CABIOS APPLICATIONS NOTES

Vol.6. no.3. 1990
Pages 289 – 290

Some programs useful for managing data in ecology, taxonomy and zoogeography

Bruno Rossaro

This paper describes programs useful in managing a database developed to analyze problems in taxonomy, zoogeography and ecology. While the ecological applications will be discussed in more detail, the same data organization is valid also in taxonomy and zoogeography.

The ecologist often has lists of species collected in sample sites. He or she is interested in the relationships between sites and species, for example to know if some species are preferably distributed in some sites or to obtain a list of the sites where one species is present. Another question is whether some environmental factors are responsible for the observed species distribution. It is useful to have information stored in a structured fashion so that some questions may be asked and some answers given.

Three files have been prepared to satisfy the questions given above: a species file, a samples file and an 'occurrences' file. The species file stores the species names, each associated with a species code; the samples file stores the sampling sites, the sampling dates and environmental variables, such as water temperature, water speed, etc., that are supposed to affect species distribution. As in the species file, in the samples file each sample site is associated with a sample code. In the 'occurrences' file samples and species codes are associated with a number that measures the absolute or relative abundance of a species in a site. If only presence–absence data are available, the occurrence of a species in a site is coded 1. In zoogeography the sampling sites are substituted by geographic regions, whereas in taxonomy the samples file is substituted by the species file and the species file considered in ecology and zoogeography is substituted by the characters file, in taxonomy the 'occurrences' are the character states. In other words, sites are described by species occurrences in ecological problems, zoogeographic regions are described by species presence in zoogeographic problems, whereas species are described by character states in numerical taxonomy.

A library of programs has been developed allowing import–export of data files between different software packages that are currently used by ecologists and taxonomists. As is shown in Figure 1 one can use the same data file to go through the different programs: in this manner one may exploit the advantages of different software products without rewriting the data. An ASCII file '***.INP' is created with a word processor (Wordstar with the non-document option). This file has the sample sites codes followed by the codes of the species that occur in the site, each code followed by the species occurrences. A BASIC program reads this file and creates a 'delimited' file that can be imported in dBASE III Plus. dBASE III Plus is used to check codes and to sort data according to a field (sample, species, character code), using INDEX files. Some dBASE III Plus programs aid in data management. The dBASE III Plus files can be exported to create ASCII files ('***.OUT') that can be used as input to BASIC programs ('can_spe.bas' and 'out_mat.bas') which write input files to multivariate data analysis programs such as NTSYS (Rohlf, 1987), DECORANA (Hill, 1979), CANOCO (Ter Braak, 1989), STATS (Rossaro, 1989), etc. The BASIC program 'out_arc.bas' allows the rebuilding of the '***.INP' file, while the 'ril_sel.bas' program prepares a file with sample sites and environmental variables that become input in 'can_env.bas'; this program builds a file '***.ENV' in the full format required by CANOCO.

The programs are written in MS GW-BASIC (Microsoft Version 3.2) and in dBASE III Plus. The programs have been developed for the Olivetti M-240 personal computer (IBM compatible) and run under the MS-DOS 3.3 operating system or later releases. Program source code is available on diskette (please enclose a 5¼-in. DOS/DD diskette) free of charge on

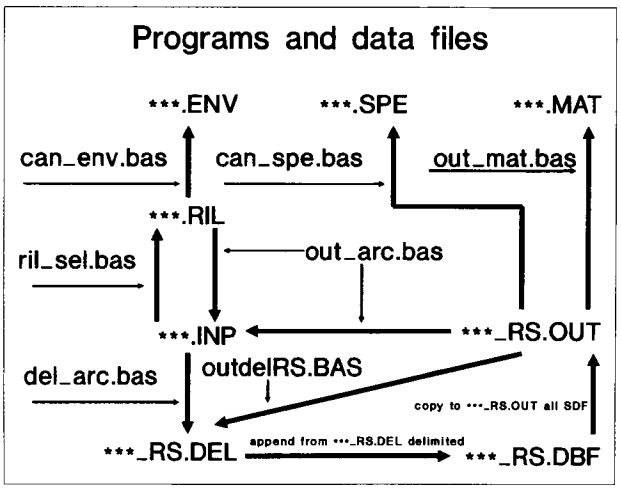

Fig. 1. Flow chart of programs and data files.

Dipartimento di Scienze Ambientali, Università di L'Aquila, Coppito, I 67010 L'Aquila, Italy

written request to the author. No compiled program has been produced at present.

References

Hill,M.O. (1979) *DECORANA: A FORTRAN Program for Detrended Correspondence Analysis and Reciprocal Averaging*. Section of Ecology and Systematics, Cornell University, Ithaca, NY.

Rohlf,F.J. (1987) *NTSYS-pc. Numerical Taxonomy and Multivariate Analysis System*. Applied Biostatics Inc., Exter Publishers, New York.

Rossaro,B. (1989) Procedura STATS: Descrizione di un programma di analisi statistica multivariata per la tassonomia numerica e l'ecologia. *Redia*, **72**, Appendice, 1–40.

Ter Braak,C.J.F. (1989) CANOCO—an extension of DECORANA to analyze species–environment relationships. *Hydrobiologia*, **184**, 169–170.

Circle No. 16 on Reader Enquiry Card