Northeast Gulf Science

Volume 5
Number 1 *Number 1*

Article 14

10-1981

The Use of Cluster Analysis in Comparing Multi-Source Feeding Studies

Stephen A. Bortone
University of West Florida

David Seigel
University of West Florida

Jerry L. Oglesby
University of West Florida

DOI: 10.18785/negs.0501.14

Northeast Gulf Science Vol. 5 No. 1 October 1981 **81**

THE USE OF CLUSTER ANALYSIS IN COMPARING MULTI-SOURCE FEEDING STUDIES.

It has been suggested that feeding studies of fishes should be conducted using at least three basic parameters: volumetric (or gravimetric) displacement, frequency of occurrence, and frequency (or number) of food items (Windell, 1971). Analysis by these three methods may avoid biases introduced by using only one method in a study. These methods, along with consistent techniques in specimen capture and data determination, permit statistical comparisons of fishes based on their feeding habits (Fritz 1974; Bray and Ebeling 1975: Bortone, Rebenack, and Siegel in press).

Often an investigator may be interested in comparing one's own data on feeding habits with data compiled by a number of other investigators. A problem immediately arises as frequently inconsistent methods have been employed to obtain the various data sets. This feature may render many parametric statistical comparisons useless or invalid in that one or more essential assumptions are not met.

Cluster analysis has been employed in ecological and taxonomic studies to compare large data sets (Sneath and Sokal 1973). Cluster analysis provides a relatively simple way of preparing a numerically based visual comparison of complex and seemingly unorganizable data. Ross (1977, 1978) demonstrated the usefulness of cluster analysis in comparing stomach contents from different species and size groups of sea robins (Triglidae).

In the present study, we demonstrate the utility of cluster analysis in making multi-source data comparisons of fishes based on food studies from several published and non-published sources.

METHODS AND MATERIALS

The operational taxonomic units (OTU's, the groups to be compared) were fish species sampled from several populations. To demonstrate the utility of cluster analysis, we chose to compare species of the family Lutjanidae. The species, areas compared and data sources were: *Lutjanus campechanus,* Texas — Louisiana (OTU 1-9), (Mosley 1966); *L. campechanus,* Texas (OTU 10-17), (Bradley and Bryan 1974); *Rhomboplites aurorubens,* North-South Carolina (OTU 18-21), (Grimes 1979); *L. campechanus, L. synagris* and *R. aurorubens*, Louisiana — NW Florida (OTU 22-26), (Siegel and Bortone unpublished). The above OTU's were identified (when data were available) by the following code:
C=*L. campechanus,* S=*L. synagris,* R= *R. aurorubens,* A=adult (> 250 mm standard length), J=juvenile (<250 mm standard length), H=captured by hook-and-line, T=trawl captured, X=specimens from Texas (Mosley 1966), M - Louisiana (Mosley 1966), Xb - Texas (Bradley and Bryan 1974), G=North-South Carolina (Grimes 1979), Mo=Mobile, Alabama, B=Orange Beach, Alabama, P=Pensacola, Florida (Siegel and Bortone unpublished), Su=Summer, F=Fall, W=Winter, Sp=Spring. The 26 OTU's were compared by the frequency of occurrence, volumetric displacement, or presence/absence of 26 food groups or types: Alphaeidae, Amphipoda, Ascidacea, Cephalopoda, Chaetognatha, Copepoda, Crustacean parts, Ctenophora, Cumacea, Decapod larvae, detritus, Euphausiacea, fish, fish larvae, Gastropoda, Isopoda, Mysidacea, Ostracoda, Pelecypoda, Penaeidea, Polychaeta, Reptantia, Scaphopoda, Sipuncula, Stomatopoda, Taenidia (Table 1).

Cluster analyses were conducted according to the methods of numerical

Table 1: Composite summary of multi-source feeding data. Data are presented as precent volume and/or percent occurrence (in parentheses) when available. Code for OTU's is presented in the text.

FOOD ITEM

	Penaeidea	Mysidacea	Stomatopoda	Alphaeidae	Reptantia	Crustacean Parts	Polychaeta	Copepoda	Pelecypoda	Cephalopoda	Fish Larvae	Ascidiacea	Fish	Detritus	Chaetognatha	Ctenophora	Gastropoda	Ostracoda	Amphipoda	Euphausiacea	Decapod Larvae	Sipuncula	Taenidia	Isopoda	Scaphopoda	Cumacea
#1	0	0	0	0	1	0	0	0	0	6	0	53	40	0	0	0	0	0	0	0	0	0	0	0	0	0
C,A,H,X,Su	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
#2	44	11	0	0	13	32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C,J,T,X,F	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
#3	37	0	0	0	0	0	10	0	0	0	0	0	53	0	0	0	0	0	0	0	0	0	0	0	0	0
C,A,H,X,F-W	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
#4	20	0	0	0	12	0	0	0	9	0	0	0	59	0	0	0	0	0	0	0	0	0	0	0	0	0
C,A,H,X,F-W	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
#5	31	25	0	0	0	10	2	0	0	11	0	0	7	14	0	0	0	0	0	0	0	0	0	0	0	0
C,J,T,X,F	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
#6	2	0	0	0	0	0	0	0	6	0	0	45	44	3	0	0	0	0	0	0	0	0	0	0	0	0
C,A,H,M,Sp	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—		—	—	—	—	—	—	—	—	—	—
#7	10	0	0	0	10	0	0	0	0	0	0	0	80	0	0	0	0	0	0	0	0	0	0	0	0	0
C,A,H,M,Sp	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
#8	29	0	0	0	0	10	0	0	0	0	0	0	21	23	17	0	0	0	0	0	0	0	0	0	0	0
C,J,T,M,W-Sp	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
#9	0	0	60	0	0	13	0	0	0	0	0	0	27	0	0	0	0	0	0	0	0	0	0	0	0	0
C,J,T,M,F	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
#10	0	0	17	0	10	0	0	0	0	41	0	0	20	12	0	0	0	0	0	0	0	0	0	0	0	0
C,J,T,Xb,Su	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
#11	0	0	0	0	46	0	0	0	0	0	0	0	34	21	0	0	0	0	0	0	0	0	0	0	0	0
C,A,H,Xb,Su	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—						
#12	34	0	0	0	0	0	0	0	0	45	0	0	16	5	0	0	0	0	0	0	0	0	0	0	0	0
C,J,T,Xb,F	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
#13	10	0	0	0	6	0	0	0	0	0	0	0	78	6	0	0	0	0	0	0	0	0	0	0	0	0
C,A,H,Xb,F	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
#14	27	0	0	11	7	21	0	9	0	0	0	0	14	11	0	0	0	0	0	0	0	0	0	0	0	0
C,J,T,Xb,W	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
#15	0	0	0	0	0	0	0	0	0	0	0	0	97	3	0	0	0	0	0	0	0	0	0	0	0	0
C,A,H,Xb,W	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
#16	75	0	0	0	0	12	0	0	0	0	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0
C,J,T,Xb,Sp	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
#17	7	0	8	0	0	0	0	0	0	0	0	0	77	8	0	0	0	0	0	0	0	0	0	0	0	0
C,A,H,Xb,Sp	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
#18	12	0	5	0	3	0	2	1	0	38	3	1	1	0	0	2	8	1	14	0	4	0	0	0	0	0
R,A,H,G,Sp	(73)	(0)	(31)	(0)	(9)	(0)	(12)	(3)	(0)	(36)	(21)	(4)	(4)	(0)	(0)	(12)	(18)	(25)	(58)	(0)	(47)	(0)	(0)	(0)	(0)	(0)
#19	6	0	3	0	0.5	0	0	0	0	34	0.5	9	8	0	0	1	24	1	9	0.5	4	0	0	0	0	0
R,A,H,G,Su	(50)	(0)	(30)	(0)	(3)	(0)	(0)	(0)	(0)	(37)	(9)	(9)	(27)	(0)	(0)	(4)	(11)	(30)	(59)	(2)	(40)	(0)	(0)	(0)	(0)	(0)
#20	5	0	9	0	2	0	0.5	2	0	37	0.5	0.5	21	0	0	1	1	1	11	0	5	0	0	0	0	0
R,A,H,G,F	(22)	(0)	(30)	(0)	(9)	(0)	(6)	(11)	(0)	(32)	(5)	(2)	(16)	(0)	(0)	(2)	(6)	(19)	(26)	(0)	(24)	(0)	(0)	(0)	(0)	(0)
#21	68	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0.5	0	6	0	0	0	0	0
R,A,H,G,W	(99)	(0)	(50)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(50)	(50)	(0)	(50)	(0)	(0)	(0)	(0)	(0)
#22	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
C,J,H,B,F	(25)	(8)	(0)	(0)	(31)	(23)	(10)	(27)	(0)	(13)	(21)	(2)	(44)	(0)	(0)	(0)	(0)	(0)	(23)	(0)	(10)	(4)	(4)	(4)	(0)	(0)
#23	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
C,J,T,Mo,Su	(17)	(7)	(31)	(10)	(1)	(9)	(1)	(1)	(0)	(7)	(0)	(1)	(46)	(0)	(0)	(0)	(0)	(0)	(3)	(7)	(0)	(0)	(0)	(0)	(0)	(0)
#24	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
S,J,T,Mo,Su	(33)	(0)	(0)	(11)	(0)	(11)	(11)	(0)	(0)	(0)	(0)	(0)	(33)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
#25	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
R,A,H,P,W	(33)	(33)	(17)	(0)	(0)	(50)	(17)	(5)	(0)	(17)	(0)	(0)	(50)	(0)	(0)	(0)	(17)	(17)	(17)	(0)	(17)	(0)	(0)	(17)	(0)	(17)
#26	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
R,J,H,P,W	(50)	(5)	(11)	(0)	(0)	(67)	(76)	(80)	(4)	(35)	(0)	(0)	(12)	(0)	(0)	(0)	(24)	(52)	(23)	(0)	(17)	(0)	(0)	(3)	(5)	(32)

taxonomy (Sneath and Sokal 1973) using the NT-SYS program (Rolf, Kispaugh, and Kirk 1971). The OTU/Character data matrices were standardized, and then used to calculate Pearson product (r) correlation coefficients (maximum similarity = +1.0, minimum =-1.0) and/or Jaccard coefficients of association (maximum =+1.0, minimum = 0). The calculated correlation and association data matrices were each clustered with the unweighted pair-group method using arithmetic averages (UPGMA) according to Sneath and Sokal (1973).

Cophenetic values, which indicate the association (maximum =+1.0, minimum = -1.0) between the original similarity matrix and the matrix produced by the clustering algorithm, were calculated for each cluster.

RESULTS

The cluster analyses of lutjanid fish populations based on stomach content studies are depicted in the cluster diagrams in Figures 1-5. Figure 1 depicts a clustering of OTU's 18-26 based on the

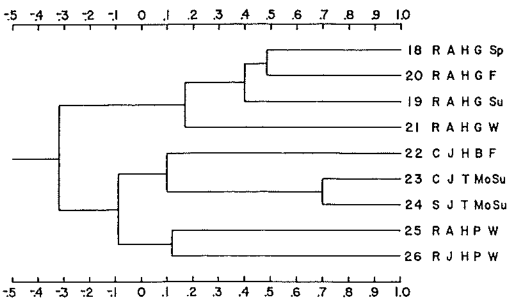

Figure 1. Cluster diagram of OTU's 18-26 based on the correlation coefficients obtained from percent occurrence data. Letter code describing each OTU is presented in the text. Cophenetic value = 0.904. (The first letter set in the code designates the species: C= *Lutjanus campechanus*, S=*L. synagris*, R=*Rhomboplites aurorubens*. The second letter set indicates life stage: J=juvenile, A=adult. The third letter set indicates method of capture: H= hook-and-line, T=trawl. The fourth indicates area of capture: X and Xb=Texas, M=Louisiana, Mo and B= Alabama, P=Florida, G=North and South Carolina. Fifth indicates season: Su=summer, F=fall, W= winter, Sp=spring).

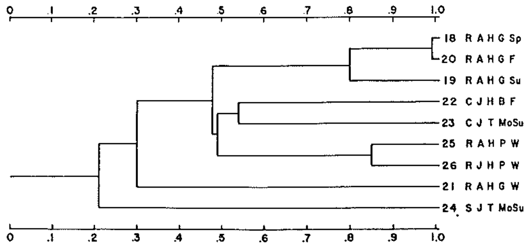

Figure 2. Cluster diagram of OTU's 18-26 based on Jaccard coefficients obtained from percent occurrence data. Letter code describing each OTU is presented in the text. Cophenetic value=0.952.

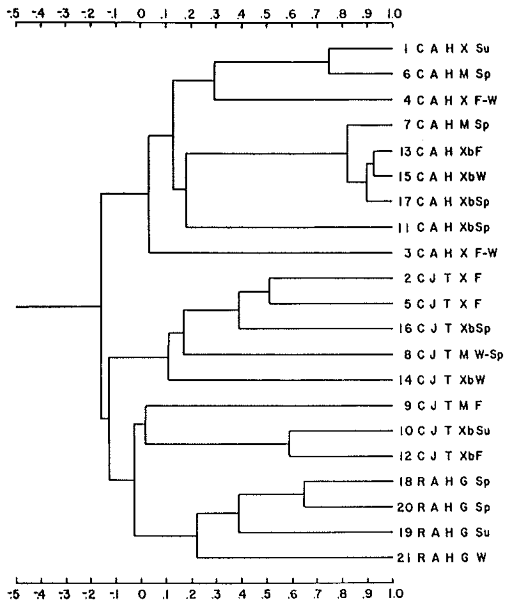

Figure 3. Cluster diagram of OTU's 1-21 based on correlation coefficients obtained from percent volume data. Letter code describing each OTU is presented in the text. Cophenetic value=0.864.

correlation coefficients of the percent occurrence of food items. *Rhomboplites aurorubens* adults from North-South Carolina (OTU's 18-21) formed a cluster separate from *R. aurorubens* juveniles and adults from Pensacola (OTU's 25 and 26) and other *Lutjanus* spp. (OTU's 23 and 24).

A presence/absence (Jaccard coefficient) clustering of the same data base (OTU's 18-26) displayed a slightly different cluster (Fig. 2). In this cluster *R. aurorubens* adults from North-South Carolina formed a distinct cluster with the exception of winter caught specimens

(OTU 21). Additionally, the high association of *L. campechanus* with *L. synagris* (Fig. 1) was not depicted in the presence/absence analysis.

OTU's 1-21 were clustered according to correlation coefficients of the percent volumetric contribution of each food item (Fig. 3). Examination of the clusters formed by this analysis reveals two major groups. The first was formed by the clustering of all adult *L. campechanus* and the second by juvenile *L. campechanus* and adult *R. aurorubens*. Within this second group there was also a cluster composed of *R. aurorubens* adults.

The cluster analysis of the same data base (percent volume of food items) as in the previous analysis was conducted, but only the presence/absence (Jaccard coefficient) of the food items was used in the comparison. Examination of this cluster analysis in Fig. 4 reveals two clusters above the 0.2 level of similarity. The first cluster was composed of *L. campechanus*, while the second cluster was composed solely of *R. aurorubens*. Within the first cluster there was, with a few exceptions (OTU's 9 and 12), a tendency for adult and juvenile *L. campechanus* to cluster separately.

Fig 5. represents a composite cluster based on the presence/absence (Jaccard coefficient) of the association of all food item data used in the present study (OTU's 1-26). At first inspection there appeared to be few, if any, discernable patterns in the data clustering. Although *R. aurorubens* (OTU's 18-20) juveniles and adult *L. campechanus* tended to cluster together, there were many inconsistencies. However, reexamination of the original data matrix (Table 1) indicates that the two major clusters which joined at the 0.2 level of similarity were separated by the number of food types consumed. The upper cluster had a larger number of food types ($\bar{x} = 13$ items) while the lower cluster had fewer food types present ($\bar{x} = 4$ items).

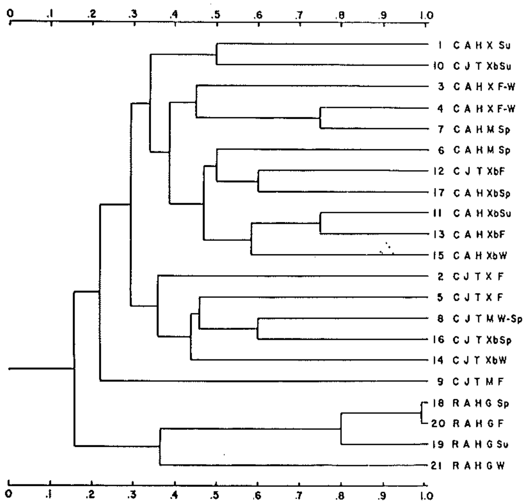

Figure 4. Cluster diagram of OTU's 1-21 based on Jaccard coefficients obtained from percent volume data. Letter code describing each OTU is presented in the text. Cophenetic value=0.807.

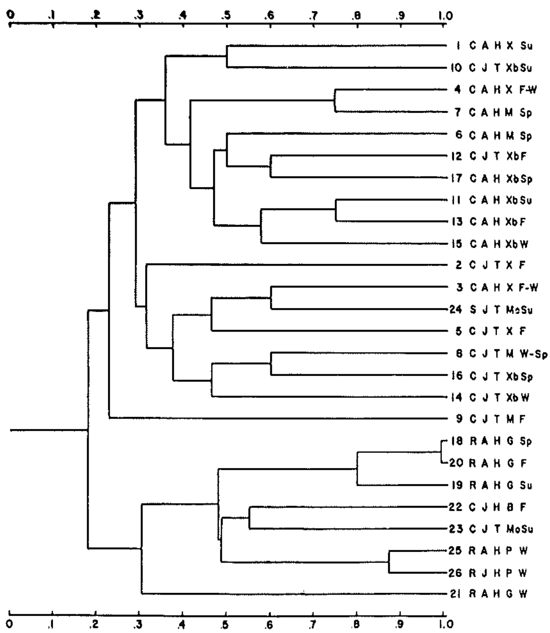

Figure 5. Cluster diagram of OTU's 1-26 based on Jaccard coefficients obtained from presence/absence of all data sources. Letter code describing each OTU is presented in the text. Cophenetic value =0.815.

DOI: 10.18785/negs.0501.14

DISCUSSION

An examination of the clusters formed in the previous analyses can be useful in interpreting the feeding relationships of fish. For example, one might expect fish of a single species to always cluster together implying an intraspecific consistency in feeding. However, examination of Fig. 1 indicates that a fish species from one area may not be similar to conspecifics from another area, indicating that local areas may differ with regard to available food or that there are real population differences among the OTU's. Also, when we clustered the OTU's with food item volume by correlation coefficients we noted an apparent discrepancy in the cluster of *R. aurorubens* adults with *Lutjanus* spp. juveniles. However, because of the nature of *R. aurorubens* to feed on small, free swimming organisms as opposed to larger, benthic organisms (Dixon 1975; Grimes 1979), we have probably described this association accurately.

There is an apparent lack of a correlated parameter to explain the cluster formed by the presence/absence analysis of all data (Fig. 5). As we have indicated, this may be due to a lack of diversity. This observation implies that the use of the Jaccard coefficient (based on presence/absence data) may be useful only when the diversity of food items is high or when all OTU's have nearly equal diversity.

Ross (1977, 1978) has demonstrated the usefulness of cluster analysis in comparing the food habits of fishes. Our analysis has demonstrated that cluster analysis may also be used to compare data from several different sources. There are, however, some qualifications one should consider before using this method: (1) food types must be put into similar taxonomic categories; (2) presence/absence clustering of OTU's which have a low diversity of food items may distort real trophic relationships; (3) each OTU should be a representative sample from the population and the number of fishes used to form each OTU should be at least 30-40 specimens Sneath and Sokal 1973).

ACKNOWLEDGMENTS

We thank S.T. Ross for his advice and constructive criticism of an early draft of the manuscript. This article was developed under the auspices of the Florida Sea Grant College with support from the National Oceanic and Atmospheric Administration, Office of Sea Grant, U.S. Department of Commerce, Grant No. 04-8-MO1-76.

LITERATURE CITED

Bortone, S.A., P. Rebenack, and D.M. Siegel. (in press). A comparative study between sympatric populations of *Diplectrum formosum* and *D. bivittatum* (Pisces: Serranidae). Florida Scientist.

Bray, R.N., and A.W. Ebeling. 1975. Food, activity, and habitat of three "picker-type" microcarnivorous fishes in the kelp forests off Santa Barbara, California. Fish. Bull. 73:815-829.

Bradley, I., and C.E. Bryan. 1974. Life history and fishery of the red snapper *Lutjanus campechanus* in the northwestern Gulf of Mexico: 1970-74. Proc. 27th Ann. Sess. Gulf Carib. Fish. 77-106.

Dixson, R.L. 1975. Evidence for mesopelagic feeding in the vermilion snapper, *Rhomboplites aurorubens.* J. Elisha Mitchell Sci. Soc. 91:240-242.

Fritz, E.S. 1974. Total diet comparison in fishes by Spearmann rank correlation coefficient. Copeia 1974: 210-214.

Grimes, C.B. 1979. Diet and feeding ecology of the vermilion snapper, *Rhomboplites aurorubens* (Cuvier), from North Carolina and South Caro-

 Gulf of Mexico Science, Vol. 5 [1981], No. 1, Art. 14

lina waters. Bull Mar. Sci. 29:53-61.

Mosley, F. 1966. Biology of the red snapper, *Lutjanus aya* Bloch, of the northwestern Gulf of Mexico. Publ. Inst. Mar. Sci. Univ. Texas. 11: 90-101.

Rohlf, F.J., J. Kishpaugh, and D. Kirk. 1971. NT-SYS. Numerical taxonomy-system of multivariate statistical programs. Tech. Rep. State University of New York at Stony Brook, New York.

Ross, S.T. 1977. Patterns of resource partitioning in sea robins (Pisces: Triglidae). Copeia 1977: 561-571.

__________. 1978. Trophic ontogeny of the leopard sea robin, *Prionotus scitulus* (Pisces: Triglidae). Fish Bull. 76:225-234.

Sneath, P.H.A., and R.R. Sokal. 1973. Numerical taxonomy, the principles and practice of numerical classification. W.H. Freeman and Co., San Francisco. 573 p.

Windell, J.T. 1971. Food analysis and rate of digestion, p. 215-226. *In* W.E. Ricker (ed.), Methods for assessment of fish production in freshwaters. 2nd ed. Blackwell Sci. Publ., Oxford.

Stephen A. Bortone and David Seigel, *Department of Biology, University of West Florida, Pensacola, FL 32504.*

Jerry L. Oglesby, *Institute for Statistics and Mathematical Modeling, University of West Florida, Pensacola, FL 32504.*

DOI: 10.18785/negs.0501.14