Reprinted from BULLETIN OF MARINE SCIENCE
Vol. 27, No. 3, July, 1977
Made in United States of America

A PACKAGE OF COMPUTER PROGRAMS FOR BENTHIC COMMUNITY ANALYSES

Stephen A. Bloom, Stuart L. Santos, and John G. Field

ABSTRACT—A package of computer programs which can sort and transfer benthic community organismal data, aid in the analysis of sedimentary data, calculate a nonparametric correlation statistic, and perform a series of multivariate techniques (ordination and classification) including plotting of the results is described. The programs are written in FORTRAN IV for use on an IBM 360/370 computer. They will be made available at cost.

Complex numerical techniques are commonly being utilized in marine ecological analyses and many have been programmed for a variety of computer systems (Hill et al., 1975; Hill, 1973; Stephenson, 1972). Unfortunately many investigators do not have access to installations which maintain such programs nor can they devote sufficient time to program or to adapt programs to available computer systems. To help alleviate this problem and in an effort to make valuable analytical tools available to potential users, we have written and compiled a package of programs which includes test data and a complete use-and-debug manual. All programs have been written in FORTRAN IV for use on an IBM 360/370 computer.

PROGRAM DESCRIPTIONS

DATPUN

DATPUN is a data conversion, compilation and punch program which converts counts per sample to a standard data matrix

as well as producing certain statistics for each station. DATPUN input consists of cards punched with counts per sample, a reference code and a species code. The codes are arbitrarily chosen by the user and can be data retrieval codes for computer storage (Field, 1971). The program creates the species list, performs the necessary summations and conversions and punches a card deck. The card deck corresponds to the usual matrix of benthic ecology, i.e. columns as stations, rows as species and tabular values on a per-square-meter basis. Printed output also includes a series of station statistics such as mean, standard deviation and coefficient of dispersion.

SPAREA

SPAREA accepts the same data deck as DATPUN and generates five plots of the species-area relationship. The first is a simple plot of the accumulating number of new species added, plotted against the increasing number of samples utilized. The succeeding four plots are generated by randomly scrambling the order of samples and repeating the plotting procedure described above.

ORDANA

ORDANA performs classification and ordination analyses with printed, punched and graphical output. Input to the program consists of the card deck produced by DATPUN and the appropriate parameter cards.

There are five data transformation options available: none, presence/absence, natural log ($x = \ln(x + 1)$), conversion to percent per sample, and an exponential transformation. The user supplies the exponent and thus the transformation can also be a root transformation. There are also four data standardizations and 13 similarity/dissimilarity indices available (Table 1). Data transformation and standardization can only be performed once in each run.

The two major options in ORDANA are classification and ordination. Each is independent of the other. The user indicates whether he wishes the resulting dendrogram (classification) to be plotted, which of the similarity/dissimilarity indices is to be used and which sorting strategy is to be utilized. The sorting strategies presently available are nearest-neighbor, furthest-neighbor, flexible and group-average. Printed output consists of the initial similarity matrix, an analysis of the distribution of the similarity values, a linkage-by-linkage summary of the similarity values for each fusion cycle and the samples involved and the dendrogram sample sequence.

The user also indicates whether he wishes the ordination analysis to be performed, and if so, whether it is to be plotted (plots of axes I to II, III, IV and V and perspective plots of axes I, II and III from two different vantage points). The user can also indicate the similarity/dissimilarity index utilized. Selection of the appropriate transformation, standardization and index allows the choice between principal components analysis and principal co-ordinates analysis (Gower, 1966; Sneath and Sokal, 1973). Printed output consists of the initial similarity/dissimilarity matrix, the latent roots and vectors of the matrix and the percent efficiency of each vector (amount of variance contributed by each axis after rotation).

The program allows normal (Q) and inverse (R) analyses. The program will also, at the user's option, punch out similarity linkage cards for the classification and the table of latent vectors from the ordination.

SPRMAN

Spearman-Rank Correlation analysis (Siegel, 1956). SPRMAN will accept two matrices and will rearrange columns and rows as directed. It will then pairwise compare all columns of one matrix to all columns of the other matrix and will generate the appropriate statistics. There is also a label option which allows labelling of all comparisons by the names (supplied by the user) of the two variables being compared.

as well as producing certain statistics for each station. DATPUN input consists of cards punched with counts per sample, a reference code and a species code. The codes are arbitrarily chosen by the user and can be data retrieval codes for computer storage (Field, 1971). The program creates the species list, performs the necessary summations and conversions and punches a card deck. The card deck corresponds to the usual matrix of benthic ecology, i.e. columns as stations, rows as species and tabular values on a per-square-meter basis. Printed output also includes a series of station statistics such as mean, standard deviation and coefficient of dispersion.

SPAREA

SPAREA accepts the same data deck as DATPUN and generates five plots of the species-area relationship. The first is a simple plot of the accumulating number of new species added, plotted against the increasing number of samples utilized. The succeeding four plots are generated by randomly scrambling the order of samples and repeating the plotting procedure described above.

ORDANA

ORDANA performs classification and ordination analyses with printed, punched and graphical output. Input to the program consists of the card deck produced by DATPUN and the appropriate parameter cards.

There are five data transformation options available: none, presence/absence, natural log ($x = \ln(x + 1)$), conversion to percent per sample, and an exponential transformation. The user supplies the exponent and thus the transformation can also be a root transformation. There are also four data standardizations and 13 similarity/dissimilarity indices available (Table 1). Data transformation and standardization can only be performed once in each run.

The two major options in ORDANA are classification and ordination. Each is independent of the other. The user indicates whether he wishes the resulting dendrogram (classification) to be plotted, which of the similarity/dissimilarity indices is to be used and which sorting strategy is to be utilized. The sorting strategies presently available are nearest-neighbor, furthest-neighbor, flexible and group-average. Printed output consists of the initial similarity matrix, an analysis of the distribution of the similarity values, a linkage-by-linkage summary of the similarity values for each fusion cycle and the samples involved and the dendrogram sample sequence.

The user also indicates whether he wishes the ordination analysis to be performed, and if so, whether it is to be plotted (plots of axes I to II, III, IV and V and perspective plots of axes I, II and III from two different vantage points). The user can also indicate the similarity/dissimilarity index utilized. Selection of the appropriate transformation, standardization and index allows the choice between principal components analysis and principal co-ordinates analysis (Gower, 1966; Sneath and Sokal, 1973). Printed output consists of the initial similarity/dissimilarity matrix, the latent roots and vectors of the matrix and the percent efficiency of each vector (amount of variance contributed by each axis after rotation).

The program allows normal (Q) and inverse (R) analyses. The program will also, at the user's option, punch out similarity linkage cards for the classification and the table of latent vectors from the ordination.

SPRMAN

Spearman-Rank Correlation analysis (Siegel, 1956). SPRMAN will accept two matrices and will rearrange columns and rows as directed. It will then pairwise compare all columns of one matrix to all columns of the other matrix and will generate the appropriate statistics. There is also a label option which allows labelling of all comparisons by the names (supplied by the user) of the two variables being compared.

characterized the fauna (SPAREA), analyze the data using a variety of multivariate numerical techniques (ORDANA) and analyze sediment data (SEDANA). Taken together, these programs should allow the investigator to devote less time to the manipulation of and more time to the creative analysis of his data.

Acknowledgments

We thank Dr. Joseph L. Simon for reviewing this paper and the computer centers at the University of South Florida and the University of Florida for providing the services and computer time necessary for this project. We especially wish to thank Cameron Bloom for her aid in preparing this manuscript.

Literature Cited

Boesch, D. F. 1973. Classification and community structure of the macrobenthos in the Hampton Roads area, Virginia. Mar. Biol. 21: 226–244.

Cormak, R. M. 1971. A review of classification. J. R. Statist. Soc. 134: 321–367.

Field, J. G. 1971. A computerized cataloguing system for handling biological survey data. S. African Mus. Ass. Bull. 19: 478–493.

Folk, R. L. 1966. A review of grain-size parameters. Sedimentology 6: 73–93.

———, and W. C. Ward. 1957. Brazos River Bar: a study of the significance of grain-size parameters. J. Sed. Petrol. 27: 325–338.

Gower, J. C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53: 325–338.

Hill, M. O. 1973. Reciprocal averaging: an eigenvector method of ordination. J. Ecol. 61: 237–249.

———, R. G. Bunce, and M. W. Shaw. 1975. Indicator species analysis, a divisive polythetic method of classification, and its application to a survey of native pinewoods in Scotland. J. Ecol. 63: 597–613.

Horn, H. 1966. Measurement of "overlap" in comparative ecological studies. Amer. Nat. 100: 419–423.

Poole, R. W. 1974. An introduction to quantitative ecology. McGraw-Hill, New York. 532 pp.

Siegel, S. 1956. Nonparametric statistics for the behavioral sciences. McGraw-Hill, New York. 312 pp.

Sneath, P. H. A., and R. R. Sokal. 1973. Numerical taxonomy. W. H. Freeman and Co., San Francisco. 573 pp.

Sokal, R. R., and F. J. Rohlf. 1969. Biometry. W. H. Freeman and Co., San Francisco. 776 pp.

Stephenson, W. T. 1972. The use of computers in classifying marine bottom communities. Pages 463–473 *in* R. Fraser, ed. Oceanography of the South Pacific. New Zealand Nat. Comm. for UNESCO, Wellington.

Whittaker, R. H. 1973. Ordination and classification of communities. W. Junk b.v. Publishers, The Hague. 737 pp.

Date Accepted: December 2, 1976.

Addresses: *(S.L.S.) Department of Biology, University of South Florida, Tampa, Florida 33620; (S.A.B.) Department of Zoology, University of Florida, Gainesville, Florida 32611; (J.G.F.) C.S.I.R. Oceanographic Research Unit, University of Capetown, Rondebosch, South Africa.*