

## **A SAS PROGRAM SUMMARIZING DATA COLLECTED BY ROTENONE SAMPLING**

DOUGLAS F. DURANT, School of Forest Resources and Conservation, University of Florida, Gainesville, FL 32611

DOUGLAS E. COLLE, School of Forest Resources and Conservation, University of Florida, Gainesville, FL 32611

JEROME V. SHIREMAN, School of Forest Resources and Conservation, University of Florida, Gainesville, FL 32611

*Abstract:* A SAS program was designed to compile and summarize data collected by rotenone sampling. The program is designed to accept data in a variety of forms, providing a statistical summary of number, weight and percentage composition for all species and also a separate summary for harvestable gamefish and/or forage fish. Options are available to calculate weights for fish collected on second and third day pickups with weight-length relationships established from first day fish for either individual fish or by size groups.

---

Proc. Ann. Conf. S.E. Assoc. Fish & Wildl. Agencies 33: 408-414

According to Hall (1962) rotenone sampling is one of the most preferred methods for obtaining fish population data in all the southeastern states and usually incorporates either cove or blocknet sampling. In many situations rotenone samples provide reliable length frequency data and can be used for accurate standing crop estimates for selected sections of the body of water sampled (Provine 1975).

The quantity of data generated by rotenone sampling requires numerous hours for manual computation and summarization of results for management decisions. Opportunities for computational errors are many when this amount of data is analyzed. This paper describes a SAS program that summarizes, simplifies, and increases the accuracy of data assemblage. The program is designed to accept data in a variety of forms and provide statistical summaries of the number, weight, percentage composition for all species, and summaries for harvestable gamefish and/or forage fish. Weight-length options are available within the program that can be used to calculate weights of second and third day fish. These data can be generated for either individual fish or by size groups. If this option is not desired, data summaries can be obtained when second and third day weights are recorded in the field.

The computer language, SAS, was chosen because of versatility of data manipulation and the availability of wide ranges of easily employed statistical analysis. An inexperienced programmer can quickly utilize SAS programs because the language does not require the user to format output or program summary techniques. Barr et al. (1976) provided a list of institutions which maintain SAS programs and if the user is unfamiliar with SAS, manuals by Helwig (1977, 1978) are suggested as useful introductions. This study was sponsored in part by an Environmental Protection Agency grant R80549 and special funds from the Center for Environmental Programs of the Institute of Food and Agricultural Sciences, University of Florida.

### **MATERIALS AND METHODS**

The following field data are needed: sample location (LAKE), date of rotenone application (DATE), sample number (SAMPLE), day of fish pickup (DAY), species collected (SPECIE), size group (SG), number in the size group (NUMBER), [equals one for an individual fish measured for weight-length], weight of all individuals within a size group or of individual fish (WEIGHT), and length of individual fish (LENGTH) [Line 2, Appendix 1]. This format allows for weight-length calculations for individual fish as well as multiple weight-length calculations for a species size group which can be assembled

without prior summarization. A size group must be assigned to individual fish measured for weight-length relationships. The sample location must be entered in the alpha mode.

The program is written to accept data placed in arbitrary fields, thereby simplifying keypunching. The cards require only field data which is punched in the order specified on the input card (Line 2, Appendix 1) with a blank space between each variable. If the weight or length variable is missing, then a decimal point must be entered. Several examples of data cards are given in Fig. 1. The first example is from Orange Lake: 15 June 1979; sample number 21, first day pickup; species code 4; size group 2; 142 fish in the size group; weight 189 g and a value missing for length (represented by the decimal point). The second example depicts a card for an individual fish weighing 1.8 g and 31 mm long. Example 3 represents a second day pickup with 72 fish in the size group where both weight and length values are missing. Example 4 is from Clear Lake; 3 October 1978; sample number 1; second day pickup of species 7; 6 inch size group; with 9 individuals weighing 1.26 pounds; with length eliminated as an input variable. Weights in pounds and ounces cannot be utilized in this program, therefore, they must be converted to decimal pound weights. Length can be eliminated from the input card (Statement 2, Appendix 1, last entry of card) if all individuals collected in the field are pooled into size groups for weighing and counting.

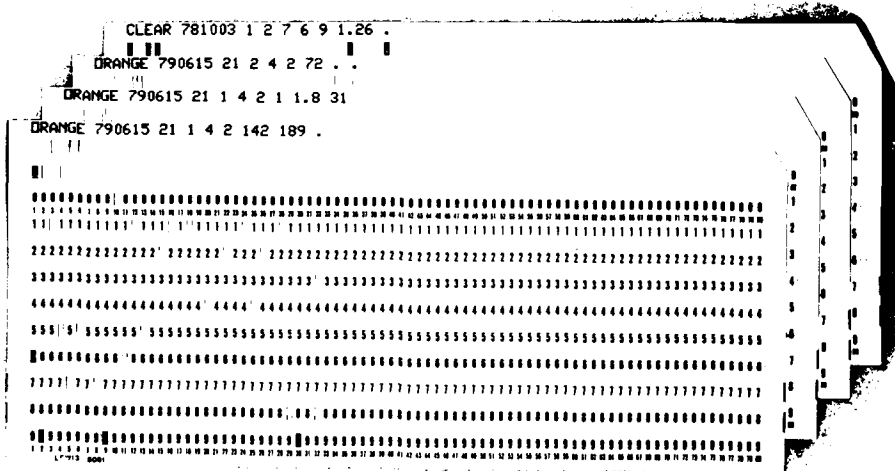


Fig. 1. Examples of data cards for SAS program.

The value assigned a species is a numeric code ranked in increasing phylogenetic order. All known species collected within a sample must be ranked. The first species is assigned number one and each succeeding species the next highest integer. If a new species is encountered later in the study, the new species' phylogenetic rank should be determined and it is given a decimal value between that of the 2 species on either side (Appendix 2). The use of a format statement (Line 5.0 to 5.7, Appendix 1) allows the species name to appear on the computer output rather than the numerical species code.

It is advisable to run only the data for 1 sample location (i.e. lake) at a time unless each rotenone sample is given a consecutive number regardless of sample location.

## DISCUSSION

When missing weights are to be computed for second and third day fish, it is important to have weights for all size groups of a species collected on the first day. When a species size group is collected on second or third day and had no weight on the first day,

then a substitute weight for the size group must be found. If a weight-length regression from the current data for the species is run at program statements 7 and 8 (Appendix 1), then the size group midpoint-length can be used to determine the missing mean size group weight (XSGWT). The mean size group weight can also be obtained from a previous study of the species or from a source such as the *Handbook of Freshwater Fishery Biology* (Carlander 1969, 1977). IF-THEN statements assigning missing mean size group weights (XSGWT) to those species size groups missing on the first day must be constructed and placed between program statements 17 and 18 (Appendix 1). Appendix 3 demonstrates an IF-THEN statement assigning a species missing mean size group weight. Recording lengths of second and third day fish for which weight-length regressions are not run will cause the use of these IF-THEN statements to have no effect. This is a result of by-passing program statement 20 (Appendix 1), because program statement 19 sends the pointer to statement 21 where the value for PWT is missing. If the lengths are recorded for such species they should be punched as missing values on the data cards.

The program presented in Appendix 1 is designed to calculate second and third day weights from field data which includes individual weight-length measurements from fish collected on the first day. If second and third day weights are missing and all data were size grouped (no individual weight-length data recorded first day), then program statements 7,8,9,17, 19 and 21 should be omitted. Various agencies sort fish collected on second and third day pickups into size groups and then record a second and third day size group weight. Program statements 6 through 25 should then be eliminated.

Program statements 7 to 9 instruct the computer to calculate and print regression statistics for weight-length relationships for species when appropriate data are available (Appendix 2). Error statements of the form "ERROR: CLASS VARIABLES OR ALL DEPENDENT VARIABLES ARE MISSING ON EVERY OBSERVATION" will appear on the printout for every species that was not measured for weight-length purposes. These error statements have no effect and can be ignored. Average size group weights by species are computed by statements 10 through 13 and are printed as instructed by statements 14 and 15. Statements 16 through 25 provide instructions for computation and printing of missing weights. Statements 26 through 30 provide for the sums of numbers and weights by species and size group for each sample. It is advisable to run the first 30 statements prior to final analysis to check for data errors and to insure that all missing weights are computed. After proofing the results the data set (FISH) can be stored.

Once the data have been proofed, the remaining program statements provide for data summary. Statements 31 through 35 provide for printing the total number and weight for each species by size group. Length frequency histograms can be constructed from these data. The data are then summarized by sample, giving total numbers and weights for each species (Statements 36 through 46). If a species was not collected in a sample, the data summarization assigns a zero value for number and weight. Tables, with all samples pooled, giving the mean, standard deviation and percentage composition by number and weight for each species are printed as instructed by statements 47 through 59.

A summary of harvestable gamefish is initiated by statement 60. The reader should refer to Appendix 4 or instructions pertaining to IF-THEN statements for harvestable gamefish. All IF-THEN statements utilized should be inserted at program statement 61.0 The harvestable gamefish data summary (Statements 63 through 85) generates the same type tables as discussed in the preceding paragraphs for all species (program statements 36 through 46, and 47 through 59). The gamefish section can also be utilized for forage fish summarization by replacing the word GAMEFISH with the word FORAGE in program statements 60 through 85. The word HARVEST (program statements 60 through 62) should also be replaced with the word PREY. The other changes needed for forage summary are creation of the IF-THEN statements inserted at program statement

61.0. They follow the format outlined in Appendix 4 with the following changes:  
 IF SPECIE=3 AND SG LE 6 THEN PREY = 1;  
 IF SPECIE = 4 AND SG LE 6 THEN PREY = 1;  
 LE represents 'less than or equal to' in the forage summary, while GE stands for  
 'greater than or equal to' in the harvestable analysis.

The utilization of SAS programs for compilation and analysis of rotenone samples allows the researcher to conduct extensive statistical analysis. Statistical comparisons between years, seasons or locations can be conducted for weight-length regressions or total standing crop by insertion of the proper program statements.

#### LITERATURE CITED

- Barr, A.J., J.H. Goodnight, J.P. Sall and J.T. Helwig. 1976. A Users Guide to SAS 76. Sparks Press, Raleigh, N. Carolina. 329 pp.
- Carlander, K.D. 1969. Handbook of freshwater fishery biology, Volume 1. Iowa State Univ. Press, Ames, Iowa. 752 pp.
- \_\_\_\_\_. 1977. Handbook of freshwater fishery biology, Volume 2. Iowa State Univ. Press, Ames, Iowa. 431 pp.
- Hall, G.E. 1962. A survey of fish population sampling methods in Southern reservoirs. Report to Reservoir Committee, So. Div., Am. Fish. Soc. 16 pp.
- Helwig, J.T. 1977. SAS supplemental library users guide. Kingsport Press, Kingsport, Tenn. 171 pp.
- \_\_\_\_\_. 1978. SAS introductory guide. SAS Institute Inc., Raleigh, N. Carolina. 83 pp.
- Provine, W.C. 1975. Evaluation of fish collection devices. Texas Parks and Wildl. Dept. D-J Proj. F-31-R-1:14 pp.
- Swingle, H.S. 1950. Relationships and dynamics of balanced and unbalanced fish populations. Ala. Agr. Expt. Sta. Bull. No. 274. 74 pp.

#### Appendix 1. SAS program for compilation and analysis of rotenone samples.

```

*PROGRAM PRECEDED BY JOB CONTROL CARDS;
1  DATA FISH;
2  INPUT LAKE $ DATE SAMPLE DAY SPECIE SG NUMBER WEIGHT
   LENGTH;
3  LOGWT=LOG10 (WEIGHT); LOGTL=LOG10 (LENGTH);
4  CARDS;
   *ALL DATA CARDS GO HERE BETWEEN STATEMENT 4 AND THIS
   CARD;
5.0  PROC FORMAT;
5.1  VALUE COMNAME
5.2  1=FLORIDA GAR
5.3  2=GOLDEN SHINER
5.4  3=MOSQUITOFISH
5.5  4=BLUEGILL
5.6  5=LARGEMOUTH BASS
5.7  ;
   *COMPUTING WEIGHT-LENGTH LINES AND PRINTING THE
   REGRESSION STATISTICS;
6  PROC SORT; BY LAKE SPECIE SG LENGTH;
7  PROC GLM; FORMAT SPECIE COMNAME.; BY LAKE SPECIE;
8  MODEL LOGWT=LOGTL; OUTPUT OUT=PREDWT PREDICTED=PWT;

```

```

9  TITLE WEIGHT-LENGTH REGRESSION STATISTICS;
  *COMPUTING AVERAGE SIZE GROUP WEIGHTS;
10 DATA MEANSGWT; SET FISH; IF WEIGHT NE .;
11 PROC MEANS NOPRINT; BY LAKE SPECIE SG; VAR NUMBER WEIGHT;
12 OUTPUT OUT=MEANSGWT SUM=NO WT;
13 DATA MEANSGWT; SET MEANSGWT; XSGWT=WT/NO;
14 PROC PRINT; FORMAT SPECIE COMNAME.;
15 TITLE AVERAGE SIZE GROUP WEIGHTS;
  *COMPUTING MISSING WEIGHTS FOR FISH;
16 DATA FISH; MERGE MEANSGWT FISH; BY LAKE SPECIE SG; DROP NO
  WT;
17 DATA FISH; MERGE PREDWT FISH; BY LAKE SPECIE; PWT=10**PWT;
18 IF WEIGHT NE . THEN GO TO JUMPA;
19 IF WEIGHT NE . THEN GO TO JUMPB;
20 WEIGHT=XSGWT*NUMBER;
21 GO TO JUMPA; JUMPB: WEIGHT=PWT;
22 JUMP A: RETURN;
23 PROC SORT; BY LAKE SPECIE WEIGHT;
24 PROC PRINT; FORMAT SPECIE COMNAME.;
25 TITLE DATA SET WITH COMPUTED MISSING WEIGHTS;
  *SUMMARIZING THE NUMBER AND WEIGHT OF THE DATA BY
  SAMPLE, SPECIE AND SG;
26 PROC SORT; BY LAKE SAMPLE SPECIE SG DATE;
27 PROC MEANS NOPRINT; BY LAKE SAMPLE SPECIE SG DATE; VAR
  NUMBER WEIGHT;
28 OUTPUT OUT=FISH SUM=NUMBER WEIGHT;
29 PROC PRINT; FORMAT SPECIE COMNAME.;
30 TITLE THE NUMBER AND WEIGHT OF FISH FROM EACH SIZE GROUP
  FOR EACH SAMPLE;
  *COMPUTING THE LENGTH FREQUENCY OF FISHES FROM ALL
  SAMPLES;
31 PROC SORT; BY LAKE SPECIE SG;
32 PROC MEANS NOPRINT; BY LAKE SPECIE SG; VAR NUMBER WEIGHT;
33 OUTPUT OUT=LENQFREQ SUM=NUMBER WEIGHT;
34 PROC PRINT; FORMAT SPECIE COMNAME.;
35 TITLE LENGTH FREQUENCY DATA OF FISH FROM ALL SAMPLES;

  *SUMMARIZING THE TOTAL NUMBER AND WEIGHT OF FISHES BY
  SAMPLE;
36 DATA FISHB; SET FISH; WEIGHT=WEIGHT*100;
37 PROC SORT; BY LAKE SAMPLE SPECIE;
38 PROC FREQ;
39 TABLES LAKE*SAMPLE*SPECIE/OUT=A SPARSE NOPRINT; WEIGHT
  NUMBER;
40 DATA A; SET A; NUMBER=COUNT; DROP COUNT PERCENT;
41 PROC FREQ DATA=FISHB;
42 TABLES LAKE*SAMPLE*SPECIE/OUT=B SPARSE NOPRINT; WEIGHT
  WEIGHT;
43 DATA B; SET B; WEIGHT=COUNT; DROP COUNT PERCENT;
44 DATA FISHB; MERGE A B; BY LAKE SAMPLE SPECIE; WEIGHT
  WEIGHT/100;
45 PROC PRINT; FORMAT SPEICE COMNAME.;
46 TITLE TOTAL NUMBER AND WEIGHT OF FISHES BY SAMPLE;
  *SUMMARIZING THE NUMBER, WEIGHT AND PERCENT

```

```

COMPOSITION OF FISH FOR ALL SAMPLES;
47 PROC SORT; BY LAKE SPECIE;
48 PROC MEANS NOPRINT; BY LAKE SPECIE; VAR NUMBER WEIGHT;
49 OUTPUT OUT=MEANS MEAN=XNO XWT STD=SNO SWT;
50 DATA MEANS; SET MEANS; XWT=XWT*100;
51 PROC FREQ;
52 TABLES SPECIE/OUT=A SPARSE NOPRINT; WEIGHT XNO;
53 DATA A; SET A; PERCENTN=PERCENT; DROP COUNT PERCENT;
54 PROC FREQ DATA=MEANS
55 TABLES SPECIE/OUT=B SPARSE NOPRINT; WEIGHT XWT;
56 DATA B; SET B; PERCENTW=PERCENT; DROP COUNT PERCENT;
57 DATA MEANS; MERGE MEANS A B; BY SPECIE; XWT=XWT/100;
58 PROC PRINT; FORMAT SPECIE COMNAME.;
59 TITLE THE NUMBER, WEIGHT AND PERCENT COMPOSITION OF
FISHES FOR ALL SAMPLES;
*SEPARATING HARVESTABLE GAMEFISH FROM THE DATA FOR
SUMMARIZATION;
60 DATA GAMEFISH; SET FISH; WEIGHT=WEIGHT*100; HARVEST=0;
61.0 IF SPECIE=4 AND SG GE 6 THEN HARVEST=1;
61.1 IF SPECIE=5 AND SG GE 10 THEN HARVEST=1;
62 IF HARVEST NE 1 THEN DELETE;
*SUMMARIZING THE NUMBER AND WEIGHT OF GAMEFISH BY
SAMPLE;
63 PROC SORT; BY LAKE SAMPLE SPECIE;
64 PROC FREQ;
65 TABLES LAKE*SAMPLE*SPECIE/OUT=A SPARSE NOPRINT;
WEIGHT NUMBER;
66 DATA A; SET A; NUMBER=COUNT; DROP COUNT PERCENT;
67 PROC FREQ DATA=GAMEFISH;
68 TABLES LAKE*SAMPLE*SPECIE/OUT=B SPARSE NOPRINT;
WEIGHT WEIGHT;
69 DATA B; SET B; WEIGHT=COUNT; DROP COUNT PERCENT;
70 DATA GAMEFISH; MERGE A B; BY LAKE SAMPLE SPECIE;
WEIGHT=WEIGHT/100;
71 PROC PRINT; FORMAT SPECIE COMNAME.;
72 TITLE THE NUMBER AND WEIGHT OF GAMEFISH BY SAMPLE;
*SUMMARIZING NUMBER, WEIGHT AND PERCENT COMPOSITION
OF GAMEFISH FOR ALL SAMPLES;
73 PROC SORT; BY LAKE SPECIE;
74 PROC MEANS NOPRINT; BY LAKE SPECIE; VAR NUMBER WEIGHT;
75 OUTPUT OUT=MEANS MEAN=XNO XWT STD=SNO SWT;
76 DATA MEANS; SET MEANS; XWT=XWT*100;
77 PROC FREQ;
78 TABLES LAKE*SPECIE/OUT=A SPARSE NOPRINT; WEIGHT XNO;
79 DATA A; SET A; PERCENTN=PERCENT; DROP COUNT PERCENT;
80 PROC FREQ DATA=MEANS;
81 TABLES LAKE*SPECIE/OUT=B SPARSE NOPRINT; WEIGHT XWT;
82 DATA B; SET B; PERCENTW=PERCENT; DROP COUNT PERCENT;
83 DATA MEANS; MERGE MEANS A B; BY LAKE SPECIE;
84 PROC PRINT; FORMAT SPECIE COMNAME.;
85 TITLE THE NUMBER, WEIGHT AND PERCENT COMPOSITION OF
GAMEFISH FOR ALL SAMPLES;
*END OF JOB CARD FOLLOWS:

```

The program statement numbers should not be punched on cards. They are included here for reference purposes. Each program statement represents one computer card.

Appendix 2. Numeric code for species identification.

PROC FORMAT;  
VALUE COMNAME

1 =FLORIDA GAR  
2 =BOWFIN  
3 =GIZZARD SHAD  
4 =GOLDEN SHINER  
4.1=TAILLIGHT SHINER  
5 =LAKE CHUBSUCKER  
6 =MOSQUITOFISH  
7 =BLUEGILL  
8 =REDEAR  
9 =LARGEMOUTH BASS  
10 =BLACK CRAPPIE

PROC FORMAT;  
VALUE SCINAME

1 =LEPISOSTEUS PLAYTYRHINCUS  
2 =AMIA CALVA  
3 =DOROSOMA CEPEDIANUM  
4 =NOTEMIGONUS CRYSOLEUCAS  
4.1 =NOTROPIS MACULATUS  
5 =ERIMYZON SUCETTA  
6 =GAMBUSIA AFFINIS  
7 =LEPOMIS MACROCHIRUS  
8 =LEPOMIS MICROLOPHUS  
9 =MICROPTERUS SALMOIDES  
10 =POMOXIS NIGROMACULATUS

Two examples of procedure FORMAT in which the species code ranks fish phylogenetically. The species code 4.1 representing taillight shiner was inserted after the original ranking was made. This section is represented by statements 5.0 through 5.7 in Appendix (1).

Appendix 3. IF-THEN statements assigning a missing specie size group weight.  
IF SPECIE=4 AND SG=2 THEN XSGWT=8;

An example of an IF-THEN statement assigning 8 grams as a mean size group weight (XSGWT) for specie 4 in size group (SG) 2. As many statments as are needed can be used for different species size groups. These IF-THEN statements should be inserted between program statements 17 and 18.

Appendix 4. IF-THEN statements for harvestable fish analysis.

IF- SPECIE= 7 AND SG GE 6 THEN HARVEST=1;  
IF SPECIE= 8 AND SG GE 6 THEN HARVEST=1;  
IF SPECIE= 9 AND SG GE 10 THEN HARVEST=1;  
IF SPECIE=10 AND SG GE 9 THEN HARVEST=1;

The size groups (SG) represent Swingle's (1950) minimum harvestable inch groups for the game fish coded in Appendix 2. The user can insert any predetermined size group after the GE statement selected. the number of cards will vary with the users needs. This section is represented by statements 61.0 and 61.1 in Appendix 1.