ORNL/TM-11061

ornl

OAK RIDGE NATIONAL LABORATORY

MARTIN MARIETTA

A Computer Program for Estimating Fish Population Sizes and Annual Production Rates

S. F. Railsback
B. D. Holcomb
M. G. Ryon

Environmental Sciences Division
Publication No. 3245

OPERATED BY
MARTIN MARIETTA ENERGY SYSTEMS, INC.
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

ORNL/TM-11061

ENVIRONMENTAL SCIENCES DIVISION

A COMPUTER PROGRAM FOR ESTIMATING FISH POPULATION SIZES AND ANNUAL PRODUCTION RATES

S. F. Railsback
B. D. Holcomb[1]
M. G. Ryon[2]

[1]Computing and Telecommunications Division
[2]Health and Safety Research Division

Environmental Sciences Division
Publication No. 3245

Date Published - October 1989

Prepared by the
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831-6285
operated by
MARTIN MARIETTA ENERGY SYSTEMS, INC.
for the
U.S. DEPARTMENT OF ENERGY
under contract DE-AC05-84OR21400

CONTENTS

ACKNOWLEDGMENTS

The fish production software that this report documents was produced for the Biological Monitoring and Abatement Program (BMAP), a requirement of the National Pollutant Discharge Elimination System permits issued for each of the three DOE Oak Ridge facilities. The BMAP is managed by Jim Loar, who also provided reviewers and direction for this work. The authors of this report wish to acknowledge the authors of the methods incorporated in the software (Carle and Strub, 1978; Garman and Waters, 1983; Newman and Martin, 1983; Zippin, 1956). Greg Garman of Virginia Commonwealth University and Raymond Newman of the University of Minnesota were very helpful in answering questions about the methods and in reviewing the work. Larry Barnthouse and Virginia Tolbert of ORNL reviewed the document, and Tommy Wright of the ORNL Engineering Physics and Mathematics Division provided statistical advice.

This work was funded by the Department of Environmental Management of the Health, Safety, Environmental, and Accountability Division, Oak Ridge Y-12 Plant. The Y-12 Plant and Oak Ridge National Laboratory are operated by Martin Marietta Energy Systems, Inc., under contract DE-AC05-84OR21400 with the U.S. Department of Energy.

ABSTRACT

STEVEN F. RAILSBACK, BRADY D. HOLCOMB, and MICHAEL G. RYON. 1989. A Computer Program for Estimating Fish Population Sizes and Annual Production Rates. ORNL/TM-11061. Oak Ridge National Laboratory, Oak Ridge, Tennessee. 68 pp.

This report documents a program that estimates fish population sizes and annual production rates in small streams from multiple-pass sampling data. A maximum weighted likelihood method is used to estimate population sizes (Carle and Strub, 1978), and a size-frequency method is used to estimate production (Garman and Waters, 1983).

The program performs the following steps: (1) reads in the data and performs error checking; (2) where required, uses length-weight regression to fill in missing weights; (3) assigns length classes to the fish; (4) for each date, species, and length class, estimates the population size and its variance; (5) for each date and species, estimates the total population size and its variance; and (6) for each species, estimates the annual production rate and its variance between sampling dates selected by the user. If data from only one date are used, only populations are estimated.

The program requires a parameter file containing parameters used by the program for each species, and a data file containing the field data. The parameter file must be named FISHPARM.DAT and is unformatted (the data need not be in specific columns). The first line of the parameter file contains the number of length classes per length class definition and the number of alternative length class definitions (the user assigns one of these definitions to each species). The file then contains one line for each length class definition, with each line containing the upper limit, in centimeters, of each length class. The remainder of the parameter file has one line for each species, containing the species code (as appears in the data file, with a maximum of six characters, in single quotes); the Cohort Production Interval, which is an estimate of the average maximum life span of the species in years; and the length class definition to be used (i.e., which of the previously defined sets of length classes to use for the

species). The length class definition should be selected so that the largest fish in the data file is not longer than the largest length class.

The program prompts the user for the name of the field data file. The field data file is formatted and contains one line per individual fish caught. The format is as follows.

Column(s)	Parameter
1-6	Date, in format MMDDYY
8-14	Site, in up to seven alphanumeric characters
16	Pass number on which the fish was caught
18-23	Species code, in up to six alphanumeric characters
24-28	Length in centimeters
29-35	Weight, in grams

The program requires the user to interactively select the site to be processed, the stream area or length sampled each date, and the combinations of sampling dates for which the production should be calculated. The program issues error or warning statements if any of the following problems occur: (1) the length of a fish is missing, zero, or negative; (2) the limits set in the program are exceeded for number of sampling dates, number of total fish in the data file, or number of species; (3) a species in the data file is not in the parameter file; (4) the weight of a fish is missing and there are insufficient data (less than two fish with lengths and weights) to estimate the weight by length-weight regression; (5) a pass number is missing; (6) a fish has a length exceeding the largest length class; (7) the population estimation method does not find a solution within the maximum allowable number of iterations (this occurs rarely when more fish were caught in later passes than in the first one); and (8) a species is found at a site on only one date, so production cannot be estimated. If only one fish is in a length class, the program attempts to move the fish into an adjacent length class to allow estimation of the variance in the production estimate.

The program output is written to a file called FISHPROG.OUT. The production output is the estimated annual production rate in grams per year per unit stream area or length and its variance. This production value is averaged over the time period for which data were used.

1. INTRODUCTION

This report describes the formulation and use of a computer program that estimates fish population sizes and annual production rates from multiple-pass sampling data. The program is designed to use multiple-pass removal method sampling data for stream fish collected two or more times in each of one or more years. The program separates data by sampling date, sampling site, and species. The program uses the maximum weighted likelihood method of Carle and Strub (1978) to estimate fish population sizes, and the size-frequency (Hynes) method for estimating annual production rates (Garman and Waters, 1983). The program calculates secondary production, defined as "the total tissue elaborated by a heterotrophic population regardless of the fate of that tissue" (Garman and Waters, 1983), for each species averaged between sampling dates chosen by the user. The variance in the production is also estimated.

The production estimation methods used in this program are applicable to aquatic insects. However, the methods used to estimate fish populations are inappropriate for standard insect sampling techniques. The program could easily be modified to accept insect data and calculate production from it.

The program is written in FORTRAN 77 and can be executed on any machine with a FORTRAN 77 compiler. The version compiled for IBM-compatible personal computers can use up to 20 species, 40 sampling dates, and 3000 total fish in the input data. The program was developed for the Biological Monitoring and Abatement Program at the Oak Ridge National Laboratory, the Oak Ridge Gaseous Diffusion Plant, and the Oak Ridge Y-12 Plant but can be used with any appropriate multiple-pass stream fish data.

Example input files are in Appendixes A and B. An example interactive input session is in Appendix C, and example output is in Appendix D. Detailed information on the program structure is in Appendix E.

2. METHODS

The computer program for estimating fish production performs the following operations: (1) organization of the input data by site, sampling date, species, and sampling pass; (2) assignment of fish to length classes; (3) estimation of missing weight data; (4) estimation of population sizes for each length class, species, and date, using the multiple-pass data; and (5) estimation of annual secondary production for each species averaged between sampling dates chosen by the user.

2.1 SEPARATION BY DATE, SITE, SPECIES, AND PASS

The program can accept data from a number of sampling sites, which are specified in the input file with a seven-character alphanumeric name. For each site, the data will include values from two or more sampling dates. The input may also include values for a number of species. Multiple-pass data also include numbers for the electroshocking pass on which each fish was caught (see Sect. 2.4 for a description of the multiple-pass removal sampling method). The user interactively selects the site to be analyzed (data from only one site is processed in a run). The program separates the data by date, species, and pass, so the lines of input data are not required to be in any particular order.

2.2 ESTIMATION OF MISSING WEIGHT DATA

Often, many fish of the same species are caught at a site, and not all of the fish are weighed in the field. The program has a routine to estimate missing weights for fish whose lengths were measured. The routine uses least-squares regression to develop a logarithmic relation between the weight and length of fish of the same species, sampling date, and site as those with missing weights. This logarithmic length-weight regression relationship is then used to estimate the weight from the length of any fish with a missing weight. The program fills in missing weights whenever there are at least two fish from which to determine the regression equation (having both length and weight measured) and issues a warning statement giving the number of fish used to determine the regression equation. A maximum of 150 fish is used to develop the regression equation. The warning statement lets the user know when missing weights are estimated so the user can make sure that sufficient data were used to develop the regression equation.

2.3 LENGTH CLASS DETERMINATION

The size-frequency method of estimating production requires that the fish populations at each site be broken into length classes. These length classes need not have constant increments between them, and in fact the method works best when the length classes are defined by increments of equal growth over time instead of equal length (Garman and Waters, 1983). The program has an internal limit on the number of classes into which the population can be broken; this limit is 16 unless it is changed by the user (Sect. 3.1). The user must define the length classes for each species. This is done by defining several alternative sets of length classes and assigning one of these length class definitions to each species. The length classes are defined by simply listing the upper limit of each length class for the 16 classes (Sect. 3.2). The number of such length class definitions can vary; the program has an internal limit of five definitions, but this limit can be changed by the user (Sect. 3.1). If the length of a fish exceeds the maximum length class, the program issues a warning statement and assigns the fish to the maximum length class.

If there is only one fish in a length class (for one species and date), the program will move the fish into an adjacent length class that has one or more fish in it, if such an adjacent class exists. This reassignment to eliminate length classes with only one fish is made because the variance in the weight of fish in a length class cannot be estimated when there is only one fish in the class. [The variance in the weight is required to estimate the variance in the production (Sect. 2.5.4)].

Because the production estimation is calculated by using changes in population and biomass in length classes over time, the selection of length classes is important. (As an example of the influence of selection of length classes on results of fisheries studies, see Cofer and Malvestuto, 1983.) The same input data used with different length class definitions produce different production estimates.

2.4 POPULATION ESTIMATION

The population densities (in numbers per unit surface area or length of stream) of each length class (for each species and date) are estimated with a modification of the Zippin (1956) method, the maximum weighted likelihood method (Carle and Strub, 1978). The total population and population density for each species, each date, and each site are also estimated with the same method. An estimate of the variance is also made for each population estimate.

The population estimates are made from input data collected with the multiple-pass removal method (Seber and LeCren, 1967). This method involves blocking off a reach of stream and making two or more electroshocking passes through it. The lengths and weights of fish caught in each pass are recorded, and the fish are not returned to the reach being sampled until all passes are completed. Assumptions of population estimates using multiple-pass removal data are that the numbers of fish in the sampled reach are changed only by sampling capture and that the probability of capture is equal for all fish in all sampling passes (Carle and Strub, 1978); these assumptions were tested in the streams for which this program was developed by Gatz and Loar (1988).

The population estimate is made from the removal method catch vector. The catch vector contains the numbers of fish caught in each pass; for example, if, at a given sampling site and date, 10 fish of a species were caught on the first pass, 5 were caught on the second pass, and 3 were caught on the third pass, the catch vector for the species would be [10, 5, 3].

The program checks the data for all species caught at a site and date to determine how many electroshocking passes were made. If, for example, fish of species A were caught only in the first two passes and fish of species B were caught in a third pass, the program assigns a zero as the third value in the catch vector for species A. This zero is very important because it significantly changes the population estimate and variance for species A. If an electroshocking pass is made in which no fish of any species are caught, a line should be included in the fish data file (Sect. 3.3) with the date, site, and pass number but without a species code; the program will then assign a zero to the catch vector for this pass for all species.

The population estimate is made by determining the population size that has the highest probability of producing the catch vector. The maximum weighted likelihood method includes weighting factors that adjust the probability of capture to prevent infinite population estimates that otherwise occur for some catch vectors. The population estimate (N) is the smallest integer that is greater than or equal to T and satisfies the inequality (Carle and Strub, 1978):

$$\frac{N + 1}{N - T + 1} \prod_i \left[\frac{KN - X - T + \beta + (K - i)}{KN - X + \alpha + \beta + (K - i)} \right] \leq 1 \quad ,$$

where K is the number of passes made, $X = \Sigma_i (K - i) C_i$ where C_i is the number of fish caught on the *i*th pass, T is the total number of fish captured $[T = \Sigma_i (C_i)]$, and α and β are the probability weighting factors. The operator Π refers to the product of the terms inside the brackets for values of i from 1 to K. The program assumes that α and β are each equal to 1.

The program evaluates N by first assuming it is equal to T, checking the above inequality, and then incrementing the assumed value of N by 1 until it satisfies the inequality. The lowest value of N that satisfies the inequality is the population estimate for the catch vector.

The variance for the population estimate is determined by the method of Zippin (1956) as referred to by Carle and Strub (1978):

$$V(N) = \frac{N (1 - q^K) q^K}{(1 - q^K)^2 - (pK)^2 q^{K-1}} \quad ,$$

where p is the probability of a fish being caught and $q = (1 - p)$. The parameter p is estimated as

$$p = \frac{T}{KN - \sum_{i=1}^{K} x_i},$$

where x_i is the number of fish caught in passes prior to the ith pass.

Population densities are calculated by dividing the estimated population by the surface area or stream length sampled. The area or length sampled is put in interactively by the user for each date. The units of the population density and final production rates depend on the units of the surface area or length sampled.

2.5 PRODUCTION RATE ESTIMATION

2.5.1 Production estimation method

Once population estimates are made for each sampling date, species, and length class, the production between sampling dates for each species is calculated with a size-frequency method. This method was originally proposed for aquatic insects by Hynes (1961) and has been developed for stream fish by Garman and Waters (1983). The size-frequency method works by estimating the mean population size and mean weight of each size class for the time interval (chosen by the program user) over which production is estimated. (Production can be estimated over two or more sampling dates.) Production is then estimated as the sum over all length classes of the loss of biomass to the next larger length class. The equation for production (Garman and Waters, 1983) is

$$\begin{aligned} P_{s-f} = \quad & \{ \tfrac{1}{2} c [\bar{w}_1(\hat{\bar{N}}_1 - \hat{\bar{N}}_2) \\ & + \sum_{k=2}^{c-1} \bar{w}_k(\hat{\bar{N}}_{k-1} - \hat{\bar{N}}_{k+1}) \\ & + \bar{w}_c(\hat{\bar{N}}_{c-1} - \hat{\bar{N}}_c)]\} (1/CPI) \quad . \end{aligned}$$

In this equation P is the production between two dates, c is the number of length classes, k is an index from 1 to c for length classes, $\bar{w}_k$ is the mean weight of length class k during the time over which production is estimated, $\hat{N}_k$ is the population density in length class k during the time over which production is estimated, and CPI is the Cohort Production Interval. The values of $\hat{N}_k$ are from the population estimates for length classes (Sect. 2.4). The value of $\bar{w}_k$ is determined from the observed weights by assuming that the mean weight of all fish estimated to be in a length class is equal to the mean weight of the fish actually caught. The value of $\hat{N}$ and $\bar{w}$ are time-weighted averages calculated by using equations presented by Garman and Waters (1983). The units of P are grams of production per year per unit surface area or length of stream [the units of surface area or length are the same as those of the input values for this parameter (Sect. 2.4)].

No correction for Pe/Pa type errors (errors in estimating the proportion of a life cycle spent in a particular length class; Hamilton, 1969) is included in this method.

This program includes negative production values when they occur. Some investigators prefer to set negative production estimates for length classes equal to zero, which would require modification of this program.

2.5.2 Cohort Production Interval (CPI)

The production estimate is proportional to 1/CPI, so changes in the estimated CPI significantly change the production estimate. Originally included in the method as developed for aquatic insects, the CPI should be an estimate, in years (or fractions of years), of the "average maximum age attained (in years) by members of the population, assuming equilibrium between year-classes" (Garman and Waters, 1983). The CPI can also be considered an estimate of the number of year classes present. Dividing the total production by the CPI gives the annual production (i.e., the production per year class). The value of the CPI must be estimated by the user. Because the maximum age of fish could vary between dissimilar sites, values of the CPI should represent the site-specific population (not just the species) for which the production estimate is made. During sampling, it may be helpful to determine the ages of some of the large fish to estimate the CPI.

2.5.3 Date Combinations for Production Estimation

The annual production rate can be estimated between any two sampling dates (dates when population measurements were made). If production is to be estimated over a time period including more than two sample dates (i.e., data are available at dates between the beginning and end of the time period), production can be estimated by using the mean population sizes and weights (by length class) for the time period, including values measured at the intermediate dates. The program uses time-averaged populations and weights from whatever sample dates are selected.

Production calculations are commonly made to determine the annual average production rate for each year, and each calculation uses only the samples taken within a year (whether or not two or more samples were taken in the year). Such values can be used to compare production rates between years. Making production rate calculations from more than one year's data may be appropriate for comparing long-term production between two different sites.

The program is written so that the user interactively selects the combinations of dates for which each production calculation is made (Sect. 3.5). More than one such combination can be made in a single run, so the calculations needed to compare production between years at a site and between sites over several years can be made at one time. If data for a site is only available from one date, the program estimates populations but cannot estimate production.

2.5.4 Estimation of Variance in Production

The program estimates the variance in the production by using the equations presented by Garman and Waters (1983):

$$
\begin{aligned}
V(P) = \quad & \{(1/2c)^2[(\bar{w}_1 + \bar{w}_2)^2 V(\hat{N}_1) \\
& + V(\bar{w}_1)(\hat{N}_1 - \hat{N}_2)^2 \\
& + \sum_{k=2}^{c-1} (\bar{w}_{K-1} - \bar{w}_{K+1})^2 V(\hat{N}_K) \\
& + V(\bar{w}_K)(\hat{N}_{K-1} - \hat{N}_{K+1})^2 \\
& + (\bar{w}_{c-1} - \bar{w}_c)^2 V(\hat{N}_c) \\
& + V(\bar{w}_c)(\hat{N}_{c-1} - \hat{N}_c)^2]\}[(1/CPI)^2]
\end{aligned}
$$

where V(P) is the variance in the production, $V(\bar{w})$ is the variance in the mean weight of a length class, and $V(\hat{N})$ is the variance in the population density for a length class. To obtain $V(\hat{N})$, the variance in the population density (numbers per area or length of stream sampled), $V(\bar{N})$ is divided by the square of the area or length sampled (Sect. 2.4). $V(\bar{w})$ is estimated from the variance in the observed weights of fish in the length class; it is assumed that the variance in the weights of all fish estimated to be in a length class is the variance of the mean of the (sample) weights of fish actually caught (Newman and Martin, 1983). The equation for the variance in weight for a length class is

$$
V(w) = \frac{\sum_{i=1}^{m} (w_i - w)^2}{m(m - 1)}
$$

where m is the number of fish caught in the length class. When the population estimate for a length class is zero, V(w) is assumed to be zero. Weights of fish that have been estimated with length-weight regression (Sect. 2.2) are included in the calculation of V(w). Although this inclusion introduces another source of variance (error in the regression estimate of weights which is not accounted for in the variance estimate), it avoids problems that would occur if no fish in a length class had measured weights.

When the population estimate for a length class is 1 (in which case the value of V(w) cannot be determined from the above equation), the program attempts to move the single fish to the next smaller or larger length class so there are no such classes with estimated populations of 1. When length classes are so combined, the program prints a statement telling which classes have been combined. When a length class with one fish in it is surrounded by length classes with no fish, the program arbitrarily sets $V(\bar{w})$ to 10^{30}. The estimated V(P) is then very high and appears as "*****" in the output. (The program will not combine more than two length classes because the resulting variances would be high. In situations where there is a length class with one fish and no fish in the adjacent classes, it is best for the analyst to examine the data and make manual changes, such as deleting the fish from the input or using different length classes.)

The values of $V(\bar{N})$ and $V(\bar{w})$ are time-weighted averages of the variances estimed for each sampling date and are calculated with equations presented by Garman and Waters (1983).

3. USE OF THE PROGRAM

This section serves as a user's guide to the program. Changes that can be made in dimensions by editing and recompiling the program are described, as are the two required input files. Execution of the program and its interactive input are described. The program's error checking features are described. It is highly recommended that the user read Sect. 2 to understand what the program does before using it.

3.1 INTERNAL DIMENSION PARAMETERS

The program has internal array dimensions that limit the amount of data that it can handle. These limits should reflect the memory available on the computer executing the program as well as the amount of data to be processed. If the existing array dimensions are larger than can be accommodated by the computer, the dimensions can be reduced. If the size of the data sets requires it, the dimensions can be increased as long as memory is available. The array dimensions can easily be changed by editing the first programming statement and recompiling the FORTRAN code. The version compiled at ORNL for use on IBM AT-compatible desktop computers is dimensioned to allow up to 20 species, 40 sampling dates, 5 length class definitions (Sect. 3.2), 16 length classes per definition, 3000 fish in the input file for the site being processed (there is no limit on the total number of lines in the field data file), and 5 electroshocking passes. A maximum of 150 fish is used to calculate the regression parameters that are then used to estimate missing weights (Sect. 2.2), but this does not limit the number of fish that can have missing weights.

3.2 PARAMETER FILE

The program requires two input files for execution. The first is a parameter file that defines the length classes and CPI values used by the program for each species. This parameter file can be used with different data files (data files are described in Sect. 3.3). The parameter file must be named FISHPARM.DAT. An example parameter file is in Appendix A.

The program uses one set of length classes for each species (Sect. 2.3). In the parameter file, one or more sets of length classes are defined; these are referred to as length class definitions. The parameter file also specifies which length class definition should be used for each species. For example, the same length class definition could be used for all sunfish in a data file, and another length class definition, with bigger classes, could be used for all trout species.

The parameter file is unformatted (the data need not be in specific columns) for ease of input; however, this means that no values can be missing. If a value is missing, the program will read the following value and assign it to the wrong parameter, even if it has to read the next line of the file to do so.

The first line of the parameter file contains two numbers. The first number is the number of length classes in each length class definition. The default maximum value for this parameter is 16; the number used in this file can be less than 16. The second number in the first line of the parameter file is the number of length class definitions in the parameter file. The default maximum value for this parameter is 5. This number tells the program how many lines follow that each contain one set of length classes. Each fish species in the data set must use length classes from one of these defined sets. (See Sect. 3.1 if the default dimensions of 5 length class definitions and 16 length classes need to be increased.)

Following this first line are lines defining the length classes; the number of such lines is equal to the second number in the first line in the file. Length classes are defined by a list of numbers which are the upper limit of each class, in centimeters. In each of these lines, the number of values must equal the number of length classes specified in the first line of the parameter file, or else the program could use incorrect values for length classes. For example, if equal length class intervals of 1, 2, or 3 cm are used for all the fish species in a data file and the program uses 16 length classes, the first 4 lines of the parameter file would be

```
16 3
3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
```

Length classes need not be evenly spaced.

The remainder of the parameter file contains one line for each species (each species in the data file must be in the parameter file), with each line containing exactly three values. The first value in each line is the species code with a maximum of seven characters. The species code must be exactly as it appears in the data file, except that it is enclosed in single quotes (see the example in Appendix A). The characters in quotes must include any leading blanks that are in the species code field of the field data file.

The second value in each line is the Cohort Production Interval (CPI), which is an estimate of the average maximum life span of the species in years (Sect. 2.5). The CPI need not be an integer, so fractional values (e.g., 2.5) are allowed.

The final value in each line is the length class definition (the set of length classes, defined in the first lines of the parameter file) to be used for the species. This value should be 1 if the first length class definition (which is on line 2 of the parameter file and, in the above example, has increments of 3 cm) should be used, or 2 if the second length class definition (which is on line 3 of the parameter file and, in the above example, has increments of 2 cm) should be used, etc. The length classes must be selected so that, for each species, the largest fish in the data file is not longer than the largest length class.

It should be noted that the values in the parameter file apply to all the sites and dates that are in the field data file (Sect 3.3); the CPI values and length classes in the fish parameter file will be applied to all sites and dates processed. Therefore, different values of the CPI or different length classes for the same species at different sites or dates can only be used by making separate program runs with different values in the fish parameter file.

Fish species in the parameter file that are not in the field data file are ignored. Once a parameter file is created, it can be used repeatedly with different field data files. The program will not calculate populations or productions of species that are not in the parameter file.

3.3 FIELD DATA FILE

The field data file contains the fish sampling data (see Appendix B for an example data file). The program interactively prompts the user for the name of this file. There are no restrictions on the name of the file, but a file extension of ".DAT" is conventional. The field data file is formatted and contains one line per individual fish caught. The data file need not be sorted by site, date, species, or sampling pass. The format of the file follows.

Column(s)	Parameter
1-6	Date, in format MMDDYY (e.g., Aug. 18, 1987 is 081887)
8-14	Site, in up to seven alphanumeric characters
16	Pass number on which the fish was caught
18-23	Species code, in up to 6 alphanumeric characters
24-28	Length in centimeters
29-35	Weight, in grams.

If there are other data in the file to the right of column 35, they will be ignored.

It is extremely important that data be placed in the proper columns. Missing length or weight values can be handled by the program (Sect. 3.4); missing lengths or weights should be left blank or set to zero.

If an electroshocking pass was made and no fish were caught, a line must be included in the data file containing the date, site, and pass number but with blanks for the species code, length, and weight. This line will tell the program that a pass was made but no fish were caught, which will give different population and production estimates than if the program assumes that no such pass was attempted (Sect. 2.4).

3.4 PROGRAM EXECUTION AND INTERACTIVE INPUT

The program executes essentially in a batch mode, though the name of the data file and the dates for which to calculate production are entered interactively. The parameter file (FISHPARM.DAT) and the data file must be in the same directory from which the executable program file (FISHPROG.EXE) is executed. The program results are written to an output file called FISHPROG.OUT.

Upon execution, interactive input is required (see Appendix C for an example interactive session). First, the program prompts the user to enter the input data file name.

Next, the program lists the sites contained in the data file and requires the user to enter the name of the site to be used. The site name must be entered exactly as it appears when listed on the screen, including leading blanks and capital letters.

The program then lists the dates when the site was sampled. The user must enter one or more combinations of sampling dates between which the program will estimate production. The program calculates production by using the population sizes and weights for each of the dates selected. Each combination entered must include at least two dates but may include more than two dates. The user is requested to enter (1) the number of dates to be used for a production calculation and (2) the desired date indexes (1, 2, 3, etc.; according to how dates are listed on the screen). This request is repeated until all desired combinations are entered. If the return key ("Enter") is hit in response to the request to enter the number of dates to use, the program uses all dates by default.

As an example of selecting sampling dates for production calculations, if a site was sampled April 1, 1986 (date 1), October 10, 1986 (date 2), April 7, 1987 (date 3), and October 12, 1988 (date 4), the program will list these dates on the computer screen. If spring-to-fall production for each of the 2 years and the average production for the entire time are desired, the combinations 1 2; 3 4; and 1 2 3 4 should be entered.

The last interactive input is the length or surface area of stream sampled for each date. These values are used to determine population densities (Sect. 2.4). Any units for stream length or area can be used (e.g., feet, meters, square feet, hectares).

After the interactive input of length or area, the program completes its execution and writes a new output file FISHPROG.OUT. If a file with this name already exists in the directory, it will be overwritten by the new output file. Output files should be renamed or moved to a different directory before the program is executed again.

3.5 ERROR CHECKING

A number of internal error checks are in the program. These error checks do not test the quality or validity of the input data but only ensure that the kinds of data required to make the population and production estimates are provided.

The program issues error or warning statements if any of the following problems occur.

1. The length of a fish is zero or negative. If this occurs, a warning statement is issued and the observation is ignored for the production calculations. The observation is used in the estimate of the total population size for the site, date, and species.
2. Any of the limits set in the program (array dimensions; Sect. 3.1) for the number of sampling dates, number of total fish in the data file, number of species, or number of length class definitions are exceeded. If any of these limits are exceeded, an error statement is issued and the program terminates.
3. A species in the data file is not in the parameter file. If this occurs, a warning statement is issued and no population or production estimates are made for that species.
4. The weight of a fish is missing and is estimated by using length-weight regression. If less than two observations with both length and weight are available for a site, species, and date, the program issues a warning statement and does not estimate

production for the species and site. If two or more observations with both length and weight are available, the program issues a warning statement and proceeds to estimate the weights with the regression method. The warning statement says how many fish were used to determine the length-weight regression equation (Sect. 2.2).

5. A pass number is missing. When an observation in the data file is missing the pass number, the observation is ignored.
6. A fish has a length exceeding the largest length class. The program issues a warning statement and assigns the fish to the maximum length class.
7. The population estimation method does not find a solution within the maximum allowable number of iterations (this occurs rarely when more fish were caught in later passes than in the first one). The program issues an error statement and does not complete the population estimate. The program proceeds to the next species.
8. Data are available on only one date, so production cannot be estimated. A warning statement is issued and the program estimates populations only.

The error or warning statements issued by the program explain which problem or problems have occurred.

It should be noted that data lines with missing species codes are accepted by the program. Such lines are used to indicate a sampling pass in which no fish were caught (Sects. 2.4 and 3.3).

4. OUTPUT

The program writes output to a file called FISHPROG.OUT. Any existing files with this name in the same directory as the program will be overwritten when the program is executed. The program output presents the input data and computational results in several ways, which provides several opportunities for a careful user to detect erroneous input data or suspicious results. An example of the output file is in Appendix D.

Observations with missing length values are not used in the estimated populations by length class or in production estimates (because they cannot be assigned to a length class) but are included in the estimates of the total population of a species by date and site. Differences between the sum of estimated populations by length class and the estimated total population may be caused by the presence of observations with missing lengths; this can be checked because missing lengths will cause warning statements in the output.

The program output includes the following information. Warning and error statements may be interspersed among the other output values.

1. The name of the input data file.
2. The number of sampling dates at the chosen site, and a list of such dates.
3. Production output tables for each combination of species, and date combinations for which production estimates were made. The tables contain the site, species, dates over which production was estimated, estimated production, and variance and standard error (square root of the variance) of the production estimate. The production values are annual production rates, in units of grams per year per unit stream surface area or length (the surface area or length units are those of the area or length value entered interactively; Sect. 3.4).

4. A table of catch data by length class. For each species and date, the table lists the total number of fish caught and the numbers in each length class. The length classes are listed for each species.
5. A table of catch data by pass. For each species and date, the table lists the number of fish caught in each sampling pass.
6. A population estimate table. For each species and date, the table lists the estimated total population and the population in each length class.
7. A population density estimate table. For each species and date, the table lists the estimated total population density (fish per unit stream area or length) and the population density for each length class.
8. A table summarizing total population data (not broken down by length class) for each date and species. The values in the table are the total number of fish caught, the sum of the estimated populations by length class, the estimated total population, the variance in the estimated total population, the percentage of the estimated total population that was actually captured in sampling, and the biomass density in grams per unit stream area or length.
9. A weight data table. For each species and date, the mean weight of fish in each length class is listed.
10. A weight variance table. For each species and date, the variance of the mean weight for each length class is listed. The variance is listed as "******" when there is only one fish in the length class.

The user should review these tables carefully and should verify that

1. The correct input file was used.
2. The number of total observations and number of dates at each site are as expected, to ensure that no data were lost.
3. The production was estimated for the desired dates.

4. Length classes were assigned correctly. Species were not assigned to length classes that are too small, or to length classes that are too large and result in fish being assigned to too few classes.
5. There is not too high a difference between the estimated total population and the sum of the estimated populations by length class. Large discrepancies may result from poor data or from observations with no length values.
6. A relatively high percentage of the estimated population was actually caught (the percentage captured depends on species and sampling conditions, but values of 80% or higher are generally good). If this is not the case, the methods used are still valid, but results probably will have large variances. A low percentage of the estimated population actually being caught results when high numbers of fish are caught in later sampling passes.
7. The mean weights increase with increasing length classes. A mean weight of a length class that is higher than the mean weight of a higher length class indicates that field data are questionable.
8. No length classes contain only one fish. If there is a length class with one fish (and the program cannot correct this problem by moving the fish to an adjacent length class that has more than zero fish), the production estimate can still be made, but the variance for the production is invalid, having been set arbitrarily high (Sect. 2.5.4). This problem can be corrected manually either by removing the fish from the data file or by adjusting the length classes (Sect. 3.2).

5. PROGRAM VERIFICATION

This program implements existing statistical methods for analyzing fish population data. Users of the program are responsible for ensuring that these methods are appropriate for their data and that their data are accurate.

All routines in the program were checked extensively against hand calculations to verify programming correctness. The population estimation routine was further checked by reproducing the results published by Carle and Strub (1978) and by comparing results to another computer program written at ORNL that uses this method. The production estimation routine was checked by reproducing the results published by Garman and Waters (1983). ORNL does not, however, guarantee the accuracy of the program output. Users should review output carefully for inconsistencies or unexpected results.

6. REFERENCES

Carle, F. L., and M. R. Strub. 1978. A new method for estimating population size from removal data. Biometrics 34:621-630.

Cofer, L. M., and S. P. Malvestuto. 1983. Centimeter vs. inch-group measurements for describing age and size composition of fish populations. N. Am. J. Fish. Manage. 3:445-450.

Garman, G. C., and T. F. Waters. 1983. Use of the size-frequency (Hynes) method to estimate annual production of a stream fish population. Can. J. Fish. Aquat. Sci. 40:2030-2034.

Gatz, A. J., Jr., and J. M. Loar. 1988. Petersen and removal population size estimates: combining methods to adjust and interpret results when assumptions are violated. Environ. Biol. Fish. 21(4):293-307.

Hamilton, A. L. 1969. On estimating annual production. Limno. Oceanog. 14:771-782.

Hynes, H. B. N. 1961. The invertebrate fauna of a Welsh mountain stream. Arch. Hydrobiol. 57:344-388.

Newman, R. M., and F. B. Martin. 1983. Estimation of fish production rates and associated variances. Can. J. Fish. Aquat. Sci. 40:1729-1736.

Seber, G. A. F., and E. D. LeCren. 1967. Estimating population parameters from catches large relative to the population. J. Anim. Ecol. 36:631-643.

Zippin, C. 1956. An evaluation of the removal method of estimating animal populations. Biometrics 12:163-189.

APPENDIX A

EXAMPLE PARAMETER FILE FISHPARM.DAT

```
  16    5
1.  2.  3.  4.  5.  6.  7.  8.  9. 10. 11. 12. 13. 14. 15. 16.
2.  4.  6.  8. 10. 12. 14. 16. 18. 20. 22. 24. 26. 28. 30. 32.
3.  6.  9. 12. 15. 18. 21. 24. 27. 30. 33. 36. 39. 42. 45. 48.
4.  8. 12. 16. 20. 24. 28. 32. 36. 40. 44. 48. 52. 56. 60. 64.
5. 10. 15. 20. 25. 30. 35. 40. 45. 50. 55. 60. 65. 70. 75. 80.
 'BNDACE'  3   2
'MRDACE'  3   2
  'ROLLER'  3   2
 'SHINER'  3   2
 'SUCKER'  3   2
 'ROCKBA'  3   2
 'BLUGIL'  3   2
 'HOGSUC'  3   2
 'SCUPIN'  3   2
 'RBTRT '  5   5
 'BRNTRT'  5   4
```

Note: Values in line 1 mean that there are 16 length classes in each definition and that there are 5 such definitions.

Lines 2 through 6 define the length classes.

The remaining lines contain the Cohort Production Interval (CPI) and length class definitions for each species. For example, the last line means that species 'BRNTRT' has been assigned a CPI of 5 years and uses the fourth length class definition (length classes of <4, <8, <12, ... cm).

APPENDIX B

EXAMPLE DATA FILE FISH.DAT

101286	NFT7.7	3	RBTRT	7	6
101286	NFT7.7	1	RBTRT	11	18
101286	NFT7.7	1	RBTRT	8	
101286	NFT7.7	1	RBTRT	13	
101286	NFT7.7	3	RBTRT	8	6
071586	NFT7.7	1	BRNTRT	19	80
071586	NFT7.7	3	BRNTRT	20	91
101286	NFT7.7	2	BRNTRT	12	23
101286	NFT7.7	2	BRNTRT	16	42
101286	NFT7.7	1	RBTRT	8	5
101286	NFT7.7	2	BRNTRT	14	28
071586	NFT7.7	2	BRNTRT	29	232
101286	NFT7.7	3	BRNTRT	17	58
101286	NFT7.7	1	RBTRT	11	
071586	NFT7.7	3	RBTRT	11	12
101286	NFT7.7	2	RBTRT	8	6
101286	NFT7.7	1	BRNTRT	14	
101286	NFT7.7	2	RBTRT	9	8
101286	NFT7.7	1	BRNTRT	17	50
101286	NFT7.7	1	RBTRT	13	25
101286	NFT7.7	2	RBTRT	7	7
101286	NFT7.7	1	BRNTRT	29	274
101286	NFT7.7	3	BRNTRT	14	
101286	NFT7.7	1	RBTRT	8	12
101286	NFT7.7	1	RBTRT	6	3
101286	NFT7.7	1	BRNTRT	21	98
101286	NFT7.7	2	BRNTRT	13	25
101286	NFT7.7	2	RBTRT		
101286	NFT7.7	2	RBTRT	12	26
101286	NFT7.7	1	BRNTRT	25	154
101286	NFT7.7	3	BRNTRT	16	45
101286	NFT7.7	1	BRNTRT	29	208
101286	NFT7.7	1	RBTRT	8	5
101286	NFT7.7	1	RBTRT	7	7

101286	NFT7.7	3	RBTRT	9	6
101286	NFT7.7	2	RBTRT	9	10
071586	NFT7.7	1	BRNTRT	15	49
071586	NFT7.7	1	RBTRT	16	40
101286	NFT7.7	2	BRNTRT	13	24
101286	NFT7.7	3	RBTRT	7	6

Note: The first field is the date, in columns 1-6. The second field is the site code, in columns 8-14. The third field is the pass number, in column 16. The fourth field is the species code, in columns 18-23. The fifth field is the length, in columns 24-28. The sixth field is the weight, in columns 29-35.

APPENDIX C

EXAMPLE INTERACTIVE INPUT SESSION

```
> fishprog                                              Execute the program
 ENTER THE NAME OF THE FISH DATA FILE:
example                                          Use data file 'example.dat'
SITES IN INPUT FILE ARE:
1234567

NFT7.7                                          Sites in data file are listed
NFT7.8
ENTER NAME OF DESIRED SITE
IN EXACTLY THE SAME COLUMNS AS ABOVE
NFT7.7                                                  Use site 'NFT7.7'
  SITE NFT7.7  HAS THE FOLLOWING SAMPLING DATES:
  INDEX      DATE
     1       71586
 ENTER STREAM AREA OR LENGTH SAMPLED ON THIS DATE
1050                                                1050 square meters were
     2      101286                                  sampled on July 15, 1986
 ENTER STREAM AREA OR LENGTH SAMPLED ON THIS DATE
1035                                                1035 square meters were
  SITE NFT7.7  HAS THE FOLLOWING SAMPLING DATES:                 sampled on
  INDEX      DATE                                          October 12, 1986
     1       71586
     2      101286                                   All sampling dates for
                                                        the site are listed
  DO YOU WISH TO SELECT ANOTHER DATE COMBINATION? (Y/N)
y                                                       Select dates to use
  HOW MANY DATES FOR THIS PRODUCTION CALCULATION?for production estimate
  OR HIT RETURN TO SELECT ALL DATES
2                                                              Use 2 dates
  ENTER THE           2 DATE INDICES SEPARATED BY A SPACE
1 2                                                  Select the 1986 data
  DO YOU WISH TO SELECT ANOTHER DATE COMBINATION? (Y/N)
n                                                    Select no more dates
  CALCULATIONS COMPLETED
```

FORTRAN STOP **Program has finished;**

> **output written to file 'FISHPROG.OUT'**

Note:

The user enters only the underlined text. Bold text on the right side of the page is explanatory and will not appear on the computer screen.

APPENDIX D

EXAMPLE OUTPUT

FISH DATA FILE example OPENED

**** WARNING: MISSING LENGTH

SPECIES : RBTRT SITE : NFT7.7 DATE : 101286 PASS : 2 LENGTH : 0.0 WEIGHT : 0.0

2 SAMPLING DATES AT NFT7.7

WARNING : FOR RBTRT 71586, FISH MOVED FROM LENGTH CLASS 5 TO 4

MISSING WEIGHT FOR: 101286 NFT7.7 RBTRT

WARNING : 103 POINTS USED IN LENGTH-WEIGHT REGRESSION

SOLUTION: WT. = 0.0360 LENGTH** 2.58

LENGTH: 8.000000 WEIGHT: 7.635090

LENGTH: 13.00000 WEIGHT: 26.67423

LENGTH: 11.00000 WEIGHT: 17.34444

LENGTH: 7.000000 WEIGHT: 5.412450

SITE NFT7.7 HAS THE FOLLOWING SAMPLING DATES:

INDEX	DATE	AREA/LENGTH SAMPLED
1	71586	1050.0
2	101286	1035.0

WARNING : FOR BRNTRT 71586, FISH MOVED FROM LENGTH CLASS 8 TO 7

WARNING : FOR BRNTRT 71586, FISH MOVED FROM LENGTH CLASS 6 TO 5

MISSING WEIGHT FOR: 101286 NFT7.7 BRNTRT

WARNING : 83 POINTS USED IN LENGTH-WEIGHT REGRESSION

SOLUTION: WT. = 0.0529 LENGTH** 2.47

LENGTH: 14.00000 WEIGHT: 35.82261

LENGTH: 14.00000 WEIGHT: 35.82261

WARNING : FOR BRNTRT 101286, FISH MOVED FROM LENGTH CLASS 10 TO 9

PRODUCTION RATE FOR DATES: 71586 101286

	SPECIES	SITE	GRAMS/YR/AREA	VAR	ST ERR
1	RBTRT	NFT7.7	1.273******************************		
2	BRNTRT	NFT7.7	2.784******************************		

ACTUAL CATCH BY CLASS

	SPECIES	SITE	DATE	TOTAL	CLASSES											
					< 5.	<10.	<15.	<20.								
1	RBTRT	NFT7.7	71586	10	0	2	4	4								
2	RBTRT	NFT7.7	101286	107	0	71	26	7	2	0	1					
					< 4.	< 8.	<12.	<16.	<20.	<24.	<28.	<32.	<36.	<40.	<44.	<48.
3	BRNTRT	NFT7.7	71586	15	0	0	2	5	5	0	2	0	0	0	0	1
4	BRNTRT	NFT7.7	101286	85	0	2	4	44	13	16	2	2	2			

TOTAL SAMPLES IS : 217

ACTUAL CATCH BY PASS

	SPECIES	SITE	DATE	TOTAL	PASSES		
					1	2	3
1	RBTRT	NFT7.7	71586	10	4	3	3
2	RBTRT	NFT7.7	101286	107	67	24	16
					1	2	3
3	BRNTRT	NFT7.7	71586	15	10	4	1
4	BRNTRT	NFT7.7	101286	85	51	25	9

POPULATION ESTIMATES BY CLASS

	SPECIES	SITE	DATE	TOTAL	CLASSES											
					< 5.	<10.	<15.	<20.								
1	RBTRT	NFT7.7	71586	10	0	2	4	4								
2	RBTRT	NFT7.7	101286	115	0	79	26	7	2	0	1					
					< 4.	< 8.	<12.	<16.	<20.	<24.	<28.	<32.	<36.	<40.	<44.	<48.
3	BRNTRT	NFT7.7	71586	15	0	0	2	5	5	0	2	0	0	0	0	1
4	BRNTRT	NFT7.7	101286	92	0	2	4	47	17	16	2	2	2			

ESTIMATED TOTAL POPULATION IS : 232

POPULATION DENSITIES BY CLASS

	SPECIES	SITE	DATE	TOTAL	CLASSES						
					< 5.	<10.	<15.	<20.			
1	RBTRT	NFT7.7	71586	0.010	0.000	0.002	0.004	0.004			
2	RBTRT	NFT7.7	101286	0.111	0.000	0.076	0.025	0.007	0.002	0.000	0.001

					< 4.	< 8.	<12.	<16.	<20.	<24.	<28.	<32.	<36.	<40.	<44.	<48.
3	BRNTRT	NFT7.7	71586	0.014	0.000	0.002	0.002	0.005	0.005	0.000	0.002	0.000	0.000	0.000	0.000	0.001
4	BRNTRT	NFT7.7	101286	0.089	0.000	0.002	0.004	0.045	0.016	0.015	0.002	0.002	0.002			

TOTAL POPULATION ESTIMATES

	SPECIES	SITE	DATE	ACTUAL CATCH	SUM OF ALL CLASS	EST. TOT. POP	VAR. IN TOTAL	% CAPTURED	BIOMASS DENSITY, GM/AREA
1	RBTRT	NFT7.7	71586	10	10	12	21.11	83.3	0.4
2	RBTRT	NFT7.7	101286	107	115	116	29.99	92.2	2.2
3	BRNTRT	NFT7.7	71586	15	15	15	0.59	100.0	2.1
4	BRNTRT	NFT7.7	101286	85	92	91	18.82	93.4	6.6

AVERAGE WEIGHT BY CLASS, GRAMS

	SPECIES	SITE	DATE	CLASSES											
				< 5.	<10.	<15.	<20.								
1	RBTRT	NFT7.7	71586	0.0	11.5	22.3	70.8								
2	RBTRT	NFT7.7	101286	0.0	8.1	30.8	50.9	112.0	0.0	222.0					
				< 4.	< 8.	<12.	<16.	<20.	<24.	<28.	<32.	<36.	<40.	<44.	<48.
3	BRNTRT	NFT7.7	71586	0.0	0.0	18.5	30.6	85.6	0.0	196.0	0.0	0.0	0.0	0.0	1200.0
4	BRNTRT	NFT7.7	101286	0.0	38.5	22.0	38.2	67.7	110.8	155.0	241.0	586.0			

WEIGHT VARIANCE BY CLASS, GRAMS

	SPECIES	SITE	DATE	CLASSES											
				< 5.	<10.	<15.	<20.								
1	RBTRT	NFT7.7	71586	0.0	2.3	15.1	298.2								
2	RBTRT	NFT7.7	101286	0.0	0.9	3.1	10.8	9.0	0.0*******						
				< 4.	< 8.	<12.	<16.	<20.	<24.	<28.	<32.	<36.	<40.	<44.	<48.
3	BRNTRT	NFT7.7	71586	0.0	0.0	6.3	23.3	105.5	0.0	1296.0	0.0	0.0	0.0	0.0*******	
4	BRNTRT	NFT7.7	101286	0.0	870.3	3.0	1.3	27.3	10.6	1.0	1089.0	24336.0			

Note: The variance cannot be calculated for either species at this site because some length classes contain only one fish. This is indicated by the asterisks in the variance column in the production rate table, the asterisks in the table of weight variance by class, and by the length classes with one fish in the table of population estimates by class. The inability to calculate the variance is caused by the capture of single very large fish of both species on the second sampling date. This common problem can be circumvented by deleting these individuals from the data file; this was done, and the results are in the following output file.

Examination of the tables of average weight by class and weight variance by class indicates a problem with species 'BRNTRT' (brown trout) at station NFT7.7. The average weight in the <8-cm length class is greater than the average weight in the <12- and <16-cm classes, and the variance in the average weight for the <8-cm class is very high. The data for this species and site should be examined for errors. (Such an examination shows that there are two brown trout with length of 8 cm and weights of 9 and 68 g; the fish with a weight of 68 g has an incorrect weight or length or a very unusual diet.)

```
FISH DATA FILE example2    OPENED

 **** WARNING: MISSING LENGTH
SPECIES : RBTRT    SITE : NFT7.7    DATE :  101286    PASS :  2    LENGTH :        0.0    WEIGHT :        0.0

   2 SAMPLING DATES AT NFT7.7

WARNING : FOR  RBTRT    71586, FISH MOVED FROM LENGTH CLASS   5 TO  4
MISSING WEIGHT FOR:   101286   NFT7.7   RBTRT
WARNING :          102 POINTS USED IN LENGTH-WEIGHT REGRESSION
  SOLUTION: WT. =  0.0349    LENGTH**   2.59
LENGTH:    8.000000     WEIGHT:    7.625941
LENGTH:    13.00000     WEIGHT:    26.81978
LENGTH:    11.00000     WEIGHT:    17.39928
LENGTH:    7.000000     WEIGHT:    5.396100

 SITE NFT7.7  HAS THE FOLLOWING SAMPLING DATES:
 INDEX     DATE       AREA/LENGTH SAMPLED
    1      71586        1050.0
    2     101286        1035.0
WARNING : FOR  BRNTRT   71586, FISH MOVED FROM LENGTH CLASS   8 TO  7
WARNING : FOR  BRNTRT   71586, FISH MOVED FROM LENGTH CLASS   6 TO  5
MISSING WEIGHT FOR:   101286   NFT7.7   BRNTRT
WARNING :            83 POINTS USED IN LENGTH-WEIGHT REGRESSION
  SOLUTION: WT. =  0.0529    LENGTH**   2.47
```

LENGTH: 14.00000 WEIGHT: 35.82261

LENGTH: 14.00000 WEIGHT: 35.82261

WARNING : FOR BRNTRT 101286, FISH MOVED FROM LENGTH CLASS 10 TO 9

PRODUCTION RATE FOR DATES: 71586 101286

	SPECIES	SITE	GRAMS/YR/AREA	VAR	ST ERR
1	RBTRT	NFT7.7	0.909	0.007	0.083
2	BRNTRT	NFT7.7	2.090	0.035	0.186

ACTUAL CATCH BY CLASS

	SPECIES	SITE	DATE	TOTAL	CLASSES								
					< 5.	<10.	<15.	<20.					
1	RBTRT	NFT7.7	71586	10	0	2	4	4					
2	RBTRT	NFT7.7	101286	106	0	71	26	7	2				
					< 4.	< 8.	<12.	<16.	<20.	<24.	<28.		
3	BRNTRT	NFT7.7	71586	14	0	0	2	5	5	0	2		
4	BRNTRT	NFT7.7	101286	85	0	2	4	44	13	16	2	2	2

TOTAL SAMPLES IS : 215

ACTUAL CATCH BY PASS

	SPECIES	SITE	DATE	TOTAL	PASSES		
					1	2	3
1	RBTRT	NFT7.7	71586	10	4	3	3
2	RBTRT	NFT7.7	101286	106	66	24	16
					1	2	3
3	BRNTRT	NFT7.7	71586	14	10	3	1
4	BRNTRT	NFT7.7	101286	85	51	25	9

POPULATION ESTIMATES BY CLASS

	SPECIES	SITE	DATE	TOTAL	CLASSES				
					< 5.	<10.	<15.	<20.	
1	RBTRT	NFT7.7	71586	10	0	2	4	4	
2	RBTRT	NFT7.7	101286	114	0	79	26	7	2

					< 4.	< 8.	<12.	<16.	<20.	<24.	<28.		
3	BRNTRT	NFT7.7	71586	14	0	0	2	5	5	0	2		
4	BRNTRT	NFT7.7	101286	92	0	2	4	47	17	16	2	2	2

ESTIMATED TOTAL POPULATION IS : 230

POPULATION DENSITIES BY CLASS

	SPECIES	SITE	DATE	TOTAL	CLASSES								
					< 5.	<10.	<15.	<20.					
1	RBTRT	NFT7.7	71586	0.010	0.000	0.002	0.004	0.004					
2	RBTRT	NFT7.7	101286	0.110	0.000	0.076	0.025	0.007	0.002				
					< 4.	< 8.	<12.	<16.	<20.	<24.	<28.		
3	BRNTRT	NFT7.7	71586	0.013	0.000	0.002	0.002	0.005	0.005	0.000	0.002		
4	BRNTRT	NFT7.7	101286	0.089	0.000	0.002	0.004	0.045	0.016	0.015	0.002	0.002	0.002

TOTAL POPULATION ESTIMATES

	SPECIES	SITE	DATE	ACTUAL CATCH	SUM OF ALL CLASS	EST. TOT. POP	VAR. IN TOTAL	% CAPTURED	BIOMASS DENSITY, GM/AREA
1	RBTRT	NFT7.7	71586	10	10	12	21.11	83.3	0.4
2	RBTRT	NFT7.7	101286	106	114	116	34.13	91.4	2.0
3	BRNTRT	NFT7.7	71586	14	14	14	0.40	100.0	1.0
4	BRNTRT	NFT7.7	101286	85	92	91	18.82	93.4	6.6

AVERAGE WEIGHT BY CLASS, GRAMS

	SPECIES	SITE	DATE	CLASSES								
				< 5.	<10.	<15.	<20.					
1	RBTRT	NFT7.7	71586	0.0	11.5	22.3	70.8					
2	RBTRT	NFT7.7	101286	0.0	8.1	30.8	50.9	112.0				
				< 4.	< 8.	<12.	<16.	<20.	<24.	<28.		
3	BRNTRT	NFT7.7	71586	0.0	0.0	18.5	30.6	85.6	0.0	196.0		
4	BRNTRT	NFT7.7	101286	0.0	38.5	22.0	38.2	67.7	110.8	155.0	241.0	586.0

WEIGHT VARIANCE BY CLASS, GRAMS

	SPECIES	SITE	DATE	CLASSES							
				< 5.	<10.	<15.	<20.				
1	RBTRT	NFT7.7	71586	0.0	2.3	15.1	298.2				
2	RBTRT	NFT7.7	101286	0.0	0.9	3.1	10.8	9.0			
				< 4.	< 8.	<12.	<16.	<20.	<24.	<28.	
3	BRNTRT	NFT7.7	71586	0.0	0.0	6.3	23.3	105.5	0.0	1296.0	
4	BRNTRT	NFT7.7	101286	0.0	870.3	3.0	1.3	27.3	10.6	1.0	1089.024336.0

Note: This output is for the same data as the previous output file, except that two large individuals (one rainbow trout and one brown trout) were deleted from the input. Comparison to the previous output shows that the production rate estimates are reduced, but the variance can be calculated and is relatively low.

APPENDIX E

PROGRAMMER INFORMATION

PROGRAM DESCRIPTION

This program performs the following functions.

1. Reads parameter file until the end is encountered or the array limit is reached.
2. Reads field data file until the end is encountered or the array limit is reached. Keeps only the records for the selected site. A check for valid data is performed. Warnings are issued for records with missing length or pass data. Each record is stored according to capture date and species with dates sorted in ascending order.
3. Assigns the proper CPI value to each species.
4. Sorts all field data by capture date and enumerates the number of species and capture dates found in the data set.

Steps 5 through 10 are performed for each species and date.

5. Forms a vector L of length LV of indices pertaining to this case.
6. When data are encountered with missing weights, uses linear regression to determine the missing values if there are at least two weight values present.
7. Sorts data for this case by length and assigns the fish to proper length classes. If length exceeds the largest class provided, a warning is issued, and the fish is included in the largest class.
8. Calculates a population estimate, its variance, and standard error for each length class.
9. Calculates the average weight and its variance for each length class.
10. Calculates one population estimate over all class sizes.
11. Gets the user to interactively specify sets of dates for use in calculating production.
12. For each species, calculates a value for annual production. These calculations are done for each set of dates supplied by the user.

13. For each set of capture dates, prints a table of production values by species.
14. Prints sorted field data by class.
15. Prints population estimates by class.
16. Prints population estimates summary table.
17. Prints average weights and variance by length class.
18. Signals end of program calculations.

PROGRAM INPUT/OUTPUT FILE DESCRIPTIONS

The program uses the following files, three of which are described below. The output file is described in Sect. 4 of the text.

OUTFIL	UNIT=6	STATUS='NEW'	Output file
INFIL2	UNIT=8	STATUS='OLD'	Parameter data file
INFILE	UNIT=10	STATUS='OLD'	Field data file
TMPFIL	UNIT=12	STATUS='UNKNOWN'	Intermediate output

FILE 8 - PARAMETER DATA FILE

Internally specified, the program will expect to find the fish parameters on a file called "fishparm.dat." All data on this file are input with a list-directed read (i.e., the data may appear in any column and each field is separated by a comma or at least one blank). Literals are enclosed in single quotes.

Three items appear in this file:

1. A single line containing these values:
   II - Number of increments in each length class definition.
   IJ - Number of length class definitions to be read.

2. A set of lines, each containing these values:
   (CIBS(I,J),I=1,II) - Upper limit for each length class (centimeters), in ascending order, for definition J.
   IJ definitions will be read.
3. A set of lines, each containing these values (with one such line for each fish species):

   SNAME(J) - Species name as it will appear on the field data file. This name consists of six alphameric characters.

   SCPI(J) - This variable is used in annual production calculations and is most easily represented by the average maximum age (in years) attained by members of this species.

   LCI(J) - Length class definition for this species. This is the index referring to one of the entries in item 2. The LCI should be selected so that the maximum specimen length in the field data falls within the definition limits.

The third item in the file will continue to be read until an end-of-file is encountered or the arrays are filled. Current maximum capacity is 20.

FILE 10 - FIELD DATA FILES

Externally specified, the program will prompt the user for the data file name. This file contains only one record type with one record per observation. The file is read until the end is encountered or the array is filled. Current maximum capacity is 3000 observations for the selected site; there is no limit on the total length of this file.

Each observation contains the following data:

Variable	Description	Type	Columns
MO_DAY_YR	Date of capture	Integer	1-6
LOC	Location of capture	Alphameric	8-14
PASS	Captured on the nth	Integer	16
SPECIES	Name of fish	Alphameric	18-23
LENGTH	Length of fish (centimeters)	Real	24-28
WEIGHT	Weight of fish (grams)	Real	29-35

FILE 12 - INTERMEDIATE OUTPUT

Internally specified, the program uses this file to store intermediate data for final report writing. This file can be used to save program output.

Variables:

ACPI	Value of CPI for the ith species
ALPHA	Coefficient for population estimates
AP	Annual production
AW	Average weight of length group
BETA	Coefficient for population estimates
CIBS	Upper limit for each length class
CMAX	Observation with greatest length for this data subset
DLIST	List of sampling dates represented
IADAY	Julian date - Days from 1/1/70 to date of sample
IDATE	Date of sample expressed as MODAYR; MO = month, DA = day, YR = year
ILOC	Index of site for sample i from LLIST
IOD	Auxiliary printing occurs when IOD is nonzero
IST	Index of species name for sample from SLIST
ITME	Index of date for a sample
KC	Maximum value for length class for this case
LCC	Total fish captured for length class i, date j
LCI	Index referring to a set of length class definitions

LCLASS	Length class assigned to an observation
LLIST	List of capture sites
LV	Number of observations for this case (location, date, and species)
MAXLC	Maximum number of length classes
NBAR	Estimate of the annual mean number of individuals within the *i*th length group
NC	Number of fish caught in pass i for this case (date, species, and site)
NDS	Number of dates sampling occurred at *i*th site
NDT	Number of sample dates
NK	Largest pass number for this case
NLC	Number of unique sites represented
NN	Population estimate, performed for each length class
NPC	Tally array, used to check for gaps in pass data
NPE	Population estimate for length class i, date j
NPETOT	Total population for this case over all class lengths
NSAMP	Number of samples
NST	Number of unique species types represented
NTC	Sum of all NC
NTX	The value of X used in paper to calculate populations
SCPI	Average maximum age attained by a species
SLIST	List of species names
SNAME	Species name (six alphameric characters)
WBAR	Estimate of the annual mean weight of individuals within the *i*th length group

PROGRAM SUBROUTINES

PROGRAM ALSQ

Purpose: To solve the linear least squares problem, norm(AB - Y) = minimum.

Variables:

A	An array containing the least squares matrix. Upon return the (m + 1)th column contains the approximating vector ab.
Y	The vector to be fit.
B	Contains upon return the coefficients of the fit.
R2	Contains upon return the residual sum of squares.
N	Number of rows in the least squares matrix.
M	Number of columns in the least squares matrix.
NA	First dimension of the array A.

PROGRAM ANDAY

Purpose: To convert month/day/year to a Julian date. Calculates days since 1/1/70.

Variables:

MONTH	Integer value for month.
IDAY	Integer value for day.
IYR	Integer value for year (2 digits).
IADAY	Number of days since 1/1/70.

PROGRAM ANPROD

Purpose: To calculate the annual production (PSF) by the size-frequency method using the annual mean abundance and weight values derived from each length group.

Variables:

PSF	Annual production (kilograms).
VPSF	Variance of the production estimate.
NPE	Number in the length class (class,date).
AW	Average weight of the length class (class,date).
DATE	Vector of sampling dates.
APBC	Annual production by size class.
ND	Total number of dates.
K	Number of length classes.

CPI	Cohort production interval (Sect. 2.5.2).
MAXD	Maximum number of sampling dates.
MAXLC	Maximum number of length classes.
NDD	Number of dates to be used in the calculation.
DINDX	Indices indicating which dates in the DATE vector will be used.
VNF	Variance of the population estimate by class (class, date).
VWT	Variance of the weight estimate by class (class, date).
IOD	Output of intermediate calculations, performed when IOD is nonzero.

PROGRAM EVAL

Purpose: Used in population estimation to evaluate the maximum weighted likelihood estimate for a given value of N.

Variables:

N	An integer greater than or equal to T.
C	Vector whose elements represent the number of individuals captured in the K removal periods.
T	Total number of individuals captured.
K	Total number of removal periods.
X	Summation over i of (K - i)Ci.
IN	A three-term sum, IN = -X' - T + BETA, where BETA = 1.
ID	A three-term sum, ID = -X + ALPHA + BETA, where ALPHA = BETA = 1.
P	The value for the maximum likelihood function.
EVAL	1 - P.

PROGRAM GETD

Purpose: To get the user to interactively specify the sampling dates to be used for the production calculations. Several production calculations may be specified up to a maximum of MAXPC. If no dates are specified, the production is calculated over all sampling dates for the site.

Variables:

ND	Number of production calculations specified.
JDI	Array of date indices input by user.
MAXDPS	Maximum number of dates per site.
MAXPC	Maximum number of production calculations.
IPV	Length of date vectors input by user.
NDD	Number of sampling dates at this site.

PROGRAM ISORT

Purpose: To rearrange an integer vector in ascending order. An associated integer vector is also rearranged according to this order.

Variables:

A	Integer vector to be sorted.
CA	Associated integer vector to be arranged according to ascending values of vector A.
NROWS	Length of vectors A and CA.
IDR	Dimension of vectors A and CA.

PROGRAM SANDL

Purpose: To calculate population estimates from a two-pass sample. The variance and 95% confidence limits are also produced. Note: While this routine remains in the code, it is not used for any of the population, production, or variance estimates.

Variables:

C	Vector with counts for passes one and two.
NN	The Seber and Le Chren population estimate.
LL	The 95% confidence interval's lower limit.
LU	The 95% confidence interval's upper limit.
PEST	The estimate of the proportion of fish captured.
VAR	The estimate variance.

PROGRAM SORT

Purpose: To create an integer vector K of indices such that A[K(i)] is the _i_th smallest element of the A vector. The physical ordering of A remains unchanged.

Variables:

A	Real vector to be sorted.
K	Integer vector of indices to be arranged according to ascending values of vector A.
L	L, a vector of indices, which is a subset of a larger vector. Standard order is accomplished with L(I) = I.
N	Dimension of vectors A and K.

INTERNAL DISTRIBUTION

1. S. M. Bartell
2. G. F. Cada
3. S. W. Christensen
4. D. L. DeAngelis
5. J. W. Elwood
6. W. M. Harris
7. J. E. Hawthorne
8. S. G. Hildebrand
9-13. B. D. Holcomb
14. M. A. Huston
15. J. M. Loar
16. D. H. Pike
17-21. S. F. Railsback
22. D. E. Reichle
23. P. S. Rohwer
24-28. M. G. Ryon
29. M. J. Sale
30. E. M. Schilling
31. J. G. Smith
32. G. W. Suter
33. Central Research Library
34-48. ESD Library
49-50. Laboratory Records Department
51. Laboratory Records, RC
52. ORNL Patent Office
53. ORNL Y-12 Technical Library

EXTERNAL DISTRIBUTION

54. Steven Ahlstedt, Division of Water Resources, Tennessee Valley Authority, Norris, TN 37828
55. R. P. Berube, Deputy Assistant Secretary for Environment, EH-20, U.S. Department of Energy, Washington, DC 20585
56. Prof. Wesley J. Birge, School of Biological Sciences, University of Kentucky, 101 Morgan Building, Lexington, KY 40506
57. C. M. Borgstrom, Director, Office of NEPA Project Assistance, EH-25, U.S. Department of Energy, Washington, DC 20585
58. J. Thomas Callahan, Associate Program Director, Ecosystem Studies Program, Room 336, 1800 G Street, NW, National Science Foundation, Washington, DC 20550
59. R. R. Colwell, Director, Maryland Biotechnology Institute, Microbiology Building, University of Maryland, College Park, MD 20742
60. W. E. Cooper, Department of Zoology, College of Natural Sciences, Michigan State University, East Lansing, MI 48824
61. N. H. Cutshall, 10461 White Granite Dr., Suite 204, Oakton, VA 22124
62. G. J. Foley, Director, Environmental Monitoring Systems Laboratory, MD-75, Research Triangle Park, NC 27711
63. A. John Gatz, Jr., Department of Zoology, Ohio Wesleyan University, Delaware, OH 43015
64. Clayton S. Gist, Environmental Protection Branch, U.S. Department of Energy, Oak Ridge Operations Office, Oak Ridge, TN 37831
65. H. Wayne Hibbitts, Chief, Environmental Protection Branch, U.S. Department of Energy, Oak Ridge Operations Office, Oak Ridge, TN 37831
66. David R. Hopkins, U.S. Environmental Protection Agency, 345 Courtland Street, N.E., Atlanta, GA 30365

67. J. W. Huckabee, Program Manager, Ecological Studies Group, Electric Power Research Institute, 3412 Hillview Avenue, P.O. Box 10412, Palo Alto, CA 94303
68. George Y. Jordy, Director, Office of Program Analysis, Office of Energy Research, ER-30, G-226, U.S. Department of Energy, Washington, DC 20545
69. Timothy Joseph, Environmental Remediation Branch, G-108, Department of Energy, Oak Ridge Operations Office, Oak Ridge, TN 37831
70. Earl C. Leming, Manager, Knoxville Basin Office, Tennessee Department of Health and Environment, 1605 Prosser Road, Knoxville, TN 37914-3434
71. G. E. Likens, Director, The New York Botanical Garden, Institute of Ecosystem Studies, The Mary Flagler Cary Arboretum, Box AB, Millbrook, NY 12545
72. Gabriel J. Marciante, Environmental Protection Branch, U.S. Department of Energy, Oak Ridge Operations Office, Oak Ridge, TN 37831
73. Helen McCammon, Director, Ecological Research Division, Office of Health and Environmental Research, Office of Energy Research, ER-75, U.S. Department of Energy, Washington, DC 20545
74. A. David McKinney, Chief, Water Quality Section, Tennessee Wildlife Resources Agency, P.O. Box 40747, Ellington Agricultural Center, Nashville, TN 37204
75. Wayne Schacher, Aquatic Habitat Protection Biologist, Tennessee Wildlife Resources Agency, 6032 W. Andrew Johnson Highway, Talbot, TN 37877
76. Frank J. Wobber, Ecological Research Division, Office of Health and Environmental Research, Office of Energy Research, ER-75, U.S. Department of Energy, Washington, DC 20545
77. Office of Assistant Manager for Energy Research and Development, Oak Ridge Operations, P.O. Box 2001, U.S. Department of Energy, Oak Ridge, TN 37831-8600

78-87. Office of Scientific and Technical Information, P.O. Box 62, Oak Ridge, TN 37831