measured with error then the relationship between fishing mortality and effort is not exact. The difference can be described by a log-normal error. This relationship is written

$$\text{predicted } f(y) = \Big[q \text{ observed } [f(y)] \exp[\epsilon_f(y)] \Big] \qquad [92]$$

where $\epsilon_f(y)$ is a random variable distributed $N(0,\sigma_f^2)$ with constant variance. Expressed in terms of the difference between observed and predicted effort [92] is written

$$\epsilon_f(y) = \Big[\Big[\text{predicted } \ln[f(y)]\Big] - \Big[\ln(q)+\text{observed } \ln[f(y)]\Big]\Big] \qquad [93]$$

If fecundity-at-age data is available then a spawner recruit relationship can be added to the model. Let recruitment be given by a Ricker spawner recruit relationship with a log-normal error as in equation [67]. In the CAGEAN model spawner-recruit data is brought into the analysis in terms of the difference between observed and predicted recruitment as in equation [68]. Recruitment, $\ln[R(r,y+r)]$, is estimated from equation [37], $SP(y)$ the number of eggs produced by the spawning stock is estimated from equation [69] in year y, and $\epsilon_{sr}(y)$ is a random variable distributed $N(0,\sigma_{sr}^2)$ with constant variance.

The CAGEAN implementation of the Doubleday model is extremely flexible with respect to which parameters can be estimated and the age and year intervals over which the parameters can be estimated. Because the CAGEAN model generalizes the model of Doubleday (1976) (see Section 3.6.1) it can accommodate the log CPUE model of Paloheimo (1980) as a special case. If effort data is available and it is assumed effort is measured without error then the log catch model can be rearranged to give [54] where [53] is substituted for F(a,y). One other benefit of generalizing Doubleday's model is the ability to carry out analyses that account for systematic changes in the selectivity with a minimum of difficulty. Deriso et al. (1987) refer to this as stratified catch-age analysis. They also discuss other useful extensions to their basic model.

4.4.2 Objective Function. The objective function to be minimized (for the CAGEAN model without any auxiliary information) is the sum of squared differences between the observed log catch-at-age data and the log catch-at-age data predicted by [37] (cf. Deriso et al. 1985, equation 8)

$$SSQ(catch) = \sum_{y=1}^{Y} \sum_{a=1}^{A} \left[\text{pred } \ln[C(a,y)] - \text{obs } \ln[C(a,y)] \right]^2 \qquad [94]$$

If effort data is available but we believe that the relationship between effort and fishing mortality is not exact then [94] is modified by adding the term (cf. Deriso et al. 1985, p. 817)

$$SSQ(effort) = L_f \, \Sigma \, [\epsilon_f(y)]^2 \qquad [95]$$

where L_f is the ratio of the variance of the observed log catch from that predicted by [37] divided by the variance of observed log effort, σ_f^2.

If fecundity-at-age data is available then more structure can be added to the model. The objective function may be modified further by adding the term (cf. Deriso et al. 1985, p. 817)

$$SSQ(spawn) = L_{sr} \Sigma [\epsilon_{sr}(y)]^2 \qquad [96]$$

where L_{sr} is the ratio of the variance of observed log catch from that predicted in [37] divided by the variance of the stock-recruitment relationship, σ_{sr}^2.

If catch, effort and fecundity data are available, then the objective function to be minimized for the full CAGEAN model is (cf. Deriso et al. 1985, equation 9)

$$\text{minimize} \left[SSQ(catch) + SSQ(effort) + SSQ(spawn)\right] \qquad [97]$$

The penalty weights determine the influence the auxiliary terms have on the parameter estimates calculated from the objective function. The penalty weight for the SSQ(catch) term is assumed to be equal to 1.0. Deriso et al. (1985) also provide two alternative objective function formulations.

4.4.3 Parameter Estimation Procedure. A non-linear least squares regression algorithm (Marquardt 1963) is used to estimate the unknown parameters of the CAGEAN model by minimizing the sum of squares. Also see Conway et al. (1970) for details on the algorithm. When auxiliary information is available the CAGEAN model can carry out Bayesian estimation (see Section 4.2.3)

To initiate the parameter estimation procedure, a guess or estimate of each parameter in the model must be supplied to the nonlinear least squares solution algorithm. Once a minimum sum of squares solution is achieved variances can be calculated for all parameters of the model.

To deal with the problem of multicolinearity between population abundance and mortality parameters (see Section 3.6.3) and the approximate nature of the variance estimates resulting from the solution algorithm the CAGEAN model calculates parameter estimates and their standard deviations by the Monte Carlo simulation method known as the bootstrap technique (Efron 1982). Briefly, the bootstrap procedure involves stochastically generating a number of different catch-at-age data sets from one minimum sum of squares solution generated by the Marquardt non-linear least squares regression algorithm. A solution produces a set of estimates for the unknown parameters of equation [37], a vector of predicted catch-at-age values, and a residual vector. The residuals measure the agreement between observed catches and catches predicted by equation [37]. Elements of the residual vector are sampled with replacement and randomly added to the predicted catch-at-age data matrix. The advantage of this approach is that the new data matrix has the same statistical properties as the original. These different catch-at-age data matrices are then resubmitted to the solution algorithm to obtain bootstrap replications of the parameter estimates. If the solution has converged to the global minimum and the fit is good (i.e. the residuals are small) then the bootstrap estimates will have small standard deviations and they will be close in value to the minimum sum of squares parameter estimates. Means and variances of the bootstrap estimates are calculated in the usual way. Differences between the minimum sum of squares estimates and the bootstrap means is a measure of bias. This feature of bootstrap estimates is useful when calculating aggregate measures of population abundance such as population biomass since corrections for bias are not required (see Section 3.6.3). Bootstrap means and standard deviations for the fishing mortality parameters require bias correction if variances are large since they are calculated on the logarithmic scale (see Section 3.6.3)

Note that recruitment parameters from the log catch model and recruitment parameters from the spawner-recruit model are estimated simultaneously if the spawner-recruit relationship is included. Deriso et al. (1985) caution that because spawners nor recruits are not observed directly (i.e. they are not data) [95] is technically not the correct contribution to the objective function. In this situation they believe that theoretically a marginal likelihood function would be more appropriate but solving a marginal likelihood equation would not be computationally practical.

By incorporating effort into the analysis as data, a reduction in the number of parameters is achieved. When effort data is not available, the model estimates Y effective effort parameters. When effort is available as data the model estimates one catchability, reducing the total number of parameters by Y-1.

5.0 COMPARISON OF METHODS

This section will compare the methods described in earlier sections. The abbreviations used to refer to each model are explained in Section 1.2. In particular, attention will be placed on data requirements, what a priori parameters are required, what parameters are estimated and their variances, the mathematical models, assumptions about errors, and parameter estimation methods. Topics described in this section are summarized in Tables 7 and 8.

5.1 Data Requirements

In all cases raw data consists of catch-at-age data. This information can be put into a matrix consisting of A rows (ages) and Y columns (years) (see Table 1). Thus the catch-at-age matrix includes AY catch observations. The DBM, VPA, SPA, CA, FA, and DU methods use the matrix directly, DO and CAGEAN use the natural log of the catch matrix, ME uses the catch ratio matrix, and the PS model uses the natural log of a catch ratio matrix. In both ME and PS methods two catch observations, C(a,y) and C(a+1,y+1), are required to make one catch ratio observation, thus the catch ratio data matrix consists of (A-1) rows and (Y-1) columns, and has (A-1)(Y-1) observations. The PLO model uses the catch-at-age matrix and effort data. The models of FA, DU also use effort data. The CAGEAN model may optionally use effort data. The FA and CAGEAN models may also use fishery independent data if it is available.

5.2 A priori/Initial Parameter Estimates

The DBM and VPA models do not require any initial parameter estimates. In the SPA, CA, and ME models the number of a priori parameters required for each cohort is one terminal fishing mortality for each cohort and an estimate of the natural mortality. To analyze the entire catch-at-age matrix two terminal fishing mortality vectors are required, one for the oldest age in each cohort for all but the last year, $F(A,y)$ $y=1,...,Y-1$ and one for all ages in the last year, $F(a,Y)$ $a=1,...,A$. So to analyze the entire data matrix, $A+Y-1$ terminal F's and one M are required. The DO and CAGEAN models analyze data from the catch-at-age matrix all at once. These methods require an initial guess for all parameters estimated by the solution algorithm. In the DO model these are supplied from the log catch ratio model. The number of initial parameter required are $A-1$ $v(a)$'s (selectivity), $Y-1$ $e(y)$'s (effective effort), and one M for a total of $A+Y-1$. The CAGEAN model is similar in that it needs $A-1$ $v(a)$'s, Y $e(y)$'s, and one M for a total of $A+Y$. Also the age at which fish are fully recruited must be supplied in the DO and CAGEAN models. The PS model analyzes the catch ratio matrix however the number of required a priori parameters is reduced from $(Y-1)$ $f(y)$'s, $(A-1)$ $s(a)$'s, and one M ($= Y+A-1$) to only three, $f(Y)$, $s(a')$, and one M. If the linear PLO model is being fit no initial parameter estimates are required. If the nonlinear PLO model is being fit, guesses for the age-specific catchabilities are required. The FA model requires a guess of an averages F experienced by a fully recruited fish and an estimate of M for a total of two. The DU model requires values for the hazard parameters π and ϕ, and one M for a total of three.

5.3 Parameter Estimates

5.3.1 Number of Parameters Estimated. In the DBM and VPA methods an F(a,y) is estimated for each age/year entry in the catch-at-age matrix except for the last year. Over the entire catch-at-age data matrix there are Y+A-3 cohorts with more than two observations giving (YA)-(Y+A-1) F(a,y)'s. In the SPA and CA models there are the same number of F(a,y)'s estimated. In addition, these two models estimate A-1 N(a,1) and (Y-2) N(1,y) cohort abundances, for cohorts with more than two observations. The ME model is density-independent so it estimates (YA)(Y+A-1) F(a,y)'s. In the DO method all cohorts are analyzed so (A-1) s(a)'s, (Y-1) e(y)'s, and Y+A-1 cohort abundance parameters are estimated. In the PS model all cohorts are analyzed so the parameters estimated are (A-2) s(a)'s, (Y-2) f(y)'s, and (Y-1)+(A-1)-1 (= Y+A-3) cohort abundance parameters. In the linear PLO model the estimated parameters are one M, one q, and (Y+A-3) cohort abundances (for cohorts with more than two observations). In the nonlinear PLO model, parameters from the linear model are supplemented by the number of age-specific catchabilities estimated from the data. In the FA model (assuming only catch and effort are submitted to the model) the estimated parameters are 3 parameters for the VB curve (b1, b2, and w), (Y+A-1) cohort abundance parameters (all cohorts are analyzed), one catchability, and Y annual deviations of F from effort. The DU model (Model I) estimates A s(a)'s, one π, one ϕ, and (Y+A-1) cohort abundance parameters (all cohorts are analyzed). The CAGEAN model estimates (A-1) v(a)'s (assuming one age group is fully recruited), Y e(y)'s, and (Y+A-1) cohort abundances (all cohorts are analyzed).

In the separable models a more restrictive assumption about fishing mortality reduces the number of parameters that need to be estimated from AY to A+Y. This results in a substantial increase in the observation-to-parameter ratio and permits a meaningful goodness-of-fit measure to be calculated. Further reductions in the number of parameters could be realized by reparameterizing mortalities. For example Fournier and Archibald use a three parameter functional relationship to describe the selectivity trend rather than estimate a selectivity parameter for each age.

Because of the flexibility of the FA and CAGEAN models, one needs to keep in mind the serious question as to whether the model has been over-parameterized relative to the information content of the data. Often the parameter estimation algorithm for a highly parameterized model will not converge properly or, because the model is so loosely specified, the model will converge on absurd parameter values. For example, Fournier and Archibald (1982) in their application, attempted to parameterize selectivity similar to Doubleday (i.e. one selectivity parameter for each age). They could not get stable parameter estimates with this approach. When they used their VB parameterization for selectivity, the reduction in parameters was sufficient to allow the parameter estimation procedure to successfully converge. Schnute (1985) provides some astute observations on the topic of model identification.

5.3.2 Bias. To obtain unbiased estimates of parameters from a given mathematical relationship requires knowledge of the variability and sources of error inherent in the data. It is difficult to evaluate if parameter estimates from various methods are biased since the degree of bias will be dependent on the model, the data, and the degree to which the underlying assumptions are violated. Given adequate data (i.e. a long time series of catch-at-age data) it is safe to say that if parameters are assumed to be constant and free from measurement error when in fact they are random variables subject to measurement error, then resulting estimates will be biased. It would seem that methods based on assumptions of constant parameter values over year and/or age are more likely to produce biased estimates. The SPA method has received the most thorough treatment in this regard. Bias could also be due to model misspecification (see below).

One final point regarding bias is that models fitted to logarithmically transformed variables are fitted to the geometric rather than the arithmetic mean and are biased towards low expected values (i.e. $\exp[\mu(\ln(X))] < \mu(X)$). In the DO and CAGEAN models the estimate of population at age is really an estimate of the natural log of the population. Since estimates of population size could range over an order of magnitude, this source of bias could be significant in some applications. This is especially true when computing variables such as total biomass which involve summing a series of age-specific exponential transformations (see Section 3.6.3).

5.3.3 Variances. Interpretation of parameter estimates not accompanied by variances is extremely difficult. Without variances there is no way of knowing the reliability of the parameters. For example, if the confidence limits of a parameter are plus or minus 100% of the actual estimate, then the parameter estimates should be viewed with a great deal of caution. It might be more desirable to accept a parameter estimate known to be slightly biased but accompanied by a reasonable variance estimate than an unbiased estimate with a coefficient of variation greater than one. McDonald and Butler's (1982) suggestion of using a biased estimation procedure such as ridge regression techniques to estimate the parameters of Paloheimo's CPUE model is particularly relevant (also see Paloheimo 1982). Using a biased parameter estimation procedure will give slightly biased parameter estimates however the precision of parameter estimates will be improved.

The DBM, VPA, SPA, ME and CA methods do not allow direct computation of variances in N and F resulting from sampling errors in the catch. Empirical relative variance estimates can be calculated but variances are highly sensitive to estimates of terminal fishing mortality. Other approximate variance estimators for the sequential methods have been proposed (Prager and MacCall 1988; Sampson 1987). The DO, FA, DU, and CAGEAN methods allow calculation of variance estimates since an approximate variance/covariance matrix is available from the nonlinear parameter estimation procedure. In the PLO method variance estimates are available directly from the least squares techniques. In the PS method no variances can be calculated from the procedure directly. However, the ratio of catches will have a higher sampling variance than the catches themselves. Thus even if variance estimates were available, they would probably be larger when compared with variance estimates from the DO or CAGEAN methods.

An key advantage of a separable ASA model is that the user has the opportunity to calculate parameter variances to get some idea of the variability in the results. The variances, however, do not reflect errors in age determination or catch estimates. These types of effects are studied more effectively with a sensitivity analysis or bootstrap procedure (see Section 4.4.3).

5.3.4 Correlation. Correlation tends to underestimate the error structure and overestimate parameter variance. Parameter variances are large since the variance expressions contain significant covariance terms (note: corr(X,Y) =cov(X,Y)/sqrt[σ^2(X)σ^2(Y)]). Doubleday (1976) found that when he analyzed the log catch-at-age matrix fishing mortality and population estimates were negatively correlated (i.e. as F increased N decreased). The problem of correlation between parameters is especially pronounced in the PS method. In this case the raw data matrix is a matrix of log catch ratios, thus successive catch ratios of the same year-class are correlated by year and age in addition to the correlation mentioned earlier. In the DO and PS methods all population numbers at age, N(a,y), can be expressed in terms of the cohorts abundance parameters (i.e. those N(a,y) that occupy the first row and first column of the catch-at-age matrix). Also as the number of parameters in a model increases so does the probability of spurious correlations. For example, if there are 5 ages and 10 years of data, the DO method would estimate 27 parameters. If the probability is 5% that two variables are correlated due to chance alone, then we could expect one random spurious correlation to occur from this data set.

5.4. Goodness-of-fit Measures

When the number of parameters are fewer than the number of observations a useful measure of the goodness-of-fit can be calculated. This measure is available in the separable ASA methods and can be used to describe the amount of the variation in the data explained by the model. In the non-separable methods each parameter estimate is supported by one observation so the observed catches are predicted exactly. The goodness-of-fit measure could be calculated for these methods, but it would be equal to 1.0 (i.e. 100% of the variation in the data is explained by the model) and not very meaningful.

5.5 Mathematical Models

The mathematical model for all methods compared are essentially based on the catch equation and exponential survival model (see Table 7). The major difference between the DBM, VPA, SPA, ME, and CA methods and the DO, PLO, PS, FA, DU, and CAGEAN methods are that the latter six use the more restrictive separability assumption to describe fishing mortality. Even though the DO, PLO, PS, FA, DU, and CAGEAN models use the separable fishing mortality they express this assumption differently. The DO and CAGEAN models express fishing mortality as $F(a,y) = \exp[v(a)+e(y)]$. This can be shown to be equivalent to $F(a,y) = s(a)f(y)$, which is identical to the PS, PLO, and FA expression. In either case the expressions can be made linear by taking logs. The model in the FA and DU methods are probability statements expressed as likelihood functions.

The more general mathematical models of the FA and CAGEAN methods provide a means of investigating model validity. Generality in the mathematical model allows the analyst a great deal of choice in the selection of the model. Dupont (1983) shows how an incorrectly specified model can have dramatic effects on estimation of population abundance. Schnute (1985) provides a lucid discussion on this topic. The generality of the FA, DU, and CAGEAN models also provides the analyst with the capability to iterate toward the "best" model; considering the unique contingencies of the fishery being analyzed, sources of data, and sources of error. At each iteration, the candidate model can be judged against some other model using the value of the objective function as a measure of how well the proposed model fits the data. If the objective function has been reduced, a likelihood ratio test can be performed to determine if the reduction is statistically significant (Schnute 1983).

A small value of the objective function should not be used as the only criteria to judge model validity. For example, sometimes a model will converge on absurd parameter estimates and still report a low objective function value indicating a good fit to the data. A reduction in the objective function should be viewed relative to realistic parameter estimates.

The flexibility of the mathematical model in the CA and FA methods allows a more integrated approach to analyzing catch and effort data. Utilizing more than one data source directly addressed the major deficiency of the separable model -- that catch-at-age data alone are not adequate to reliably estimate stock abundance.

5.6 Consideration of Errors

Table 7 summarizes how the different models incorporate error. The FA, DU, and CAGEAN models are the most generalized with respect to addressing various sources of error. In their comparative study Deriso et al. (1985) considered several different error models and concluded that the CAGEAN model performed similarly despite differences in the model assumptions regarding the source of stochasticity in the data. It seems more important to have the model recognize that there is stochastic error in the data than to know the precise magnitude or source of the error.

5.7 Parameter Estimation Methods

Table 7 summarizes parameter estimation methods used by each method. Generally the early methods used simple approaches and the more recent methods require more complicated techniques. In the methods of DBM, VPA, CA, SPA, and CA the equations are solved sequentially, linking successive age groups. The PLO method uses multiple regression least squares techniques to simultaneously estimate parameters while the DO and nonlinear PLO methods use linearization and multiple regression to simultaneously estimate parameters. The PS method was developed in response to computational problems encountered when trying to use the DB method. The PS method uses both a simultaneous and serial approach to solving for the parameters. Mortality parameters are estimated simultaneously. Cohort abundance estimates are estimated serially and they are conditional on mortality estimates. The CAGEAN method uses nonlinear regression methods to simultaneously estimate parameters. Both FA and DU methods use maximum likelihood methods to simultaneously estimate parameters. The FA model uses a serial estimation approach to fit models of increasing complexity.

The main difference between maximum likelihood and nonlinear regression parameter estimation procedures is that in the former the model is fit to the data and in the later the data are fit to the model. When errors are normally distributed these two procedures are equivalent since maximizing the likelihood objective function is identical to minimizing the sum of squares (least squares) objective function. Maximum likelihood estimates are desirable because under very general conditions they are consistent (converge in probability to the correct value), asymptotically normal, and asymptotically attain the smallest possible variance. The key consideration should be that the parameter estimation procedure should provide a goodness-of-fit statistic and residuals so that the relative merits of various models can be compared and evaluated. Through examining a variety of models the analyst may achieve a greater understanding of the of the important factors that affect the population dynamics of the stock under study. Finally, it should be emphasized that "... no single method has emerged which is the best for the solution of all nonlinear programming problems" (Bard 1974, p. 84).

5.8 Extensions

All models discussed up to this point analyze data from one species only. Stock assessment methods have been extended to work with multispecies data sets. In this type of approach (called Legion Analysis by Pope and Woolner 1981) the natural mortality rate, normally considered a constant in single species ASA models, is partitioned into mortality components associated with inter/intra specific species interactions. These may include predation mortality, starvation mortality, or mortality arising because of competition for limited resources. The data requirements for these mutispecies ASA models (Helgasson and Gislason 1979; Pope 1979a; Majkowski 1981; Pope and Woolner 1981; Lleonart et al. 1985) are especially demanding and include food habits data as well as the usual catch-at-age data. Some theoretical work has been done on multispecies ASA models (Dekker 1982).

ASA methods have also been extended to work with length data instead of catch-at-age data. Jones (1981) shows how standard ASA methods can be adapted to use number-at-length as input data. Pope (1980) extended Jones' length-based analysis to the multispecies situation and called it Phalanx Analysis. Very recently Fournier and Doonan (1987) and Schnute (1987) proposed very flexible length-based population dynamics models.

Work in other areas has also been taking place. Zhang and Sullivan (1987) and Zhang (1988) have proposed a biomass-based version of Pope's Cohort Analysis. Shepherd (1983) recently proposed two measures of overall fishing mortality that remove age class effects. Also time series methods are being applied to ASA models (Kettunen and Hilden 1983; Gudmundsson 1986, 1987; Fournier and Doonan 1987).

6.0 CONCLUSIONS

The early catch-at-age models were quite simple, required several restrictive assumptions and were solved sequentially in a completely deterministic fashion. Newer models, operate on fewer assumptions, permit much more flexibility in the way the models can be formulated and constrained, estimate more parameters and on the whole require far more complex statistical and mathematical procedures than did their predecessors.

Undoubtedly many of the newer techniques will supplant and/or augment older methods since the newer methods offer important advantages. The question still remains as to which method may be the best one to use. This is a difficult question to answer. The analyst needs to 1) carefully consider what data is available, 2) evaluate the sources of error contained in the data and which of these are tractable, 3) consider what fishery-specific contingencies might be relevant towards trying to describe the population dynamics of the exploited stock, and 4) select mathematical models and parameter estimation methods that permit maximum extraction of information from the data. There should be about ten years of data before attempting to use a separable method, although this figure is somewhat arbitrary. If the fishery has undergone a transition a longer time series might be needed so that a time series of adequate length is available over each unique period in the fisheries history. If the fishery has been fairly static a separable method may be usable with less than ten years of data. Whenever possible, diagnostic methods (i.e. residual analysis etc.) should be employed to the fullest extent possible. When selecting a model, follow the rule of parsimony; try not to overparameterize the model relative to the available data.

The widespread availability of high-speed computing resources, especially powerful desktop microcomputers, has drastically changed stock assessment research. In earlier times, simple models and crude approximations were absolutely essential when attempting to derive tractable solutions to complex sets of nonlinear equations. In earlier eras much effort was required to solve the equations. The situation is altogether different in todays research environment. Currently our ability to propose and solve complex systems of nonlinear equations via complicated statistical methods and sophisticated numerical algorithms exceed our biological knowledge base or our data gathering ability. Nonetheless, use of these software tools should be encouraged because it allows the analyst to devote their energies to critical analysis rather than tedious arithmetic.

Several highly specialized computer programs are available to carry out several of the stock assessment analyses described in this paper (Doubleday 1975b; Rivard and Doubleday 1979; Rivard 1982; Archibald et al. 1983a; Megrey and Lynde 1986; Megrey 1987). Also Dupont (1983) and Deriso et al. (1985) report the availability of computer software to implement their models.

ACKNOWLEDGEMENTS

The author would like to express his thanks to William E. Ricker for supplying translated copies of two of Baranov's early papers and to Debbie Dwyer for help in translating sections of Derzhavin's paper. I am also grateful to Richard B. Deriso, Terrance J. Quinn and Phillip R. Neal for allowing me the opportunity to read a copy of their manuscript on the CAGEAN model prior to publication. Finally thanks to Robert C. Francis for introducing me to the subject of stock assessment. The careful and thoughtful reading of this manuscript by two anonymous reviewers led to a much improved paper. This work was supported by the Resource and Conservation Engineering Division of the Northwest and Alaska Fisheries Center of the National Marine Fisheries Service.

LITERATURE CITED

Abramson, N.J. 1971. Computer programs for fish stock assessment. FAO Fisheries Technical Paper 101:1-154.

Agger, P., I. Boetius and H. Lassen. 1971. On errors in the virtual population analysis. ICES CM 1971/H:16 (mimeo).

Agger, P., I. Boetius and H. Lassen. 1973. Error in the virtual population: the effect of uncertainties in the natural mortality coefficient. Journal du Conseil International pour l-exploration de la Mer 35:93.

Allen, K.R. 1966. Some methods for estimating exploited populations. Journal Fisheries Research Board Canada 23:1553-1574.

Allen, K.R. 1969. An application of computers to the estimation of exploited populations. Journal Fisheries Research Board Canada 26:179-189.

Aldenberg, T. 1975. Virtual population analysis and migration: A theoretical treatment. ICES CM 1975/F:32.

Archibald, C.P., D. Fournier and B.M. Leaman. 1983a. A user's guide to Fournier's catch-at-age model. Canadian Manuscript Report of Fisheries and Aquatic Sciences No. 1688, 81 pp.

Archibald, C.P., D. Fournier and B.M. Leaman. 1983b. Reconstruction of stock history and development of rehabilitation strategies for Pacific ocean perch in Queen Charlotte Sound. North American Journal Fisheries Management 3:283-294.

Bajkov, A. 1932. Fish population and productivity of lakes. Transactions of the American Fisheries Society 62:307-316.

Bannerot, S.P. and C.B. Austin. 1983. Using frequency distributions of catch per unit effort to measure fish-stock abundance. Transactions of the American Fisheries Society 112:608-617.

Baranov, F.I. 1918. On the question of the biological basis of fisheries. Nauchn. Issled. Ikhtiol. Inst. Izv. 1:81-128 (In Russian, translated by W.E. Ricker).

Bard, Y. 1974. Nonlinear Parameter Estimation. Academic Press, N.Y., 341 pp.

Beamish, R.J. 1986. Overview. p. 5-11. In: Miles S. Alton (compiler), A workshop on comparative biology, assessment, and management of gadoids from the North Pacific and Atlantic oceans. Unpublished Proceedings. U.S. Department of Commerce, National Marine Fisheries Service, Northwest and Alaska

Fisheries Center, 7600 Sand Point Way NE, BIN C15700, Bldg. 4, Seattle, WA 98115.

Beauchamp, J.B. and J.S. Olson. 1973. Corrections for bias in regression estimates after logarithmic transformations. Ecology 54:1403-1407.

Beverton, R.J.H. 1954. Notes on the use of theoretical models in the study of the dynamics of exploited fish populations (mimeo). U.S. Fisheries Laboratory, Beaufort, N.C., Miscellaneous Contribution 2:1-186.

Beverton, R.J.H. and S.J. Holt. 1957. on the dynamics of exploited fish populations. Fishery Investigations London Ser. 2, 19. 533 pp.

Bishop, Y.M.M. 1959. Errors in estimates of mortality obtained from virtual populations. Journal Fisheries Research Board Canada 16:73-90.

Boiko, E.G. 1934. Estimation of the supply of Kuban zanders. Rab. Dono-Kubanskoi Nauch. Rybokhozyaistvennoi Sta. 1:1-43. (In Russian).

Brownlee, K.A. 1965. Statistical theory and methodology in science and engineering. John Wiley and Sons, NY, 590 pp.

Burd, A.C. and J. Valdivia. 1970. The use of virtual population analysis on the Peruvian anchoveta data. Instituto del Mar, Peru.

Butler, S.A., and L.L. McDonald. 1979. A simulation study of a catch-effort model for estimating mortality rates. Transactions of the American Fisheries Society 108:353-357.

Chapman, D.G. 1961. Statistical problems in dynamics of exploited fish populations. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability. Berkeley and Los Angeles, University of California Press 4:153-168.

Chuganov, N.L. 1935. An attempt at a biostatistical determination of the stocks of fishes in the North Caspian. Ryb. Khoz. 15(6): 24-29; 15(8):17-21. (In Russian).

Collie, J.S. and M.P. Sissenwine. 1983. Estimating population size from relative abundance data measured with error. Canadian Journal of Fisheries and Aquatic Sciences 40:1871-1879.

Conway, G.R., N.R. Glass, and J.C. Wilcox. 1970. Fitting nonlinear models to biological data by Marquardt's algorithm. Ecology 51(3):503-513.

Cooke, J.G. 1985. On the relationship between catch per unit effort and whale abundance. Rep. Int. Whal. Commn. 35:511-519.

Cox, D.R. 1972. Regression models and life tables (with discussion). Journal of the Royal Statistical Society, Series B 34:187-220.

Cushing, D.H. 1975. Marine Fisheries Ecology. Cambridge University Press, Cambridge, London. 278 p.

DeLury, D.B. 1947. On the estimation of biological populations. Biometrics 3:145-167.

Dekker, W. 1982. Theoretical aspects of the multispecies virtual population analysis. ICES CM 1982/ G:32.

Deriso, R.B. 1980. Harvesting strategies and parameter estimation for an age-structured model. Canadian Journal of Fisheries and Aquatic Sciences 37:268-282.

Deriso, R.B., T.J. Quinn II, and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Canadian Journal of Fisheries and Aquatic Sciences 42:815-824.

Deriso, R.B., T.J. Quinn II, and P.R. Neal. 1987. Further aspects of the theory of catch-age analysis with auxiliary information. (unpublished manuscript). International Pacific Halibut Commission, Seattle, Washington, 9 pp.

Derzhavin, A.N. 1922. The stellate sturgeon (Acipenser stellatus Pallas), a biological sketch. Byull. Bakinskoi Ikhtiol. Sta. 1:1-393 (In Russian).

Doubleday, W.G. 1975a. A simple iterative solution to the catch equation. International Commission for the Northwest Atlantic Fisheries Research Document 75/42: 7 pp.

Doubleday, W.G. 1975b. Two computer programs for analysis of catch-at-age data. Fisheries and Marine Service, Research and Development Directorate, Technical Report 520:49 pp.

Doubleday, W.G. 1976. A least squares approach to analysing catch at age data. International Commission for the Northwest Atlantic Fisheries, Research Bulletin 12:69-81.

Doubleday, W.G. 1981. A method of estimating the abundance of survivors of an exploited fish population using commercial fishing catch at age and research vessel abundance indicies. p. 164-178. In: (W.G. Doubleday and D. Rivard, eds.) Bottom Trawl Surveys. Canadian Special Publication in Fisheries and Aquatic Sciences 58.

Draper, N.R. and H. Smith. 1981. Applied Regression Analysis, Second Edition. John Wiley and Sons, N. Y., 709 pp.

Dupont, W.D. 1983. A stochastic catch-effort method for estimating animal abundance. Biometrics 39:1021-1033.

Efron, B. 1982. The jackknife, the bootstrap, and other resampling plans. CBMS-NSF Regional Conference Series in Applied Mathematics SIAM No 38:92 pp.

Fournier, D.A. and I.J. Doonan. 1987. A length-based stock assessment method utilizing a generalized delay-difference model. Canadian Journal of Fisheries and Aquatic Sciences 44:422-437.

Fournier, D. and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Canadian Journal of Fisheries and Aquatic Sciences 39:1195-1207.

Fox, W.W. 1974. An overview of production modelling. National Marine Fisheries Service, Southwest Fisheries Center, Administrative Report Lf-74-10, La Jolla, California.

Fry, F.E.J. 1949. Statistics of a lake trout fishery. Biometrics 5:26-67.

Fry, F.E.J. 1957. Assessment of mortalities by use of the virtual population. Proceedings of the Joint Scientific Meeting ICNAF/ICES/FAO on Fishing Effort, the Effects of Fishing on Resources and the Selectivity of Fishing Gear, Lisbon. Contribution No. P15 (mimeographed).

Garrod, D.J. 1967. Population dynamics of the Arcto-norwegian cod. Journal Fisheries Research Board Canada 24:145-190.

Garrod, D.J. 1976. Catch per unit effort in long range North Atlantic demersal fisheries, and its use in conjunction with cohort analysis. FAO Fisheries Technical Paper No. 155.

Garrod, D.J. 1977. The north Atlantic cod. p. 216-242. In: (J.A. Gulland, ed.) Fish Population Dynamics. John Wiley and Sons, Inc., New York, New York. 372 pp.

Graham, M. 1935. Modern theory of exploiting a fishery and application to North Sea trawling. Journal du Conseil International pour l-exploration de la Mer 10:264-274.

Graham, M. 1939. The sigmoid curve and the overfishing problem. International pour l-exploration de la Mer Rapports et Proces-verbaux des Reunions 110:15-20.

Gray, D.F. 1977. An iterative derivation of fishing and natural mortality from catch and effort data giving measurements of goodness of fit. ICES CM 1977/F:33.

Gray, D.F. 1979. Some extensions to the least squares approach to deriving mortality coefficients. Investigacion Pesquera 43:241-243.

Gudmundsson, G. 1986. Statistical considerations in the analysis of catch-at-age observations. Journal du Conseil International pour l-exploration de la Mer 43:83-90.

Gudmundsson, G. 1987. Time series models of fishing mortality rates. Raunvìsindastofnun Hàskòlans, Science Institute, University of Iceland, Report RH-02-87, Reykjavik, Iceland, 36 pp.

Gulland, J.A. 1965. Estimation of mortality rates. Annex to Arctic Fisheries Working Group Report (meeting in Hamburg, January 1965), ICES CM 1965, Doc. No. 3, 9 pp. (mimeo).

Gulland, J.A. 1977. The analysis of data and development of models. p. 7-95. In: (J.A. Gulland, ed.) Fish Population Dynamics. John Wiley and Sons, Inc. New York, 372 pp.

Hayman, R.A., A.V. Tyler, and R.L. Demory. 1980. A comparison of cohort analysis and catch per unit effort for Dover sole and English sole. Transactions of the American Fisheries Society 109:35-53.

Helgasson, T. and H. Gislason. 1979. VPA-analysis with species interaction due to predation. ICES CM 1979/G:52.

Hoag, S.H. 1978. Trends in abundance and fishing mortality of Pacific halibut, cohort analysis. ICES CM 1978/G:13.

Jones, R. 1961. The assessment of long-term effects of changes in gear selectivity and fishing effort. Marine Research, Department of Agriculture and Fisheries for Scotland 2:1-19.

Jones, R. 1964. Estimating population size from commercial statistics when fishing mortality varies with age. International pour l-exploration de la Mer Rapports et Proces-verbaux des Reunions 155:210-214.

Jones, R. 1981. The use of length composition data in fish stock assessments (with notes on VPA and cohort analysis). FAO Fisheries Circular No. 734, 55 pp.

Kettunen, J. and M. Hilden. 1983. The extended Kalman filter approach to VPA. ICES CM 1983/D:17.

Laurec, A. and J.G. Shepherd. 1983. On the analysis of catch and effort data. Journal du Conseil International pour l-exploration de la Mer 41:81-84.

Lewy, P. 1985. Integrated stochastic virtual population analysis (ISVPA) applying catch, effort and survey information. ICES CM 1985/D:4.

Lleonart, J., J. Salat, and E. MacPherson. 1985. CVPA, an expanded VPA with cannibalism. Application to a hake population. Fisheries Research 3:61-79.

Lotka, A.J. 1925. Elements of Physical Biology. Williams and Wilkins, Baltimore.

McDonald, L.L. and S.A. Butler. 1982. Estimation of fish mortality. Transactions of the American Fisheries Society 111:535-537.

MacCall, A.D. 1976. Density dependence of catchability coefficient in the California Pacific sardine, Sardinops sagax caerula, purse seine fishery. CALCOFI Reports 18:136-148.

MacCall, A.D. 1986. Virtual population analysis (VPA) equations for nonhomogeneous populations, and a family of approximations including improvements on Pope's cohort analysis. Canadian Journal of Fisheries and Aquatic Sciences 43:2406-2409.

Majkowski, J. 1981. Application of a multispecies approach for assessing the population abundance and the age-structure of fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 38:424-431.

Marquardt, D.W. 1963. An algorithm for least squares estimation of non-linear parameters. SIAM Journal 11:431-441.

Megrey, B.A. 1987. A users guide to the Lotus age-structured stock analysis template. Unpublished manuscript, 60 p. Northwest and Alaska Fisheries Center, National Marine Fisheries Service, 7600 Sand Point Way NE, BIN C15700, Bldg. 4, Seattle, WA 98115.

Megrey, B.A. and C.M. Lynde. 1986. A users guide to GENCOH - A generalized cohort analysis program. Unpublished manuscript, 69 p. Northwest and Alaska Fisheries Center, National Marine Fisheries Service, 7600 Sand Point Way NE, BIN C15700, Bldg. 4, Seattle, WA 98115.

Mesnil, B. 1978. Resolution of the catch equation by a simple iterative method. International Commission for the Northwest Atlantic Fisheries, Research Document 78/VI/19: 5 pp.

Miller, D.S. 1977. A simple iterative method to determine fishing mortality rates associated with varying catch levels. International Commission for the Northwest Atlantic Fisheries, Selected Papers 2:7-9.

Mohn, R.K. 1983. Effects of error in catch and effort data on tuning cohort analysis, with a postscript on logistic production models. p. 141-150. In: (W.G. Doubleday and D. Rivard eds.) Sampling Commercial Catches of Marine Fishes and Invertebrates. Canadian Special Publication in Aquatic Sciences 66.

Monastyrsky, G.N. 1935. A method of making long-term forecasts for the vobla fishery of the North Caspian. Ryb. Khoz. 15(5):11-18; 15(6):18-23. (In Russian).

Murphy, G.I. 1965. A solution of the catch equation. Journal Fisheries Research Board Canada 22:191-202.

Norden. R.H. 1972. A survey of maximum likelihood estimation. International Statistics Review 40:329-354.

Norden. R.H. 1973. A survey of maximum likelihood estimation. Part 2. International Statistics Review 41:39-58.

Paloheimo, J.E. 1958. A method of estimating natural and fishing mortalities. Journal Fisheries Research Board Canada 15(4):749-758

Paloheimo, J.E. 1961. Studies on estimation of mortalities. I. Comparison of a method described by Beverton and Holt and a new linear formula. Journal Fisheries Research Board Canada 18(5):645-661.

Paloheimo, J.E. 1980. Estimation of mortality rates in fish populations. Transactions of the American Fisheries Society 109:378-386

Paloheimo, J.E. 1982. Estimation of fish mortality. Transactions of the American Fisheries Society 111:535-537.

Paloheimo, J.E. and L.M. Dickie. 1964. Abundance and fishing success. International pour l-exploration de la Mer Rapports et Proces-verbaux des Reunions 155:152-163.

Peterman, R.M. and G.J. Steer. 1981. Relation between sportfishing catchability coefficients and salmon abundance. Transactions of the American Fisheries Society 110:585-593.

Pope. J.G. 1971. An investigation of the accuracy of the virtual population analysis. International Commission of the Northwest Atlantic Fisheries, Research Document 71/116: 11 pp. (mimeo).

Pope, J.G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. International Commission for the Northwest Atlantic Fisheries, Research Bulletin 9:65-74.

Pope, J.G. 1974. A possible alternative method to virtual population analysis for the calculation of fishing mortality from catch at age data. Annual Meeting International Commission for the Northwest Atlantic Fisheries 1974. Research Document No. 74/20. Serial No. 3166 (mimeographed).

Pope, J.G. 1977. Estimation of fishing mortality, its precision and implications for the management of fisheries. p. 63-76. In: (J.H. Steele ed.) Fisheries Mathematics. Academic Press, New York, 198 pp.

Pope, J.G. 1979a. A modified cohort analysis in which constant natural mortality is replaced by estimates of predation level. ICES CM 1979/H:16.

Pope, J.G. 1979b. Population dynamics and management: current status and future trends. Investigacion Pesquera 43:199-221.

Pope, J.G. 1980. Phalanx analysis: an extension of Jones' length cohort analysis to multispecies cohort analysis. ICES CM 1980/G:19.

Pope, J.G. 1982. Response to the Note by M.S.M. Siddeek. Journal du Conseil International pour l-exploration de la Mer 40:306.

Pope, J.G. and L. Woolner. 1981. A simple investigation into the effects of predation on recruitment to some North Sea fish stocks. ICES CM 1981/G:14.

Pope, J.G. and J.G. Shepherd. 1982. A simple method for the consistent interpretation of catch-at-age data. Journal du Conseil International pour l-exploration de la Mer 40:176-184.

Pope. J.G. and J.G. Shepherd. 1984. On the integrated analysis of catch-at-age and groundfish survey or CPUE data. ICES CM 1984/G:16.

Pope, J.G. and J.G. Shepherd. 1985. A comparison of the performance of various methods for tuning VPAs using effort data. Journal du Conseil International pour l-exploration de la Mer 42:129-151.

Prager, M.H. and A.D. MacCall. 1988. Sensitivities and variances of virtual population analysis as applied to the mackerel, Scomber japanicus. Canadian Journal of Fisheries and Aquatic Sciences 45:539-547.

Ricker, W.E. 1948. Methods of estimating vital statistics of fish populations. Indiana Univ. Publication Science Series 15, 101 pp.

Ricker, W.E. 1971. Derzhavin's biostatistical method of population analysis. Journal Fisheries Research Board Canada 28:1666-1672.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bulletin Fisheries Research Board Canada 191, 382 pp.

Rivard, D. and W.G. Doubleday. 1979. APL programs for stock assessment including a sensitivity analysis with respect to input parameters. Fisheries and Marine Service Technical Report No. 853: 80 p.

Rivard, D. 1982. APL programs for stock assessment (revised). Canadian Technical Report of Fisheries and Aquatic Sciences, 1091: 146 p.

Rivard, D. 1983. Effects of systematic, analytical, and sampling errors on catch estimates: A sensitivity analysis. p. 114-129. In: (W.G. Doubleday and D. Rivard eds.) Sampling Commercial Catches of Marine Fishes and Invertebrates. Canadian Special Publication in Aquatic Sciences 66.

Rothschild, B.J. 1977. Fishing Effort. p. 96-115. In: (J.A. Gulland, ed.) Fish Population Dynamics. John Wiley and Sons, Inc. New York, 372 pp.

Sampson, D.B. 1988. The stability of virtual population analysis cohort size estimates. Journal du Conseil International pour l-exploration de la Mer 44:135-142.

Sampson, D.B. 1987. Variance estimators for virtual population analysis. Journal du Conseil International pour l-exploration de la Mer 43:149-158.

Schaaf, W.F. 1975. Fish population models: potential and actual links to ecological models. p. 221-239. In: (C.S. Russel ed.) Ecological Modelling in a Resource Management Framework. Resources for the Future, Working Paper QF-1, Johns Hopkins University Press, Baltimore, Maryland.

Schaefer, M.B. 1954. Some aspects of the dynamics of populations important to the management of commercial marine fisheries. Bulletin of the Inter-American Tropical Tuna Commission 1:25-56.

Schaefer, M.B. 1957. A study of the dynamics of the fishery for yellowfin tuna in the eastern tropical Pacific Ocean. Bulletin of the Inter-American Tropical Tuna Commission 2:245-285.

Schnute, J. 1983. A new approach to estimating populations by the removal method. Canadian Journal of Fisheries and Aquatic Sciences 40:2153-2169.

Schnute, J. 1985. A general theory for analysis of catch and effort data. Canadian Journal of Fisheries and Aquatic Sciences 42:414-429.

Schnute, J. 1987. A general fishery model for a size-structured fish population. Canadian Journal of Fisheries and Aquatic Sciences 44:924-940.

Schumacher, A. 1970. Bestimmung der fischereilichen Sterblichkeit biem Kabeljaubestand vor Westgrönland. Ber. Dtsch. Komm. Meeresforsch 21(1-4):248-259. (Fisheries Research Board Canada Translation Series No. 1690.)

Seber, G.A.F. 1982. Estimation of animal abundance and related parameters. MacMillian Publishing Co., Inc., New York. 653 pp.

Shepherd, J.G. and M.D. Nicholson. 1986. Use and abuse of multiplicative models in the analysis of fish catch-at-age data. The Statistician 35:221-227.

Shepherd, J.G. 1983. Two measures of overall fishing mortality. Journal du Conseil International pour l-exploration de la Mer 41:76-80.

Siddeek, M.S.M. 1982. A note on Pope's cohort analysis. Journal du Conseil International pour l-exploration de la Mer 40:209-210.

Sims, S.E. 1982a. Algorithms for solving the catch equation forward and backward. Canadian Journal of Fisheries and Aquatic Sciences 39:197-202.

Sims, S.E. 1982b. The effect of unevenly distributed catches on stock-size estimates using virtual population analysis (cohort analysis). Journal du Conseil International pour l-exploration de la Mer 40:47-52.

Sims, S.E. 1984. An analysis of the effect of errors in the natural mortality rate on stock size estimates using virtual population analysis (cohort analysis). Journal du Conseil International pour l-exploration de la Mer 41:149-153.

Tereshchenko, K.K. 1917. The bream (Abramis brama L.) of the Caspian-Volga region, its fishery and biology. Tr. Astrakhan. Ikhtiol. Lab. 4:1-159 (In Russian).

Tomlinson, P.K. 1970. A generalization of the Murphy catch equation. Journal Fisheries Research Board Canada 27:821-825.

Ulltang, O. 1976. Catch per unit of effort in the Norwegian purse seine fishery for Atlanto-Scandian (Norwegian spring spawning) herring. FAO Fisheries Technical Paper 155:91-101.

Ulltang, O. 1977. Sources of errors in and limitations of Virtual Population Analysis (Cohort Analysis). Journal du Conseil International pour l-exploration de la Mer 37:249-260.

Voevodin, I.N. 1938. On the problem of estimating the supply of zanders in the North Caspian. Tr. Pervoi Vsekaspiiskoi Nauch. Rybokhozyaistvennoi Konferentsii 2:3-12. (In Russian).

Widrig, T. 1954. Method of estimating fish populations, with applications to Pacific sardines. Fisheries Bulletin 56:141-166.

Zhang, C.I. 1987. Biology and population dynamics of Alaska Plaice, Pleuronectes quadrituberculatus, in the Eastern Bering Sea. Ph.D. Thesis, University of Washington, Fisheries Department, Seattle, WA, 225 pp.

Zhang, C.I. and P.J. Sullivan. 1988. Biomass-based cohort analysis incorporating growth. Transactions of the American Fisheries Society (in press).

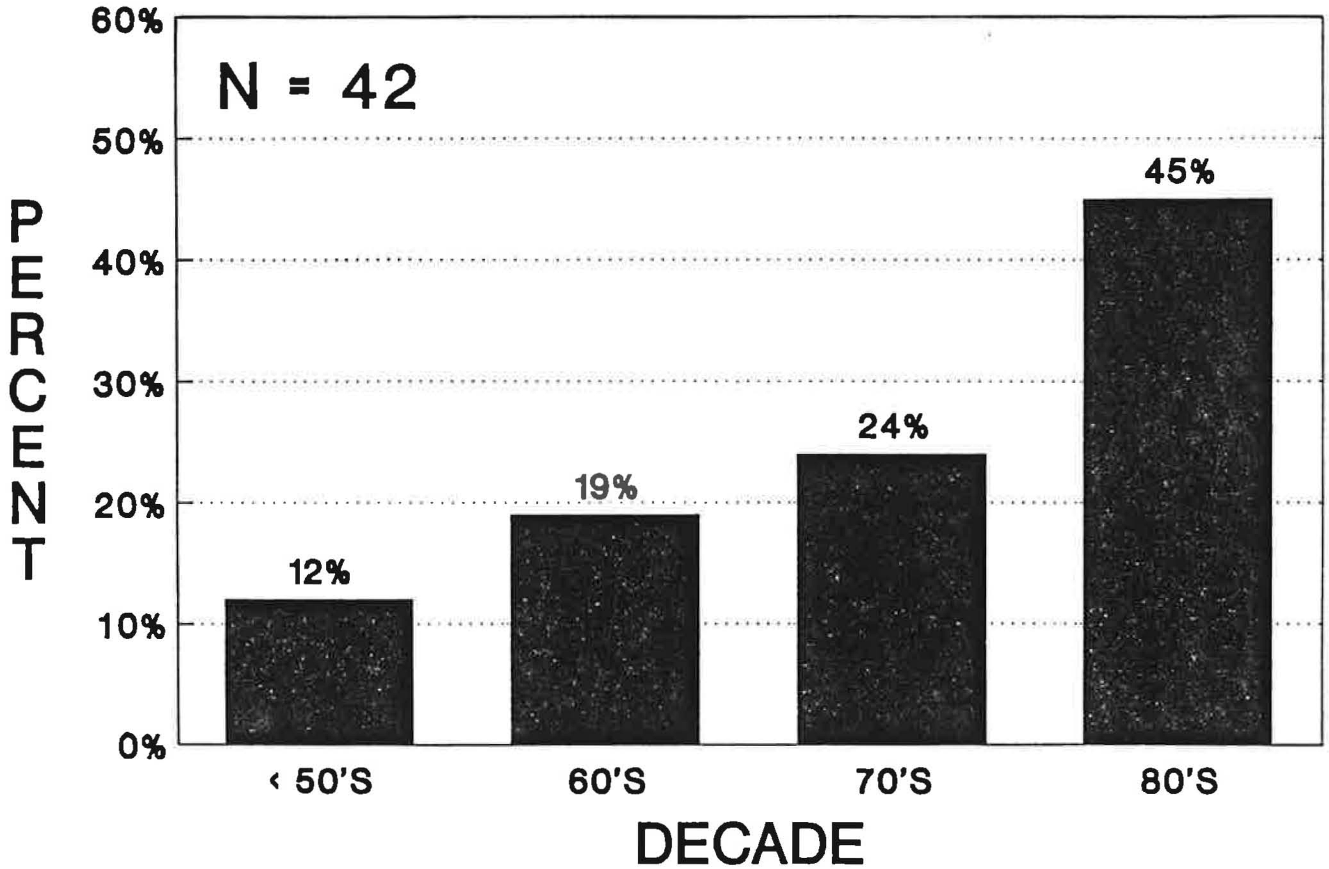

Figure 1. Age-structured stock assessment research activity in the past thirty years.

Table 1. Catch-at-age matrix from a hypothetical fishery.

A B S O L U T E
Y E A R

		1977	1978	1979	1980	1981		
A	2	C(2,77)	C(2,78)	C(2,79)	C(2,80)	C(2,81)	1	R
B	3	C(3,77)	C(3,78)	C(3,79)	C(3,80)	C(3,81)	2	E
S A	4	C(4,77)	C(4,78)	C(4,79)	C(4,80)	C(4,81)	3	L A
O G	5	C(5,77)	C(5,78)	C(5,79)	C(5,80)	C(5,81)	4	A G
L E	6	C(6,77)	C(6,78)	C(6,79)	C(6,80)	C(6,81)	5	T E
U	7	C(7,77)	C(7,78)	C(7,79)	C(7,80)	C(7,81)	6	I
T	8	C(8,77)	C(8,78)	C(8,79)	C(8,80)	C(8,81)	7	V
E	9	C(9,77)	C(9,78)	C(9,79)	C(9,80)	C(9,81)	8	E
		1	2	3	4	5		

R E L A T I V E
Y E A R

Table 2. Catch-at-age matrix from a hypothetical fishery arranged by cohorts.

Y E A R C L A S S

1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
C(9,77)	C(8,77)	C(7,77)	C(6,77)	C(5,77)	C(4,77)	C(3,77)	C(2,77)	C(2,78)	C(2,79)	C(2,80)	C(2,81)
	C(9,78)	C(8,78)	C(7,78)	C(6,78)	C(5,78)	C(4,78)	C(3,78)	C(3,79)	C(3,80)	C(3,81)	
		C(9,79)	C(8,79)	C(7,79)	C(6,79)	C(5,79)	C(4,79)	C(4,80)	C(4,81)		
			C(9,80)	C(8,80)	C(7,80)	C(6,80)	C(5,80)	C(5,81)			
				C(9,81)	C(8,81)	C(7,81)	C(6,81)				

Table 3. Definition of symbols, indicies and notation.

Symbol	Definition
Indicies:	
a	age-class index (a=1,...,A) for a relative index scheme (a=r,...,s) for an absolute index scheme
a'	arbitrary reference age
amax	the oldest age in a cohort
amin	the youngest age in a cohort
A	total number of age-classes in the catch-at-age matrix (A=s-r+1)
fy	first year of fisheries data
g	parameter group or strata index (g=1,...,G)
G	total number of groups or strata
j	time interval (j=1,...,Y-1)
j(k)	time interval relative to cohort k
k	cohort index (k=1,...,K)
K	total number of cohorts in the catch-at-age matrix (K=A+Y-1)
ly	last (most recent) year of fisheries data
n	total number of catch observations in the catch-at-age matrix (n=AY)
r	the age at which an age-class first recruits to the fishery; the youngest exploited age-class
s	the oldest exploited age-class
t	time (years)
$t_{j(k)}$	time that individuals from cohort k are first caught
y	calendar year index (y=1,...,Y) for a relative index scheme (y=fy,...,ly) for an absolute index scheme
ymax	the last (most recent) year of catch data in a cohort

Table 3 (con't). Definition of symbols, indicies and notation.

Symbol	Definition
Indicies:	
ymin	the first year of catch data in a cohort
Y	total number of years of fisheries data in the catch-at-age matrix (Y=ly-fy+1)
Parameters, Observations, and Symbols:	
a(w)	VB parameterization of the selectivity trend with age relationship
b_1, b_2	parameters of the selectivity trend with age function
b(w)	selectivity trend with age curve
B(y)	total population biomass in year y
C(a,y)	catch of fish from age-class a in year y
C(y)	catch of all ages in year y
CB(y)	catch biomass (all ages) in year y
d(a)	age-specific deviation from average catchability
d	realization of the random variable D
d_k	realization of the random variable D_k
d_{kj}	realization of the random variable D_{kj}
$\|d_k\|$	total catch from cohort k
D_{kj}	catch from cohort k in time interval j
D_k	catch vector from cohort k
D	total catch vector
e(y)	log (base e) of f(y)
exp	exponential function
E(a,y)	exploitation fraction on age a fish in year y (=F(a,y)/Z(a,y) (1-exp[-Z(a,y)]))
f( )	functional relationship

Table 3 (con't). Definition of symbols, indicies and notation.

Symbol	Definition
Parameters, Observations, and Symbols:	
f(y)	full recruitment fishing mortality or effective effort coefficient in year y
fec(a)	fecundity of age group a
F(a,y)	annual instantaneous fishing mortality for fish in age-class a in year y
G(a,y)	true age composition of age a fish in year y
H(a,y)	estimated age composition; percent of the total catch observed to be of age a in year y from a random sample of the catch
ln [ ]	natural log function
L_c	penalty weight for catch data
L_f	penalty weight for effort data
L_{sr}	penalty weight for spawner-recruit data
M	annual instantaneous rate of natural mortality (assumed constant for all a and y)
N(a,y)	abundance (number) of age-class a at the beginning of year y
N(y)	total population abundance (numbers) at the beginning of year y
N_k	size of cohort k at time $t_{j(k)}$
$N_k(t)$	size of cohort k at time t
O(y)	estimate of the total number of fish taken in year y
p(a,i)	probability that a fish from age class i is judged during the aging determination to be age a
p_{kj}	probability that a fish from cohort k is caught in the j th time interval
p_k	probability that a fish from cohort k is caught in any time interval
P	availability of the exploited stock to the fishery

Table 3 (con't). Definition of symbols, indicies and notation.

Symbol	Definition
<u>Parameters</u>,	<u>Observations</u>, <u>and</u> <u>Symbols</u>:
q	average catchability coefficient
q(y)	catchability (time-dependent) in year y
$r_k(t)$	probability that a member of cohort k is alive at time t
R(r,y)	number of individuals in age-class r newly recruited the fishery at the beginning of year y
Re(a)	residual for age summed over years
Re(y)	residual for year summed over ages
s(a)	selectivity coefficient for age-class a
s(a,g)	selectivity coefficient for group g age-class a
S(a,y)	survival rate for age a fish in year y (S(a,y)=exp[-Z(a,y])
SP(y)	number of eggs produced by the mature spawning stock in year y
v(a)	log (base e) of s(a)
v(a,g)	log (base e) of s(a,g)
V(a,y)	virtual population of age a fish in year y
w	parameter in nonlinear scaling of the age index in the VB parameterization
W(a)	average weight of an individual in age-class a
We(a)	empirical weighting factor for age
We(y)	empirical weighting factor for year
z(a,g)	dummy variable for stratified selectivity parameterization
z(y)	dummy variable for stratified effort parameterization
Z(a,y)	annual instantaneous rate of total mortality for fish in age-class a in year y [Z(a,y)=F(a,y)+M]

Table 3 (con't). Definition of symbols, indicies and notation.

Symbol	Definition
Parameters,	**Observations, and Symbols:**
$\epsilon_c(a,y)$	random variable for catch distributed $N(0,\sigma_c^2)$ with constant variance.
$\epsilon_f(y)$	random variable for fishing effort in year y distributed $N(0,\sigma_f^2)$ with constant variance.
$\epsilon_{sr}(y)$	random variable for recruitment in year y distributed $N(0,\sigma_{sr}^2)$ with constant variance.
σ_c^2	variance of the catch random variable
σ_f^2	variance of the effort random variable
σ_{sr}^2	variance of the spawner-recruit random variable
$\pi_k(t)$	fishing mortality hazard parameter for cohort k
$\phi_k(t)$	natural mortality hazard parameter for cohort k
α	density-independent Ricker spawner recruit parameter
β	density-dependent Ricker spawner recruit parameter
θ	population density parameter of catchability submodel
Φ	gear saturation parameter of catchability submodel

Table 4. Summary of the assumptions underlying each of eleven ASA methods.

ASSUMPTIONS	DBM	VPA	SPA	ME	CA	DO	PS	PLO	FA	DU	CAGEAN
I. DATA SOURCES:											
1) The age composition of the stock is not constant from year to year. The H(a,y) are obtained by aging a random sample of the total catch. The random variables H(a,y) and O(y) are independent.	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
2) There are errors associated with estimating age composition percentages. Errors in estimating percentage at age have unequal variances and are correlated.	N	N	N	N	N	N	N	N	Y	N	N
3) There are errors associated with estimating total catch	N	N	N	N	N	Y	Y	Y	Y	N	Y
4) A long series of well sampled catches are required.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
II. DETERMINISTIC MORTALITY MODELS:											
1) All significant mortality components are accounted for in F and/or M	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
2) Natural mortality does not vary with respect to age, year and size of the stock. It represents all mortality components not associated with a directed fishery (i.e. predatory mortality, starvation mortality, incidental catch mortality). .	NA	NA	Y	Y	Y	Y	Y	Y	Y	Y	Y
3) Fishing mortality does not vary with respect to size of stock .	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
4) There are no random components in F and/or M.	Y	Y	Y	Y	Y	Y	Y	Y	N	N	O
5) Fishing and natural mortality operate concurrently and are independent of one another (Type II fishery, Ricker (1975, p. 11)).	NA	NA	Y	Y	Y	Y	Y	Y	Y	Y	Y
6) Natural mortality is not a significant mortality component; all removals from the population are accounted for in the catch (i.e. M=0).	Y	Y	N	N	N	N	N	N	N	N	N
7) Natural mortality is known or can be estimated independently. .	NA	NA	Y	Y	Y	Y	Y	NA	Y	Y	Y
III. SEPARABLE FISHING MORTALITY MODEL:											
1) The fishing mortality rate can vary between years and within any one year it can vary with age	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
2) The variation in F can be represented as a product of an age and year factor	N	N	N	N	N	Y	Y	Y	Y	Y	Y
3) Age-specific selectivity factors are constant for each age group over all years.	N	N	N	N	N	Y	Y	Y	Y	O	O
4) Year-specific exploitation pattern can change between years but within any one year it is constant.	N	N	N	N	N	Y	Y	Y	Y	Y	Y
IV. EFFORT/CATCHABILITY FISHING MORTALITY MODELS:											
1) The catchability of the gear is constant and does not vary by age and year. That is, one unit of fishing effort catches the same percentage of the stock. This assumption holds regardless of when and where the effort is applied	NA	NA	NA	NA	NA	NA	NA	Y	O	Y	Y

Y - Yes; N - No; O - Optional; NA - Not Applicable

Table 4 (con't). Summary of the assumptions underlying each of eleven ASA methods.

ASSUMPTIONS	DBM	VPA	SPA	ME	CA	DO	PS	PLO	FA	DU	CAGEAN
IV. EFFORT/CATCHABILITY FISHING MORTALITY MODELS:											
2) Units of effort are independent of how much effort there already is or of effort on other fish stocks (i.e there is no gear interference or saturation (Rothschild 1977)	NA	NA	NA	NA	NA	NA	NA	Y	Y	Y	Y
3) There has been no change in gear or vessel efficiency (fishing power) over the period that fisheries data are available. . . .	NA	NA	NA	NA	NA	NA	NA	Y	Y	Y	Y
4) The relationship between fishing effort and fishing mortality is independent of stock size.	NA	NA	NA	NA	NA	NA	NA	Y	O	Y	Y
V. CATCH EQUATION MODEL:											
1) The population is closed to immigration and emigration.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
2) The fishery is operating on a single unit stock over its entire geographic range	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3) The main human influence on the abundance of a group of recruits is the amount of fishing	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
4) All removals from the population are accounted for in the catch (resulting from directed effort), except for losses due to natural mortality.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
5) The entire catch is taken exactly half way through the year . .	N	N	N	N	Y	N	N	N	N	N	N
VI. STOCHASTIC MORTALITY MODEL:											
1) The probability that a fish dies from fishing (i.e. is caught) in the interval (t,t+dt) is $\pi_k(t)+o(dt)$	NA	NA	NA	NA	NA	NA	NA	NA	NA	Y	NA
2) The probability that a fish dies from natural causes in the interval (t,t+dt) is $\phi_k(t)+o(dt)$.	NA	NA	NA	NA	NA	NA	NA	NA	NA	Y	NA
3) The probability that two or more fish dies in the interval (t,t+dt) is o(dt)	NA	NA	NA	NA	NA	NA	NA	NA	NA	Y	NA
4) The catch hazard is proportional to the product of effort and selectivity	NA	NA	NA	NA	NA	NA	NA	NA	NA	Y	NA
VII. MISCELLANEOUS:											
1) Values for the penalty weights are fixed.	NA	NA	NA	NA	NA	NA	NA	NA	Y	NA	Y
2) Recruitment is described by a Ricker spawner-recruit function multiplied by a log-normal random error term	NA	NA	NA	NA	NA	NA	NA	NA	Y	NA	Y

Y - Yes; N - No; O - Optional; NA - Not Applicable

Table 5. Summary of the advantages for each of eleven ASA methods.

ADVANTAGES	DBM	VPA	SPA	ME	CA	DO	PS	PLO	FA	DU	CAGEAN
I. THE MODEL:											
- Information submitted to the model are subject to error.	N	N	N	N	N	Y	Y	Y	Y	Y	Y
- The mathematical model is generalized and extremely flexible . .	N	N	N	N	N	N	N	N	Y	Y	Y
- The separability assumption results in a substantial reduction in the number of parameters that need to be estimated.	N	N	N	N	N	Y	Y	Y	Y	Y	Y
- Data from different cohorts are linked	N	N	N	N	N	Y	Y	Y	Y	Y	Y
- Rather than assuming that natural mortality is known, the catch/effort model allows this parameter to be estimated	N	N	N	N	N	N	N	Y	O	O	O
- No assumptions are required regarding catchability	Y	Y	Y	Y	Y	Y	Y	N	N	N	O
- The log transform makes the equation more nearly linear; Also heteroscedasticity in the catch error variance is removed. . . .	N	N	N	N	N	Y	Y	Y	Y	N	Y
- Arbitrary choice of the terminal fishing mortality is removed. .	NA	NA	N	N	N	Y	Y	Y	Y	Y	Y
- Capable of predicting future catches with error bounds	N	N	N	N	N	N	N	N	N	Y	N
- Using a competing risk approach to modeling mortalities provides greater flexibility in representing each mortality factor as a stochastic process	N	N	N	N	N	N	N	N	N	Y	N
- Ability to integrate many sources of data into a simultaneous parameter estimation scheme.	N	N	N	N	N	N	N	N	Y	N	Y
- Ability to incorporate fishery-independent data simultaneously into the estimation procedure precludes the need for ad hoc tuning methods	N	N	N	N	N	N	N	N	Y	N	Y
- Ability to stratify data by gear type or vessel type increases the number of observations submitted to the analysis	N	N	N	N	N	N	N	N	N	Y	Y
II. PARAMETER ESTIMATION METHOD/SOLUTION ALGORITHM:											
- A measure of the variation explained by the model is available .	N	N	N	N	N	Y	Y	Y	Y	Y	Y
- Variance estimates of the parameters are available so a determination as to their reliability can be made.	N	N	N	N	Y	Y	N	Y	Y	Y	Y
- The variance/covariance matrix is available to examine correlation between the independent and dependent variables. . .	N	N	N	N	N	Y	N	Y	Y	Y	Y

Y - Yes; N - No; O - Optional; NA - Not Applicable

Table 5 (con't). Summary of the advantages for each of eleven ASA methods.

ADVANTAGES	DBM	VPA	SPA	ME	CA	DO	PS	PLO	FA	DU	CAGEAN
II. PARAMETER ESTIMATION METHOD/SOLUTION ALGORITHM:											
- Use of statistical methods permits use of many diagnostic methods to better evaluate the assumptions of the model.	N	N	N	N	N	Y	N	Y	Y	Y	Y
- All parameters in the model are estimated simultaneously	N	N	N	N	N	Y	N	Y	Y	Y	Y
- An efficient flexible parameter estimation algorithm is used . .	N	N	N	N	N	N	N	N	Y	Y	Y
- The maximum likelihood parameter estimation algorithm is powerful and generalizes easily.	N	N	N	N	N	N	N	N	Y	Y	N
- Use of the bootstrap variance estimation method permits evaluation of statistical bias and statistical uncertainty in the parameter estimates.	N	N	N	N	N	N	N	N	N	N	Y
III. GENERAL:											
- The method is extremely easy to carry out.	Y	Y	Y	Y	Y	N	N	N	N	N	N
- The analysis provides a minimum estimate of population abundance and rates of exploitation with a minimum of data . . .	Y	Y	N	N	N	N	N	N	N	N	N
- The concept of estimating a population by summing the catches is intuitively attractive.	Y	Y	NA	NA	NA	NA	NA	NA	NA	NA	NA
- The method is independent of errors associated with measures of CPUE.	Y	Y	Y	Y	Y	Y	Y	N	N	N	O
- It is very valuable in understanding a fishery in a historic sense, for explaining its population dynamics, and is potentially of great value in showing up large, and possible detrimental, changes in fishing mortality soon after they have happened.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
- Estimates of F can be used to test more effectively the proportionality of F to effort (i.e. F=qf) and the validity of CPUE data (Garrod 1976; Hayman et al. 1980)	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y
- The method is fairly robust to violations of underlying assumptions.	Y	Y	Y	Y	Y	Y	N	N	N	N	N
- All parameters in the model are estimated simultaneously	N	N	N	N	N	Y	Y	Y	Y	Y	Y

Y - Yes; N - No; O - Optional; NA - Not Applicable

Table 6. Summary of the disadvantages for each of eleven ASA methods.

DISADVANTAGES	DBM	VPA	SPA	ME	CA	DO	PLO	PS	FA	DU	CAGEAN
I. DATA:											
· Errors in the data are not considered (i.e. the method is not stochastic).	Y	Y	Y	Y	Y	N	N	N	N	N	N
· The assumption of constant interannual age composition is strong	Y	N	N	N	N	N	N	N	N	N	N
· A long series of catch-at-age observations are needed.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
II. THE MODEL:											
· No provision is made for natural mortality	Y	Y	N	N	N	N	N	N	N	N	N
· The assumption of constant natural mortality is extremely strong	NA	NA	Y	Y	Y	Y	Y	Y	Y	Y	Y
· The assumption that separability at age does not change with time may often not be valid.	NA	NA	NA	NA	NA	Y	Y	Y	Y	Y	Y
· Catch-at-age data are analyzed one cohort at a time.	Y	Y	Y	Y	Y	N	N	N	N	N	N
· Catch-at-age information alone are not sufficient to reliably estimate stock abundance because fishing mortality and stock size are highly negatively correlated (Doubleday 1976; Pope 1977) (i.e. solutions are not unique)	NA	NA	Y	Y	Y	Y	Y	Y	Y	Y	Y
· The natural mortality and catchability parameters are strongly negatively correlated.	N	N	N	N	N	N	Y	N	Y	N	O
· The model is sensitive to trends in catchability with effort or time.	N	N	N	N	N	N	Y	N	Y	N	O
· If data from just one cohort is being analyzed the linear model will not do well at estimating the year class abundance parameters since this parameter is highly correlated with the catchability parameter (Butler and McDonald 1979).	N	N	N	N	N	N	Y	N	N	N	N
III. PARAMETER ESTIMATION METHOD/SOLUTION ALGORITHM:											
· Because the number of parameters equals the number of data points, there is no measure of the variability about the parameter estimates nor a measure of the amount of variation in the data explained by the model	Y	Y	Y	Y	Y	N	N	N	N	N	N

Y - Yes; N - No; O - Optional; NA - Not Applicable

Table 6 (con't). Summary of the disadvantages for each of eleven ASA methods.

DISADVANTAGES	DBM	VPA	SPA	ME	CA	DO	PLO	PS	FA	DU	CAGEAN
III. PARAMETER ESTIMATION METHOD/SOLUTION ALGORITHM:											
• Even when catches are explained well (as measured by a low residual sum of squares or a high R-squared value), parameter estimates have large variances and wide confidence intervals . .	N	N	N	N	N	Y	Y	Y	Y	Y	Y
• The variance of the random variable, predicted catch, is assumed approximately independent of the actual magnitude of catch . . .	N	N	N	N	N	Y	Y	Y	Y	Y	Y
• Variance estimates of s(a) and f(y) are difficult because of the parameter estimation algorithm. Variance estimates of population at age are conditional on estimates of s(a) and f(y)	NA	NA	NA	NA	NA	N	N	Y	N	N	N
• The linearization solution algorithm often does not monotonically converge	NA	NA	NA	NA	NA	Y	Y	NA	NA	NA	NA
• There is no guarantee of a global minimum and different starting values may produce different solutions.	NA	NA	NA	NA	NA	Y	Y	N	Y	Y	Y
• Maximum likelihood estimates do not possess any optimal properties for small samples	N	N	N	N	N	N	N	N	Y	Y	N
IV. GENERAL:											
• Population estimates are minimum estimates since natural mortality is not accounted for	Y	Y	N	N	N	N	N	N	N	N	N
• Estimates of F for the most recent year are the least accurate .	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
• Relative strength of strong and weak cohorts will be biased if M varies with cohort strength (Ulltang 1977).	NA	NA	Y	Y	Y	Y	Y	Y	Y	Y	Y
• If trends exist in natural mortality (say decreasing with age) bias in parameter estimates results since actual increases in natural mortality (when M was erroneously assumed to be constant) would show up as increasing fishing mortality (Ulltang 1977)	NA	NA	Y	Y	Y	Y	Y	Y	Y	Y	Y
• The limited accuracy of Pope's approximation would preclude application of cohort analysis to stocks that are exploited heavily ($F > 1.2$) or those with high levels of natural mortality ($M > 0.3$), but only if catch data are available in yearly intervals	NA	NA	NA	NA	Y	NA	NA	NA	NA	NA	NA
• The complexity of the model essentially requires a computer program to carry out the analysis.	N	N	N	N	N	Y	Y	Y	Y	Y	Y

Y - Yes; N - No; O - Optional; NA - Not Applicable

Table 7. Summary of the mathematical models, error models, parameter estimation method, and parameter estimation sequence underlying each of eleven ASA methods.

ITEM	DBM	VPA	SPA	ME	CA	DO	PS	PLO	FA	DU	CAGEAN
I. MATHEMATICAL MODELS:											
1) Variation on Catch Equation and Exponential Survival Model. . .	Y	Y	N	N	Y	N	N	N	N	N	N
2) SPA Model	N	N	Y	N	N	N	N	N	N	N	N
3) Catch Ratio Model	N	N	N	Y	N	N	N	N	N	N	N
4) Log Catch Ratio Model	N	N	N	N	N	N	Y	N	Y	N	N
5) Log Catch Model	N	N	N	N	N	Y	N	N	N	N	Y
6) Log CPUE Model.	N	N	N	N	N	N	N	Y	N	N	N
7) Specialized Stochastic Model.	N	N	N	N	N	N	N	N	Y	Y	N
II. ERROR MODEL:											
1) Measurement of Catch Error.	N	N	N	N	N	Y	Y	Y	Y	N	Y
2) Measurement of Effort Error	N	N	N	N	N	N	N	N	Y	N	Y
3) Measurement of Spawner-recruit Error.	N	N	N	N	N	N	N	N	O	N	O
4) Stochastic Mortality.	N	N	N	N	N	N	N	N	N	Y	N
III. PARAMETER ESTIMATION METHOD:											
1) Simple Data Transformations	Y	Y	N	N	Y	N	Y	N	N	N	N
2) Root Finding (Newton-Raphson) Iterative Approach.	N	N	Y	Y	N	N	Y	N	N	N	N
3) Linearization	N	N	N	N	N	Y	N	Y	N	N	N
4) Multiple Linear Least Squares Regression.	N	N	N	N	N	Y	N	Y	N	N	N
5) Nonlinear Least Squares Regression.	N	N	N	N	N	Y	N	Y	N	N	Y
6) Maximum Likelihood.	N	N	N	N	N	N	N	N	Y	Y	N
7) Bayesian Estimation	N	N	N	N	N	N	N	N	O	N	O
IV. PARAMETER ESTIMATION SEQUENCE:											
1) Sequential.	Y	Y	Y	Y	Y	N	N	N	N	N	N
2) Simultaneous.	N	N	N	N	N	Y	Y	Y	Y	Y	Y
3) Serial.	N	N	N	N	N	N	Y	N	Y	N	N

Y - Yes; N - No; O - Optional

Table 8. Data requirements, parameters estimated, and required initial starting values for each of eleven ASA methods.

	DBM	VPA	SPA	ME	CA	DO	PS	PLO	FA	DU	CAGEAN
I. DATA REQUIRED:											
- Catch by age and year	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
- Total catch by year	N	N	N	N	N	N	N	N	Y	N	N
- Effort by year	N	N	N	N	N	N	N	Y	Y	Y	O
- Fecundity by age	N	N	N	N	N	N	N	N	O	N	O
- Aging error	N	N	N	N	N	N	N	N	O	N	N
- Catch estimate error	N	N	N	N	N	Y	Y	Y	Y	Y	Y
II. PARAMETERS ESTIMATED:											
- Fishing mortality by age and year, F(a,y)	Y	Y	Y	Y	Y	N	N	N	N	N	N
- Catchability, q	N	N	N	N	N	N	N	Y	Y	O	O
- Catchability by age, q(a)	N	N	N	N	N	N	N	O	N	N	N
- Catchability by year, q(y)	N	N	N	N	N	N	N	N	O	N	N
- Density-dependent catchability, θ	N	N	N	N	N	N	N	N	O	N	N
- Effort by year, f(y)	N	N	N	N	N	Y	Y	N	O	O	O
- Population abundance by age and year, N(a,y)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
- Abundance of recruits, R(r,y) and R(a,fy)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
- Natural mortality, M	N	N	N	N	N	N	N	Y	O	O	O
- Selectivity by age, s(a)	N	N	N	N	N	Y	Y	N	Y	O	Y
- Selectivity trend with age	N	N	N	N	N	N	N	N	O	N	N
- Spawner-recruit, α and β	N	N	N	N	N	N	N	N	O	N	O
- Hazard mortality, π and ϕ	N	N	N	N	N	N	N	N	N	Y	N
- Annual deviations of F from effort	N	N	N	N	N	N	N	N	Y	N	N
III. INDEPENDENT INITIAL ESTIMATES REQUIRED:											
- Natural mortality	Y	Y	Y	Y	Y	Y	Y	N	O	O	O
- Terminal F, each cohort	Y	Y	Y	Y	Y	N	N	N	N	N	N
- Selectivity at reference age	N	N	N	N	N	Y	Y	N	N	N	Y
- Catchability estimates for other ages	N	N	N	N	N	N	N	Y	N	N	N
- Effort for last year	N	N	N	N	N	Y	Y	N	N	N	N

Y - Yes; N - No; O - Optional