

**SOME MULTIDIMENSIONAL SCALING AND RELATED PROCEDURES**  
**DEvised AT BELL LABORATORIES, WITH ECOLOGICAL APPLICATIONS**

J. Douglas Carroll  
AT&T Bell Laboratories  
Murray Hill, New Jersey 07974, USA

**Abstract** – A large number of multidimensional scaling (MDS) and related models, methods, and computer programs (for all of which we use the generic term “MDS procedures”) have been developed over the years at Bell Laboratories. This paper focuses on probably the most widely known and used subset of Bell Labs MDS procedures involving spatial (as opposed to tree structure, overlapping or non-overlapping clustering, or other “discrete and hybrid”) models. These are: the MDSCAL and KYST family, for two-way (metric or nonmetric) MDS of proximities (e.g., similarities or dissimilarities); INDSCAL, SINDSCAL and IDIOSCAL, for three-way MDS, primarily of proximities (but also applicable to more general multiway data, in a manner to be described); MDPREF, for “internal analysis” of preference (or other “dominance”) data for different individual “subjects” (or other data sources) in terms of a vector model; and the PREFMAP family for “external analysis” of such data (where the “stimulus” or other “object” dimensions are externally provided by prior analysis or theory, only “subject” vectors, ideal points and/or other parameters being determined from preference/dominance data). A number of these Bell Labs MDS procedures are applied to some ecological data on seaworm species due to E. Fresi and collaborators.

## INTRODUCTION

In this paper are presented descriptions of some of the major models, methods, and computer algorithms for multidimensional scaling (MDS) and related techniques developed at Bell Laboratories. Most of the computer programs implementing the procedures described in this paper are available in one of two tapes available at a nominal cost from the AT&T Bell Labs Computer Information Library. These two tapes are referred to as the MDS-1 and MDS-2 tapes. These programs are all written in FORTRAN. Most of those on the MDS-1 tape are written for IBM equipment, while those on the MDS-2 tape should be machine independent. (It should be emphasized that no guarantee is implied that any of these programs will continue to be distributed on this basis by AT&T Bell Laboratories.) All of the programs discussed here, except SINDSCAL, and PREFMAP-3 are on the MDS-1 tape (which has already been very widely distributed). SINDSCAL is on the MDS-2 tape. It is hoped that PREFMAP-3 will soon be available.

While this paper is explicitly limited to procedures for which programs are (or are hoped soon to be) available through the Bell Labs computer information library, a large number of other MDS and related procedures have been developed at Bell Labs which are not so available (and thus are not described here). A supplementary bibliography citing papers relevant to such other procedures developed (totally or in part) at Bell Labs is available by request from the author. Space limitations also require omission of many of the programs included in the Bell Labs package of MDS programs. These include a procedure for maximum likelihood nonmetric 2-way MDS appropriate for proximity data collected by a certain ranking process, called MAXSCAL4.1 (Takane and Carroll 1980, 1981); SIMULES (*SIM*ultaneous *L*inear *E*quation *S*caling) (Carroll and Chang 1972b, Chang and Carroll 1972c); MONANOVA (*MON*otonic *AN*alysis of *VA*riance, which implements a procedure for fitting an additive conjoint measurement model to data from a factorial design) (Kruskal 1965, Kruskal and Carmone 1968); Categorical Conjoint Measurement (CCM) (Carroll 1969, Chang 1971); CANCOR (Generalized *CAN*onical *COR*relation Analysis) (Carroll 1968, Chang 1971); PROFIT (*PRO*perty *FIT*ting) (Carroll and Chang 1964, Chang and Carroll 1968); PARAMAP (*PAR*ametric *MAP*ping of nonlinear data structures) (Carroll 1965, Shepard and Carroll 1966, Chang 1968); POLYFAC (*POLY*nomial *FAC*tor Analysis) (Carroll 1969); HICLUS (*H*ierarchical *CLU*stering via ultrametric tree models) (Johnson 1967), MAPCLUS (A *M*athematical *P*rogramming method for fitting the ADCLUS overlapping *CLU*stering model) (Arabie and Carroll 1980a,b); INDCLUS (*IN*dividual *D*ifferences *CLU*stering) (Carroll and Arabie 1982, 1983); and others. (Most of the programs on the MDS-1 tape, including all of those just named with the exceptions of MAXSCAL4.1, MAPCLUS and INDCLUS, which are on the MDS-2 tape, and MONANOVA, are synopsised and described briefly in Chang, 1971. This paper also includes brief synopses of early versions of MDSCAL, as well as INDSCAL, INDSCALS, NINDSCAL, MDPREF and PREFMAP, all of which are discussed in the body of this paper.) We focus here on two-way and three-way (or Individual Differences) MDS methods for proximity data, and on methods for individual differences preference (or other dominance) data. (For a general discussion of MDS, including many of those models and methods not discussed in detail in the present paper, see Carroll and Arabie 1980.)

The procedures to be discussed here are organized under 3 general headings. These are: I. Two-Way (Nonmetric or metric) Multidimensional Scaling (MDS) procedures; II. Three-Way Multidimensional Scaling (MDS) procedures; III. MDS Analysis of Preference (or other Dominance) Data.

A complete outline of the text of this paper follows, including names of programs and their authors.

- I. Two-Way MDS of Proximity Data (Theoretical references: Shepard 1962a,b; Kruskal 1964a,b).
  - I.A. MDSCAL-5 (Kruskal and Carmone 1969) and KYST, KYST2 and KYST-2A (Kruskal, Young and Seery 1973).
  - I.B. Some ecological data on 88 species of seaworms analyzed by KYST-2A.
- II. Three-Way MDS of Proximity (and other) Data.
  - II.A. INDSCAL (Carroll and Chang 1969, 1970; Chang and Carroll 1969) and SINDSCAL (Pruzansky 1975).
  - II.B. IDIOSCAL (Carroll and Chang 1972a; Chang and Carroll 1972a; Carroll and Wish 1974a).
  - II.C. An application of three-way MDS to the ecological data on seaworms due to Fresi *et al.*
- III. MDS and Multidimensional Analysis of Preference (or other Dominance) Data.
  - III.A. MDPREF (Carroll and Chang 1964; Chang and Carroll 1968).
  - III.B. PREFMAP and PREFMAP-2 (Carroll and Chang 1967; Chang and Carroll 1972b) and PREFMAP-3 (Meulman, Heiser and Carroll 1986).
  - III.C. MDPREF analysis of the Fresi *et al.* seaworm data, and relation to previous analyses via KYST-2A and SINDSCAL.

## I. TWO-WAY MDS OF PROXIMITY DATA

### I.A MDSCAL5, KYST, KYST2 and KYST-2A

The multidimensional scaling (MDS) programs, known as MDSCAL5, KYST, KYST2 and KYST-2A, most closely associated with J. B. Kruskal, are highly versatile in the sense that they can be used for a large variety of scaling problems. These programs have gone through several versions so far, and the discussion below relates to the fifth version of MDSCAL, and to all three versions of KYST. The original theoretical discussion of the procedure can be found in Kruskal (1964a,b). Detailed documentation of these programs can be found in Kruskal and Carmone (1969) or in Kruskal, Young and Seery (1973). A complete discussion of the computational method used is given in Kruskal (1977).

#### *Discussion of Computational Procedure*

The problem attacked by Kruskal (1964a,b), following Shepard's (1962a,b) pioneering work in this area, and known generically as nonmetric (or, with use of certain options, metric) two-way MDS, is that of deriving a configuration of objects in a prespecified number of dimensions, given a set of proximity data among pairs of objects. Let  $\delta_{ij}$  represent the original measure of dissimilarity between pairs of objects  $i$  and  $j$ . Assume for the moment that the dissimilarities can be strictly rank ordered. (The way that ties are handled by the program will be discussed later.) The objective is to represent the  $n$  objects by  $n$  points in an  $r$ -dimensional space, such that the rank order of distances between pairs of points best reproduces the rank order of the  $\delta$ 's. Let the coordinates of the  $i$ -th point in that space be defined by a vector  $\mathbf{x}_i = (x_{i1}, \dots, x_{ir})$ . Let  $d_{ij}$  denote the distance from  $\mathbf{x}_i$  to  $\mathbf{x}_j$ . Let  $\mathbf{X}$  be the matrix whose  $i$ -th row is  $\mathbf{x}_i$ ; thus,  $\mathbf{X} \equiv (x_{it})$ , for  $i = 1, 2 \dots n$  (objects) and  $t = 1, 2 \dots r$  (dimensions).

The criterion is that of minimizing the function called "Stress" given by one of two alternate formulas. The one now known as Formula 2 is:

$$STRESSFORMULA 2 = \sqrt{\frac{\sum_{ij} (d_{ij} - \hat{d}_{ij})^2}{\sum_{ij} (d_{ij} - \bar{d})^2}}, \quad (\text{I.A.1})$$

where  $\bar{d}$  is the average of all the  $d_{ij}$ 's. The one known as Formula 1 is:

$$STRESSFORMULA 1 = \sqrt{\frac{\sum_{ij} (d_{ij} - \hat{d}_{ij})^2}{\sum_{ij} d_{ij}^2}}. \quad (\text{I.A.2})$$

The problem, then, can be expressed as that of finding the matrix  $\mathbf{X}$  such that Euclidean distances, defined as  $d_{ij} = \sqrt{\sum_t^r (x_{it} - x_{jt})^2}$  computed from that matrix, best match the  $\delta_{ij}$ 's. The  $\hat{d}_{ij}$ 's are a set of numerical values chosen to be as close to their  $d_{ij}$  counterparts as possible, subject to being monotone with the original  $\delta_{ij}$ 's. The  $\hat{d}_{ij}$ 's are simply fitted values in the monotone regression procedure.

The two formulas above will be abbreviated here as  $S2$  and  $S1$ , respectively.  $S2$  is, in  $\text{MDSCAL5}$ , the "normal" or default option. In the various versions of  $\text{KYST}$ ,  $S1$  is the default option. It should be mentioned that the two Stress formulas differ only in the normalizing factor in the denominator. In all cases the  $\sum_{ij}$  implies summation over all values of  $i$  and  $j$  for which there are data. For example, if a half-matrix option with diagonal absent is used, the sum would be only over that off-diagonal half-matrix, while if, say, the whole matrix option with diagonal present is used, summation is over all  $n^2$  values. If there are missing cells the summation skips these cells. Furthermore, in the case of  $S2$ ,  $\bar{d}$  is the average over these same values of  $i$  and  $j$ .

The procedure used for obtaining the  $x$ 's is the method of steepest descent. Briefly stated, the method involves improving the starting configuration a bit by moving it slightly in the direction of the negative gradient, or direction of steepest descent. The direction of steepest descent is the direction in the configuration space (the space defined by all  $n \cdot r$  parameters of the  $\mathbf{X}$  matrix) along which stress is decreasing most rapidly. This direction corresponds to the (negative) gradient which is defined by evaluating the partial derivatives of the function  $S$  ( $S1$  or  $S2$ , depending on which option is used).

Letting  $S$  stand for either  $S1$  or  $S2$ , then the gradient will be a vector of  $n \cdot r$  components whose general entry is

$$-\frac{\partial S}{\partial x_{it}} \quad (i = 1, 2, \dots, n; t = 1, 2, \dots, r).$$

The  $n \cdot r$  components of this vector can be "packed" into a matrix  $G$  of the same row and column order as the  $\mathbf{X}$  matrix; thus  $\mathbf{G} \equiv \left[ -\frac{\partial S}{\partial x_{it}} \right]$ . On each iteration a step size  $\alpha$  is defined in a way described in Kruskal's original paper [1964b] and  $\alpha$  times  $\mathbf{G}$  is added to  $\mathbf{X}$  to get an "improved" estimate of  $\mathbf{X}$ . Using a subscript  $I$  for the  $I$ -th iteration, the

iterative process can be described as follows:

Given  $\mathbf{X}_I \equiv (x_{it})$ , the  $I$ -th estimate of  $\mathbf{X}$ ,

1. Compute  $\mathbf{G}_I \equiv \left[ -\frac{\partial S}{\partial x_{it}} \right]$  (evaluated at  $\mathbf{X} = \mathbf{X}_I$ ),
2. Compute  $\alpha_I$  (as described in the above-cited Kruskal paper), and then,
3. Define  $\mathbf{X}_{I+1} = \mathbf{X}_I + \alpha_I \mathbf{G}_I$ .

$\mathbf{X}_{I+1}$  is, then, the improved estimate of the  $\mathbf{X}$  matrix corresponding to iteration  $I + 1$ . This iterative process continues until convergence occurs, as determined by convergence criteria specified in detail in Kruskal and Carmone (1969) or Kruskal, Young and Seery (1973). The  $\mathbf{X}_0$  matrix defines the "initial configuration," (corresponding to the 0-th iteration) which may be defined in a number of different ways. One option is to generate a starting configuration by a procedure that puts all points along the coordinate axes in a systematic way (but one that results in an "essentially random" placement of the points along these axes). This is the one that is sometimes referred to as the "L shaped" starting configuration because, in two dimensions, the configuration does, indeed, resemble an "L."

A second option involves a more fully random configuration ("filling" the space more completely) which can be used by providing a "seed" number for a random number generator. The configuration, in this case, is generated by choosing points randomly from a spherical multivariate normal distribution. By choosing different seeds for the random number generator, of course, different random starts can be used.

A third option is for the user to provide a starting configuration. This may be a "rational" start provided by using some other procedure, an *a priori* configuration of some kind, or one provided by a previous run of the same program which requires additional iterations.

As a fourth option, if one is securing solutions in several dimensionalities in one run, the first  $r$  dimensions of the  $(r + 1)$ st dimensional solution can be used to define the starting configuration for the  $r$ -dimensional solution.

All of the options listed above are available in both MDSCAL-5 and in the various versions of KYST. KYST, KYST2, and KYST-2A have an additional option for the starting configuration, which is probably the most important algorithmic distinction between the KYST and the earlier MDSCAL family of MDS programs. This option entails using an adaptation of the classical metric MDS technique associated with Torgerson (1958) or Gower (1966) to derive a starting configuration. This starting configuration is similar to, but not quite identical with, that in programs by F. W. Young called (generically) TORSCA (Young and Torgerson, 1967, Young, 1968). In the variant of this "TORSCA" starting configuration used in KYST, KYST2 and KYST-2A, a linear transformation of the data is implemented to assure the data values are all positive and that the ratio between the smallest and largest values has a reasonable value. (This provides a practical solution to what is sometimes called the "additive constant" problem in metric MDS methods.)

#### *Special Features of the MDSCAL5 and KYST Programs*

The MDSCAL5 and KYST programs can cope with a variety of problems arising in the original dissimilarities data. We shall discuss them in this section.

*Missing Data* – the program can be set to identify missing observations by reading in a cut-off value below which data will be treated as missing. The Stress function is modified by simply omitting, both in the numerator and denominator, the terms which correspond to the missing cells.

*Nonsymmetry* – either because of inherent nonsymmetry of measurement procedures or errors in measurement, the values of  $\delta_{ij}$  and  $\delta_{ji}$  may not be equal in some cases. In such a situation the Stress function is computed over all cells (i.e.,  $i, j$  and  $j, i$ ) and minimized in the algorithm.

*Ties* – two approaches are possible for resolving ties between dissimilarities (a tie arises wherever  $\delta_{ij} = \delta_{k\ell}$ ). These are called primary and secondary approaches.

In the primary approach, when  $\delta_{ij} = \delta_{k\ell}$  no restriction is placed on the corresponding  $\hat{d}$ 's. Thus, if  $\delta_{ij} = \delta_{k\ell}$ ,  $d_{ij}$  may be greater than, less than, or equal to  $d_{k\ell}$ , without a necessary penalty in the Stress function (since  $\hat{d}_{ij}$  may be greater than, less than, or equal to  $\hat{d}_{k\ell}$ ).

The secondary approach is appropriate when  $\delta_{ij} = \delta_{k\ell}$  is taken to mean that  $\hat{d}_{ij} = \hat{d}_{k\ell}$ . Then if  $d_{ij} \neq d_{k\ell}$ , the terms  $(d_{ij} - \hat{d}_{ij})^2$  and  $(d_{k\ell} - \hat{d}_{k\ell})^2$  cannot both be zero so that Stress might be lowered by making a correction to the configuration, tending to bring  $d_{ij}$  and  $d_{k\ell}$  more nearly into agreement (at least those two components of Stress would be lowered, although, of course, other components may be increased at the same time).

*Non-Euclidean Distances* – the user of the MDSCAL5 or KYST programs can choose any Minkowski- $p$  metric, by specifying the value of  $p$  ( $\geq 1.0$ ), thus causing the program to use the following formula for computing  $d_{ij}$ :

$$d_{ij} = \left[ \sum_{t=1}^r |x_{it} - x_{jt}|^p \right]^{1/p}. \quad (\text{I.A.3})$$

This option enables one to use this specific class of non-Euclidean distances. The Stress and gradient formulas are changed accordingly. (While  $p$  is usually restricted to be  $\geq 1.0$ , values between 0 and 1 can in fact be used, and may be meaningful in some circumstances. If the “ $1/p$ ” power is omitted, this formula does, in fact, yield a metric. For discussion of this, see Carroll and Wish 1974a.)

#### *Definition of Gradient*

It is possible to write a general equation for the (negative) gradient of  $S1$  or  $S2$  for any of the Minkowski- $p$  metrics (recalling that  $p = 2$  corresponds to the Euclidean case). Letting  $S_a$  ( $a = 1$  or  $2$ ) stand for either  $S1$  or  $S2$ , this equation is:

$$g_{it} \equiv \frac{\partial S_a}{\partial x_{it}} = \left[ \frac{2}{S_a \sum_{jk} (d_{jk} - d_a)^2} \right] \sum_k \left[ \frac{d_{ik} - \hat{d}_{ik} - S_a^2 (d_{ik} - d_a)}{d_{ik}^{p-1}} \right] |x_{kt} - x_{it}|^{p-2} (x_{kt} - x_{it}), \quad (\text{I.A.4})$$

where  $p$  is the parameter of the Minkowski  $p$  metric, and where

$$\begin{cases} d_1 = 0 \\ d_2 = \bar{d}, \end{cases} \quad (\text{I.A.5})$$

while  $\hat{d}_{ik}$  is the current “estimate” of  $d_{ik}$  derived from a least squares monotone regression (or from some other least squares regression procedure, options for which have already been described).



Both the definition of Stress and of the gradient are necessarily different for the various “split data” options, which will be described in the section on splitting data below.

### *Options for Regression*

Four basic options exist for performing the regression of  $d_{ij}$  on  $\delta_{ij}$ . These are:

1. Regression-Ascending – for performing monotone regression when the original data are dissimilarities.
2. Regression-Descending – for performing monotone regression when the original data are similarities.
3. Regression-Polynomial – specified integer (degree of polynomial) for performing polynomial regression. If the degree of polynomial is equal to one, it becomes linear regression. An integer from 1 through 4 can be used. In the linear case one has the option of including or excluding a constant term (i.e., the linear function may be non-homogeneous or homogeneous).
4. Regression-Multivariate – integer (number of variates) for performing a prespecified regression by supplying a separate FORTRAN subroutine for same. This option, in principle, allows essentially any linear regression function of the form  $\hat{d} = \sum_{c=1}^C a_c g_c(\delta)$  ( $C \leq 5$ ) to be used, so long as an algorithm is available in the form of the above-mentioned FORTRAN subroutine for computing the functions  $g_c(\delta)$ .

### *Options for Data Input*

The input matrices can be in one or more of these forms:

1. Full, matrix, diagonal present
2. Lower half matrix, diagonal present
3. Lower half matrix, diagonal absent
4. Upper half matrix, diagonal present

5. Upper half matrix, diagonal absent
6. Lower corner matrix\*
7. Upper corner matrix\*

\* A corner matrix is a rectangular ( $MXN$ ) matrix which is treated as an off-diagonal submatrix of a larger  $(M+N) \times (M+N)$  full (square) matrix with the remaining entries handled as missing data.

*Initial Configuration* – the user may supply a starting configuration for scaling the objects. If not, two varieties of a random start can be used, as discussed above. Also, as discussed earlier, other options exist if solutions in more than one dimensionality are obtained. Finally, as discussed earlier, in the KYST programs, the “TORSCA”-like start is another option.

*Splitting Data* – four options exist for using parts of the data as separate sublists and then performing separate regressions for each of these sublists. They are:

1. Split by rows
2. Split by groups
3. Split by decks
4. Split no more (a control phrase used to indicate that no more “split” options are to be specified).

The first three options make each row of every data deck, each group of rows (see Kruskal and Carmone 1969, for explanation of this) or each data deck a separate sublist, respectively. The “split no more” option is relevant only when several data decks are used. It causes all subsequent data decks to be joined into a single sublist until further indication.

In case any of the “split data” options are used, it is necessary to redefine Stress as follows:

$$S^*_a = \sqrt{\frac{1}{N_B} \sum_b S^2_{ab}}, \quad (\text{I.A.6})$$

where  $b$  stands for a data “block” (which may be a row, group, or deck, depending on

options used),  $N_B$  is the number of such blocks, while  $S_{ab}^2$  is  $S_1$  or  $S_2$  (for  $a = 1$  or  $2$ , respectively) defined on block  $b$ .  $S^*_a$  is then the overall Stress (of type  $a$ ), defined simply as the *root mean square* of the individual Stresses.

The gradient can be defined easily. Dropping the “ $a$ ” from  $S$  and  $S^*$ , the overall gradient is simply:

$$G^* = \frac{1}{S^*N_B} \sum_b S_b G_b. \quad (\text{I.A.7})$$

*Data Saving* – it is possible to use the same data for performing different methods of scaling by using the option called “Save Data.”

*Weighting of Data* – the MDSCAL5 and KYST programs allow for differential weighting of the original data values. This can be done either by supplying a matrix of weights in the same way as the data are laid out or by using a FORTRAN subroutine for generating weights internally. The standard weights are taken as 1.0 for each observation. Further details on this and other aspects of these programs can be found in Kruskal and Carmone (1969) or Kruskal, Young and Seery (1973). More information and a general introductory overview of “two-way” multidimensional scaling generally (as well as a brief summary of “three-way” MDS) can be found in Kruskal and Wish (1978).

### **I.B Some Ecological Data on 88 Species of Seaworms Analyzed by KYST-2A**

Some ecological data collected over a period of two years at 5 sites in the harbor of Ischia in the Bay of Naples are described in detail in a later section (II.C) of this paper. Also described in that section is the computation of a number of different proximity (derived dissimilarity) matrices (one for each of the 5 sites, a number for various time periods, and an “overall” dissimilarity matrix).

While leaving details of this measure for later, we will describe briefly here the results of applying KYST-2A to the “overall” dissimilarity measure calculated for the Fresi *et al.* data. Before KYST-2A could be applied to this data, a subset of the seaworm species had to be eliminated. The reason for this is that our version of KYST-2A would handle only 60 objects (in the present case, the species of seaworms). Inspection of the original data in the Fresi *et al.* (1983) paper indicated that 33 species

were observed only twice in the entire study (i.e., at any one of the sites in any one of four time periods). Thus these 33 were eliminated, leaving a total of 55 species to be analyzed by KYST-2A.

**Table 1.** Biological names of 88 seaworm species in data from Fresi *et al.* (1983). These marked with asterisks were the 55 most frequent species in that data, which were analyzed via KYST2-A.

* 1	<i>Lepidonotus clava</i> (Montagu)	* 45	<i>Nereidae</i> gen. sp. 3
2	<i>Pholoe synophthalmica</i> (Claparède)	* 46	<i>Lysidice ninetta</i> Audouin et Milne-Edwards
3	<i>Paleontus debilis</i> (Grube)	* 47	<i>Lumbrineris coccinea</i> (Renieri)
4	<i>Eteone</i> sp.	* 48	<i>Lumbrineris funchalensis</i> (Kinberg)
5	<i>Phyllococe</i> (cfr.) <i>vittata</i> Ehlers	* 49	<i>Lumbrineris inflata</i> (Moore)
6	<i>Eulalia sanguinea</i> Oersted	50	<i>Lumbrineris</i> sp.
* 7	<i>Eulalia viridis</i> Linneo	* 51	<i>Arabella geniculata</i> (Claparède)
8	<i>Hesionidae</i> gen.sp.	* 52	<i>Arabella iricolor</i> (Montagu)
* 9	<i>Syllis</i> (cfr.) <i>vivipara</i> Krohn	53	<i>Dorvillea rudolphii</i> (Delle Chiaje)
* 10	<i>Syllis gracilis</i> Grube	* 54	<i>Polydora ciliata</i> (Johnston)
* 11	<i>Syllis hyalina</i> Grube	* 55	<i>Polydora caeca</i> (Oersted)
* 12	<i>Syllis armillaris</i> (Müller)	* 56	<i>Polydora</i> sp.
* 13	<i>Syllis proliferans</i> Krohn	57	<i>Spio filicornis</i> (Müller)
* 14	<i>Syllis spongicola</i> Grube	* 58	<i>Dodecaceria concharum</i> Oersted
15	<i>Syllis cirropunctata</i> Michel	* 59	<i>Cauleriella bioculatus</i> (Keferstein)
16	<i>Syllis amica</i> Quatrefages	60	<i>Cirratulus cirratus</i> (Müller)
17	<i>Syllis cornuta</i> Rathke	* 61	<i>Cirratulus chrysoderma</i> Claparède
18	<i>Syllis kronii</i> Ehlers	* 62	<i>Cirriiformia filigera</i> (Delle Chiaje)
* 19	<i>Trypanosyllis zebra</i> (Grube)	63	<i>Ctenodrilus serratus</i> (O. Schmildt)
20	<i>Odontosyllis ctenostoma</i> Claparède	* 64	<i>Cirratulidae</i> gen. sp. 1
21	<i>Odontosyllis fulgurans</i> Claparède	* 65	<i>Cirratulidae</i> gen. sp. 2
22	<i>Pionosyllis</i> sp.	66	<i>Theostoma oerstedii</i> (Claparède)
23	<i>Eurysyllis tuberculata</i> Ehlers	67	<i>Capitellidae</i> gen. sp.
* 24	<i>Brania clavata</i> Claparède	* 68	<i>Streblosoma hesslei</i> Day
* 25	<i>Brania pusilla</i> Dujardin	69	<i>Thelepus cincinnatus</i> (Fabricius)
26	<i>Exogone verugera</i> (Claparède)	* 70	<i>Nicolea venustula</i> (Montagu)
* 27	<i>Exogone gemmifera</i> Pagenstecher	* 71	<i>Amphiglena mediterranea</i> (Leydig)
* 28	<i>Exogone</i> sp.	* 72	<i>Potamilla tarelli</i> Malmgren
* 29	<i>Sphaerosyllis hystrix</i> Claparède	* 73	<i>Mixicola aestetica</i> (Claparède)
30	<i>Sphaerosyllis claparedi</i> Ehlers	* 74	<i>Fabricia sabella</i> (Ehrenberg)
* 31	<i>Autolytus aurantiacus</i> Claparède	* 75	<i>Oriopsis</i> (cfr.) <i>eimeri</i> (Langerhans)
32	<i>Autolytus prolifer</i> (Müller)	76	<i>Sabellidae</i> gen. sp.
* 33	<i>Autolytus</i> sp.	* 77	<i>Pileolaria</i> sp.
34	<i>Syllidae</i> gen.	78	<i>Pomatoceros triqueter</i> (Linneo)
* 35	<i>Platynereis dumerilii</i> Audouin et Milne-Edwards	* 79	<i>Hydroides pseudouncinata</i> Zibrowius
* 36	<i>Platynereis coccinea</i> Delle Chiaje	* 80	<i>Hydroides elegans</i> (Haswell)
* 37	<i>Nereis zonata</i> Malmgren	81	<i>Hydroides dianthus</i> (Verrill)
* 38	<i>Nereis persica</i> Fauvel	82	<i>Serpula concharum</i> Langerhans
* 39	<i>Nereis</i> sp.	* 83	<i>Vermiliopsis striaticeps</i> (Grube)
* 40	<i>Ceratonereis costae</i> (Grube)	84	<i>Vermiliopsis</i> sp.
* 41	<i>Perinereis macropus</i> (Claparède)	* 85	<i>Filograna implexa</i> Berkeley
* 42	<i>Perinereis cultrifera</i> Grube	* 86	<i>Spirobranchus polytrema</i> (Philippi)
* 43	<i>Nereidae</i> gen. sp. 1	87	<i>Protula</i> sp.
* 44	<i>Nereidae</i> gen. sp. 2	88	<i>Serpulidae</i> gen. sp.

Table 1 indicates the names of all 88 seaworm species analyzed in this paper. The sequential numerical code on the left is actually used in the various plots in this paper. Asterisks indicate the 55 most frequent species analyzed by KYST-2A. The “regression ascending” option was used, with the primary option for ties, and STRESSFORMULA1. Analyses were done in 6 down to 1 dimension(s).

In MDSCAL or KYST analyses, a plot of STRESS vs. dimensionality is often used as an aid in deciding on the most appropriate dimensionality.

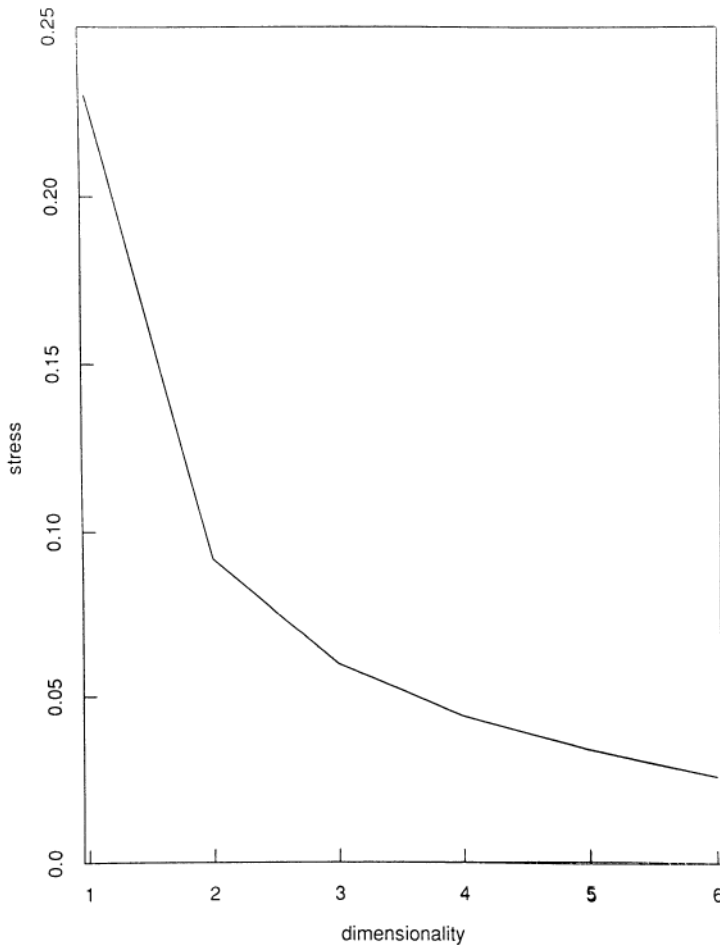


Figure 1. Plot of STRESS (formula 1) vs. Dimensionality for KYST-2A analysis of Fresi *et al.* data.

Figure 1 shows this plot. One often looks for a clear “elbow” in such a plot; that is, a dimensionality after which STRESS falls off only minimally (and more or less linearly) with dimensionality. While inspection of Figure 1 does not yield an absolutely clear

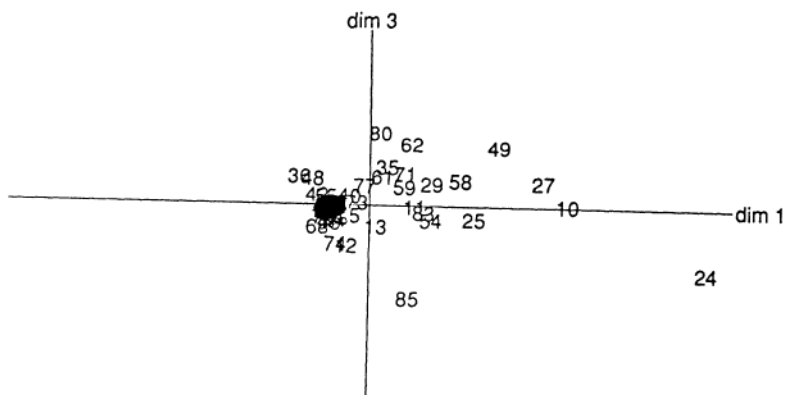


Figure 2. The dimension one-three plane of the four dimensional KYST-2A solution for the 55 most frequent of the 88 seaworm species from the Fresi *et al.* data. "Overall" derived dissimilarity matrix used as input.

"elbow," it was decided that the most appropriate dimensionality was four.

For reasons to be discussed later, the four dimensions were plotted in two planes, the plane defined by dimensions one and three (in Figure 2) and that defined by dimensions two and four (in Figure 3).

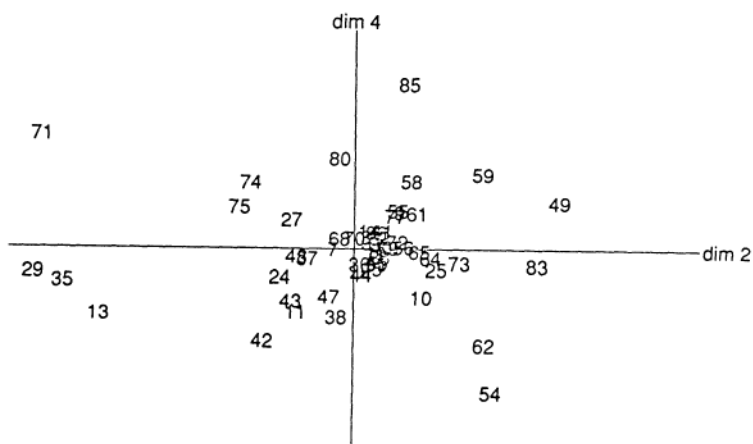


Figure 3. The dimension two-four plane of the four dimensional KYST-2A solution for the 55 most frequent of the 88 seaworm species from the Fresi *et al.* data.

The 55 seaworm species included in this analysis are shown in these figures, using the sequential coding indicated in Table 1. Since the present author is not a biologist, and has no knowledge whatever about these particular species of seaworms, we leave substantive interpretations of these (and other dimension plots to be seen later) to subject matter experts.

## II. THREE-WAY MDS OF PROXIMITY (OR OTHER) DATA

Before a detailed discussion of three-way (and possibly even *higher-way*) MDS or other data analysis models and methods, some terminology is needed. Because of our psychological roots we often speak of “stimuli” and of individual “subjects”. A more neutral pair of terms, however, is “objects” (which can be entities – of any type whatsoever – one is interested in studying; e.g., species of seaworms, variables, sites, times, countries in Western Europe, epistemological theories, numerical ecologists, monads, Hilbert spaces, or brands of soap) and “data sources” (which, as the phrase suggests, comprise any *source of data* about these objects; e.g., individual numerical ecologists, who may make judgments of similarity among various *species* of Polychaetes, different times – say in a longitudinal study – in which measures of correlation over species are computed to provide measures of proximity among various *variables* – with *species* and *variables* comprising “objects” in these two examples, respectively). Clearly, what may be a “stimulus” or other “object” in one context may be an “individual” or other “data source” in another!

We also often speak of the number of *ways* of a data array, as when we refer to two-way, three-way, or higher-way *models* and *methods*, for two-way, three-way or higher-way *data*. The simplest “way” (pardon the ambiguity!) to think of this use of the term “way” is that it is the number of indices, or subscripts, necessary to index the data. The Fresi *et al.* data to be described in detail shortly can be viewed as *four-way* data (species  $\times$  sites  $\times$  months  $\times$  years) since we would need *four* indices to keep track of these four different modalities. If, however, one were to argue (as one well may) that months and years should readily be thought of as a single *mode*, and thus a single *way* of the data array (indexed by only one subscript, ranging systematically – say sequentially in time – over all month-year combinations) then it might as easily be formalized as a *three-way* data array.

Our point here is that the number (and nature) of “ways” in a data array is largely “in the mind of the beholder” (or, more to the point, is dependent on the aims of and/or conceptual structure imposed by the data analyst/researcher trying to understand a particular batch of data).

Another term often used in reference to data arrays, already alluded to tangentially above, is “modes”. A data *mode* is a *type* or *category* of entity (e.g., the “species” mode, “time” mode, “site” mode, or “variable” mode) which may or may not correspond to the “ways” of the data array. In general, the number of “ways” will be at least as great as the number of “modes”, but may be greater (because two or more different “ways” of the data may correspond to the same “mode”). The best example of the latter phenomenon is the case (already considered in Section I) of a two-way, but one-mode  $n \times n$  (usually, but not necessarily symmetric) matrix of proximities (similarity or other proximity measures among seaworm species, for example, or correlation coefficients among variables).

Another case in which the number of “ways” exceeds that of “modes” – which we shall soon encounter – will be a data set that is two-mode (seaworm *species* – the “objects” – by “data sources” derived, as will be described, from data corresponding to various combinations of sites, months and years), but three-way (species  $\times$  species  $\times$  data source). As will be seen in detail shortly, we shall begin with a data array that is either four-mode and four-way, or three-mode and three-way (depending on whether one feels “month” and “year” should be treated as separate modes/ways or as a single mode/way), and *derive* from this another data array that can be conceived as being two-mode, three-way data of proximities among the 88 seaworm species (“objects”) for 14 different “data sources”.

## II.A INDSCAL

The INDSCAL approach, standing for INdividual Differences SCALing of proximity (or other) data by means that retain information on individual differences, was developed by Carroll and Chang (1970). Two basic options exist in the INDSCAL program: a) INDSCAL analysis *per se* (called INDIFF in the program) – for scaling stimuli (or other objects) for which symmetric matrices of proximity measures are available for a number of individuals or other data sources, in terms of a *weighted*



Euclidean model often called the INDSCAL *model* and b) CANDECOMP analysis – for scaling stimuli (or other objects) for which (for example) measurements are available on a number of variables (i.e., the input matrices are, in general, rectangular and non-symmetric) in a number of different *conditions* (e.g., observational contexts, experimental variations, times, sites, or other “modes” or scenarios distinguishing the various object  $\times$  variable matrices). The CANDECOMP part of the algorithm uses Carroll and Chang’s method of canonical decomposition of  $N$ -way tables. INDSCAL analysis (option a) actually corresponds to using *symmetric* CANDECOMP with pre- and post-processing, to be described below.

### *The INDSCAL Model*

The INDSCAL model of individual differences is based on two major assumptions which are stated below. While (since this model was originally devised in a psychological context) we will refer to “stimuli” and “individuals,” and to individual differences in perception as reflected in similarity judgments, this model can be applied to individual differences, among any type of data sources, in similarity or dissimilarity measures defined on all pairs of objects (e.g., species), from domains other than that of psychological “stimuli.”

*Assumption 1* – a set of  $r$  dimensions or “factors” underlie the  $n$  stimuli. These dimensions are assumed to be common for all  $m$  individuals making similarity judgments, i.e., they are sufficient to account for (except for “noise” or error) the similarity judgments (or other proximity values) associated with all  $m$  subjects, or other data sources. Let  $\mathbf{X} \equiv (x_{jt})$  represent the matrix of stimulus coordinates in the common, or group space;  $x_{jt}$  is the coordinate value for the  $j$ -th stimulus on the  $t$ -th dimension;  $j = 1, 2, \dots, n$  and  $t = 1, 2, \dots, r$ .

*Assumption 2* – the similarity judgments for each individual are related in a simple way to a “modified” Euclidean distance in the group stimulus space. In particular, the relationship is assumed to be linear (in the metric version) or monotone (in a quasi-nonmetric version). We shall describe the metric version which is the one used predominantly. (The quasi-nonmetric version is implemented in a program called NINDSCAL, available on the MDS-1 tape, but this will not be discussed further here.) We assume that the dissimilarity measure,  $\delta_{jk}^{(i)}$ , provided by the  $i$ -th individual for the

pair of stimuli  $j$  and  $k$ , is related to a modified or weighted Euclidean distance,  $d_{jk}^{(i)}$ , by:

$$L^{(i)} \left( \delta_{jk}^{(i)} \right) \cong d_{jk}^{(i)} \quad (\text{II.A.1})$$

where  $L^{(i)}$  is a linear function with positive slope. The subscripts  $j$  and  $k$  (for stimuli or other objects) range from  $1, 2, \dots, n$  and the superscript  $i$  (for individuals or other data sources) ranges from  $1, 2, \dots, m$ .

The “modified” Euclidean distance for the  $i$ -th subject is given by the formula:

$$d_{jk}^{(i)} = \sqrt{\sum_{t=1}^r w_{it} (x_{jt} - x_{kt})^2}. \quad (\text{II.A.2})$$

This formula differs from the usual Euclidean distance formula only in the presence of the weights  $w_{it}$ , which represent the saliences or “perceptual importances” for the  $i$ -th individual of the  $t$ -th dimension of the group perceptual space, represented by the matrix  $\mathbf{X}$ . Another way to express the  $d_{jk}^{(i)}$ 's are as ordinary Euclidean distances computed in a “private” space for individual  $i$  whose coordinates are:

$$y_{jt}^{(i)} = w_{it}^{1/2} x_{jt}. \quad (\text{II.A.3})$$

This is a space that is like the  $\mathbf{X}$ -space except that the configuration has been expanded or contracted (differentially) in directions corresponding to the coordinate axes. This can be seen to be a linear transformation with the transformation matrix restricted to be diagonal (the diagonals being square roots of the  $w$ 's). This class of transformations is sometimes referred to as a “strain.”

The above model is sufficiently general to accommodate individuals with widely divergent perceptions of a set of stimuli in terms of a common perceptual space. For example, consider a two-dimensional perceptual space of a set of automobile brands whose axes are identified as luxuriousness and sportiness. Let us now imagine two individuals,  $P$  and  $Q$ , who view the brands entirely differently, individual  $P$  considering the brands only on one dimension (luxuriousness) while individual  $Q$  views the brands only on the other (sportiness). The INDSCAL model has the capability of accommodating the judgments of both persons  $P$  and  $Q$ , by allowing dimension weights of  $(1, 0)$  for  $P$  and  $(0, 1)$  for  $Q$ .

The same basic model, but without a method for fitting the model to data, was proposed independently by Horan (1969). Alternative methods of fitting the INDSCAL

model (sometimes called simply the “weighted Euclidean model”) to data have been proposed by Bloxom (1968), Takane *et al.* (1977) and Ramsay (1977).

### *Estimation of Parameters*

We now briefly discuss the procedures by which the parameters of the model, namely, the  $n \times r$  elements of the  $\mathbf{X}$ -matrix and the  $m \times r$  elements of the matrix  $\mathbf{W} \equiv (w_{it})$  are estimated from dissimilarity judgments on all possible  $n(n-1)/2$  distinct pairs of stimuli by  $m$  individuals.

The first step in the method of estimation is to convert the dissimilarities into distance estimates. In view of the linearity assumptions made above, this is done using the standard procedure described in Torgerson (1958). This method entails estimation of an additive constant which converts the comparative distances (i.e., the original dissimilarity judgments) into absolute distances between pairs of stimuli. The method estimates the smallest value of the constant which guarantees satisfaction of the triangle inequality for all triples of points. This can easily be shown to be  $c_{\min}^{(i)} = \max_{jkl} \left[ \delta_{j\ell}^{(i)} - \delta_{jk}^{(i)} - \delta_{k\ell}^{(i)} \right]$ . This constant guarantees that the triangle inequality will be satisfied for all triples of points, with the inequality being a precise equality for at least one triple (the one for which the expression above attains its maximum). It is as though these three points lie precisely on a straight line in the multidimensional space. This is why this scheme is sometimes called the “one-dimensional subspace” method of estimating the additive constant. Any constant larger than  $c_{\min}^{(i)}$  would certainly suffice also, but  $c_{\min}^{(i)}$  is, as its name implies, the smallest constant guaranteeing this. While there are a number of other schemes of estimating the so-called additive constant (see Torgerson 1958), this one is one of the simplest (both conceptually and numerically) and most assumption-free. Having estimated  $c^{(i)}$  in this way, distance estimates,  $\hat{d}_{jk}^{(i)}$ , are calculated as  $\hat{d}_{jk}^{(i)} = \delta_{jk}^{(i)} + c^{(i)}$ .

The distance estimates are then converted for each subject to scalar products between the points represented as vectors issuing from an origin at the centroid of all the points. This is done by double centering the matrix whose entries are  $-1/2 \left[ \hat{d}_{jk}^{(i)} \right]^2$ . The resulting numbers,  $\hat{b}_{jk}^{(i)}$ , can be regarded as the estimated scalar products between the vectors  $\mathbf{y}_j^{(i)} \equiv (y_{j1}^{(i)}, y_{j2}^{(i)}, \dots, y_{jr}^{(i)})$ , and  $\mathbf{y}_k^{(i)}$ . This step is the same as in the “metric”

phase of the TORSCA (Young, 1968) algorithm, and in generating the "TORSCA" starting configuration in KYST, KYST2 and KYST-2A (Kruskal, Young and Seery, 1973).

The derivation below shows that these numbers are, in fact, estimated scalar products. (Readers not interested in this derivation are advised to skip to the section entitled "Scalar Product Form of INDSCAL Model.")

### *Derivation of Scalar Products from Distances*

Given exact squared Euclidean distances

$$d_{jk}^2 = \sum_{t=1}^r (x_{jt} - x_{kt})^2 \quad (\text{II.A.4})$$

assume:

$$\sum_{j=1}^n x_{jt} = 0 \text{ for all } t = 1, 2, \dots, r \quad (\text{II.A.5})$$

(We may do this without loss of generality, since the origin of the  $x$  space is arbitrary, and this just fixes it at the centroid of all  $n$  points.) Expanding (II.A.4),

$$\begin{aligned} d_{jk}^2 &= \sum_t (x_{jt}^2 - 2x_{jt}x_{kt} + x_{kt}^2) \quad (\text{II.A.6}) \\ &= \sum_t x_{jt}^2 - 2\sum_t x_{jt}x_{kt} + \sum_t x_{kt}^2 \\ &= \ell_j^2 + \ell_k^2 - 2b_{jk} \end{aligned}$$

where

$$\ell_j^2 = \sum_t x_{jt}^2 \quad (\text{II.A.7})$$

and

$$b_{jk} = \sum_t x_{jt}x_{kt} \text{ (the scalar product)} \quad (\text{II.A.8})$$

Because of II.A.5

$$b_{.k} = b_{j.} = b_{..} = 0 \quad (\text{II.A.9})$$

(e.g.,  $b_{.k} = \frac{1}{n} \sum_j b_{jk} = \frac{1}{n} \sum_j \sum_t x_{jt}x_{kt} = \frac{1}{n} \sum_t x_{kt} \sum_j x_{jt} = 0$ ).

From (II.A.6) and (II.A.9) we have

$$d_{\cdot k}^2 = \ell_{\cdot}^2 + \ell_k^2 \quad (\text{II.A.10})$$

$$d_{j \cdot}^2 = \ell_j^2 + \ell_{\cdot}^2 \quad (\text{II.A.11})$$

$$d_{\cdot \cdot}^2 = 2\ell_{\cdot}^2 \quad (\text{II.A.12})$$

where

$$\ell_{\cdot}^2 = \frac{1}{n} \sum_j \ell_j^2 \quad (\text{II.A.13})$$

Then (II.A.6), (II.A.10), (II.A.11) and (II.A.12) together imply that

$$d_{ij}^2 - d_{\cdot k}^2 - d_{j \cdot}^2 + d_{\cdot \cdot}^2 = -2b_{jk} \quad (\text{II.A.14})$$

Multiplying both sides by  $-1/2$  gives the desired result.

Note that we didn't have to know anything about geometry to derive this result. The law of cosines, for example, was never mentioned.

Note also that this is an exact result for deriving exact scalar products (about an origin at the centroid) from exact Euclidean distances. In practice, of course, we derive *estimated* scalar products ( $\hat{b}$ 's) from estimated distances ( $\hat{d}$ 's).

### *Scalar Product Form of INDSCAL Model*

A *scalar product* form of the INDSCAL model can be devised by substituting vectors  $\mathbf{y}_j^{(i)}$  and  $\mathbf{y}_k^{(i)}$  in the "private space" for individual  $i$  into the definition of the scalar product, as follows:

$$b_{jk}^{(i)} \equiv \sum_{t=1}^r y_{jt}^{(i)} y_{kt}^{(i)} = \sum_{t=1}^r w_{it} x_{jt} x_{kt} \quad (\text{II.A.15})$$

Thus, the three-way matrix of individuals by stimulus pairs, where general entries are the values of  $b_{jk}^{(i)}$  derived from the dissimilarity data, can, if the INDSCAL model holds, be decomposed into the *trilinear* form in equation (II.A.15). The problem now is one of estimating values of the  $\mathbf{X}$ -matrix and the  $\mathbf{W}$ -matrix where elements enter into the right-hand side of equation (II.A.15). This estimation (in a least squares sense) can be achieved by a procedure called "canonical decomposition of  $N$ -way tables" (now usually abbreviated CANDECOMP). In this particular case,  $N = 3$ , since there are three ways, two for stimuli and one for individuals. The CANDECOMP procedure, for the general

$N$ -way case ( $N \geq 3$ ) is described in detail in Carroll and Chang (1970).

CANDECOMP is actually designed to analyze data in terms of a more general *trilinear* model (or multilinear, in the  $N$ -way case for  $N > 3$ ) which (in the 3-way case) is of the form:

$$z_{ijk} \cong \sum_{t=1}^r a_{it} b_{jt} c_{kt} \quad (\text{II.A.16})$$

where  $z_{ijk}$  represents data, the  $a$ 's,  $b$ 's and  $c$ 's are parameters to be estimated and " $\cong$ " here implies least squares estimation. The CANDECOMP procedure provides least squares estimates of these parameters (the  $a$ 's,  $b$ 's and  $c$ 's) via what is now called an *Alternating Least Squares* procedure but was originally called a NILES (*Nonlinear Iterative Least Squares*) or NIPALS (*Nonlinear Iterative Partial Least Squares*) procedure (see Carroll and Chang, 1970).

The INDSCAL special case is obtained by making the following identifications:

$$\begin{cases} z_{ijk} = b_{jk}^{(i)} \\ a_{it} = w_{it} \\ b_{jt} = c_{jt} = x_{jt} \end{cases} \quad (\text{II.A.17})$$

However, when CANDECOMP is applied to a 3-way table of scalar products (which are symmetric in the  $j, k$  indices), no special constraint is imposed to make "matrix 2" [ $B \equiv (b_{jt})$ ] equal to "matrix 3" [ $C \equiv (c_{kt})$ ]. This will be true (up to the class of admissible transformations, namely that of a "strain," or linear transformation given by a diagonal matrix) when the iterative process has converged. That it be exactly true is guaranteed by setting them equal in a final stage of the program. For a general discussion of CANDECOMP, including theory and applications of its higher (than three) way generalizations, and an extension called CANDELINC enabling linear constraints on parameters, see Carroll and Pruzansky (1984). Carroll, Pruzansky and Kruskal (1980) provide a general discussion of CANDELINC.

#### *Normalization of Data and Solution*

In the algorithm for the INDIFF part of the program (the part that does the actual INDSCAL analysis using CANDECOMP as a subroutine), the original data and final solutions are normalized. In the case of the original data, the scalar product matrices

are normalized such that the sum of squares of the scalar product matrix is set equal to unity for each subject (or “data source”). In the case of INDSCAL analysis, the final stimulus space is normalized such that the variance of projections of stimuli on the coordinate axes is equal to unity and the centroid is at the origin. The appropriate companion normalization is applied to the subject-matrix.

The combination of these two different procedures has one interesting outcome: the square of the Euclidean distance of a subject’s point from the origin can be (approximately) interpreted as the (proportion of) total variance accounted for in the scalar products data for that subject. If the dimensions of the stimulus space are orthogonal, then the square of the Euclidean distance of the subject’s point will *exactly* equal the proportion of variance accounted for. No normalization of the data is done for the CANDECOMP option; there is, however, a normalization of the solution. Specifically, all matrices but the first are normalized to have unit sums of squares for each dimension. All the differences in sums of squares are then absorbed in the final matrix. When using CANDECOMP the origins of the various spaces are not constrained at all.

### *Input Parameters*

The various input parameters of the INDSCAL program are enumerated below:

*Data Input Options* – these are controlled by a parameter called IRDATA. Eight alternatives are provided in the program, corresponding to integer values of 0 to 7 for IRDATA.

IRDATA	Input Option
0	Rectangular matrices (this is the CANDECOMP option)
1	Lower half of similarities matrix without diagonal
2	Lower half of dissimilarities matrix without diagonal
3	Lower half of Euclidean distance matrix without diagonal
4	Lower half of correlation matrix without diagonal
5	Lower half of covariance matrix with diagonal
6	Full symmetric matrix of similarities
7	Full symmetric matrix of dissimilarities

In cases 1-5 the matrix can also be read in as an ordered vector.

The user can obtain either a simultaneous or a successive  $r$ -dimensional solution. In the simultaneous case all dimensions in the matrices are computed at one time, whereas in the successive case, as the name indicates, only one dimension is estimated at a time. In general, unless there is good reason to do otherwise, a "simultaneous" solution should be obtained. The user can control stringency of convergence of the iterative process by two parameters, namely, maximum number of iterations and another specifying a convergence criterion based on changes in the fit measure from iteration to iteration.

An option exists for not setting matrix 2 equal to matrix 3. In the general CANDECOMP analysis this option must always be chosen since, in general, the input matrices are different. In the case of INDSCAL analysis, however, matrix 2 is set equal to matrix 3 since, by symmetry, these input matrices should be equal. When done in the latter fashion, we refer to the CANDECOMP analysis (say of the derived scalar products) as *symmetric* CANDECOMP.

The INDSCAL program can also be used in solving for the weights assigned by subjects to a prespecified configuration. The program also has the ability to use a prespecified configuration as a rational start even in the case in which all matrices are to be solved for.

More complete details of how to use the INDSCAL program can be found in Chang and Carroll (1969).

### *SINDSCAL*

SINDSCAL (Pruzansky, 1975) is another computer program that implements the procedure of Carroll and Chang (1970) for fitting the INDSCAL individual differences model for multidimensional scaling. It is a modification of the more general INDSCAL program of Chang and Carroll (1969) described above. (See also INDSCALS described in Chang (1971).) INDSCAL was written to allow as input either rectangular or symmetric matrices of proximities. Since almost all of the applications of INDSCAL to date have used similarities or dissimilarities, Euclidean distances, correlations or covariances, SINDSCAL was written to handle only these symmetric data. It is also limited to the three-way case.



The method of analysis used in SINDSCAL is essentially the same as the method of Carroll and Chang (1970) used in INDSCALS (Chang 1971). Therefore, the final stimulus and weights configurations should be identical (except for possible differences due to different convergence criteria, starting configurations, or other numerical details). The principal differences between SINDSCAL, INDSCALS or INDSCAL used with three-way "INDIFF" options lie in the computational procedure and user options.

Modifications in the internal program structure have yielded:

- (1) a considerable reduction in memory requirements achieved by storing the input data in symmetric storage mode. (An  $n \times n$  stimulus matrix in symmetric storage mode is reduced to a vector of length  $n(n + 1)/2$ .)
- (2) considerable simplification of the computational algorithm in the main computation subroutine, called CANDE.
- (3) some reduction in computation at various stages in the procedure.

These changes along with the use of the global optimization feature of the Fortran-IV compiler result in significant savings in computer charges. Additional savings may result because SINDSCAL uses dynamic storage allocation. Small data sets may be run with proportionately smaller computer memory and, therefore, some savings in cost.

Some user-oriented changes include:

- (1) a reduction in the number of input parameters required,
- (2) additional plotting and printing options,
- (3) provision for a user-supplied subroutine to preprocess the input data,
- (4) sufficient printout throughout the computation so that most of the information from a run can be recovered if the program gets cut before completion,
- (5) no limitation on the number or size of the input matrices due to the use of dynamic storage allocation,
- (6) access to the program in both batch and time-sharing modes.

Since most features of SINDSCAL have already been described in the discussion of INDSCAL above, we highlight only those features in which it is most distinct from that earlier program/procedure.

### *Input Data Type*

The input to SINDSCAL consists of many different matrices, corresponding to different individuals or other “data sources”, but all pertaining to the same stimuli or other “objects”. Thus SINDSCAL deals *only* with two-mode but three-way data! Since each matrix is assumed, in SINDSCAL, to be symmetric, only a *half* matrix need be provided (and stored) in each case. These matrices may be:

- (1) similarities, dissimilarities, Euclidean distances or correlations, represented as lower-half matrices without diagonals;
- (2) covariances or scalar product matrices in the form of lower-half matrices with diagonals;
- (3) full symmetric matrices of similarities or dissimilarities. The program ignores the values on the diagonal. In this case, although the upper half of each matrix is (redundantly) provided as input, only a half matrix is stored, thus allowing the greater efficiency in memory storage and computation which is the principal hallmark of SINDSCAL.

### *Maximum Number of Iterations*

SINDSCAL uses the same basic iterative procedure as is used in INDSCAL to estimate parameters. The program ends when convergence is achieved, or the maximum number of iterations has been reached; the reason for ending the analysis is printed on the standard output. The convergence criterion is based on the difference between the fit on the current iteration and the previous iteration. When this difference is less than a certain value, the process is considered to have converged. An important advantage of the INDSCAL model, as already discussed, is that the orientation of coordinate axes is uniquely determined. However, the solution must have reached a global minimum for the axes to be in the correct orientation. Since it is relatively inexpensive to run SINDSCAL (as compared to running the INDSCAL program using the “INDIFF” options), it is recommended that, if possible, a very large number, such as 200, be used

for this option in order to prevent the program from stopping before convergence has been reached.

### *Plot Options*

The program generates plots of all possible planes (defined by pairs of SINDSCAL coordinates) of the final *group stimulus space* and *weights space*. The points may be numbered or the user may supply either the stimulus or subject labels or both sets of labels. It is also possible to suppress all plotting.

### *Relaxation Factor*

A “relaxation” factor was introduced in the parameter estimation procedure (subroutine CANDE). This technique was originally described by Harshman (1970). Its effect is to move the parameters being estimated in a direction beyond the value which is optimum for the current iteration and, hopefully, towards the final overall optimum value. In practice, the number of iterations generally is reduced by at least one-half, and the final solutions are identical to solutions obtained without the relaxation factor.

For a description of preprocessing and normalization options available for certain data types, output options, and other details of SINDSCAL, see Pruzansky (1975), or Arabie, Carroll and DeSarbo (in press).

## **II.B IDIOSCAL**

IDIOSCAL (*Individual Differences In Orientation SCALing*), is a generalization of INDSCAL allowing *IDIO*syncratic reference systems as well as an analytic approximation to INDSCAL. Equations originally formulated in the 1970 Carroll-Chang *Psychometrika* paper describing INDSCAL and CANDECOMP (the method of canonical decomposition of *N*-way tables on which INDSCAL is based) have been implemented in a computer program called IDIOSCAL. This amounts to a generalization of INDSCAL in which each individual (or “data source”) is allowed an idiosyncratic orthogonal rotation of the coordinate system prior to differential weighting of this (rotated) reference system. A classical example of (conceptual) rotation of such a coordinate system is provided by the debate in the early part of this century among educational psychologists about the *factors* or *dimensions* underlying intelligence. (To

simplify matters, let us suppose for now that all agreed that there were exactly *two* dimensions of intelligence.) One school proposed a first (primary) dimension (often called “G”) corresponding to “General Intelligence”, with a second (and secondary) dimension contrasting verbal with quantitative ability. A second school countered that – quite to the contrary (they felt) – there were two *independent*, sovereign and equally theoretically valid dimensions – one a dimension of *verbal* and a second of *quantitative* intelligence! From the perspective of our modern sophisticated multivariate point of view, replete with manifold degrees of rotational freedom, we see quite clearly that these two schools were arguing, quite vociferously as it happens, about nothing more than *different rotations* of coordinate systems describing the *same space* of intellectual “objects” (e.g., specific “abilities” measured by equally specific tests; or, in a dual manner, specific individuals exemplifying different degrees of these abilities, as measured by their respective “factor scores”). To derive the IDIOSCAL model as a description of the perceptual structure of intelligence for these different educational psychologists, we need only add the assumption that, within each of these “schools” different adherents attached different saliences, or “perceptual importances”, to the two dimensions characterizing the particular “school” to which that particular scholar subscribed. In practice, the IDIOSCAL model means that each individual is allowed a generalized Euclidean metric defined by a positive definite quadratic form. Another (seemingly different, but mathematically equivalent) interpretation of this quadratic form is possible in terms of different “subjective intercorrelations” of the same set of coordinate axes. This latter interpretation is favored by Tucker, Harshman and others.

The model includes as special cases Tucker’s (1972) Three-Mode Scaling, based on three-mode factor analysis, the PARAFAC-2 model and method of R. Harshman (1972), and a generalization of INDSCAL proposed by Sands and Young (1980). The method of solution is closely related to one proposed by P. Schönemann (1972), based on earlier work of Meredith’s (1964).

Inspired by Schönemann’s (1972) “Analytic Solution” for the INDSCAL model (which provides an exact solution in the errorless case, but has uncertain properties with errorful data), we have incorporated a second phase that allows an analytic approximation to INDSCAL based on a modification of Schönemann’s procedure. Basically, it differs in that, rather than choosing some arbitrary particular subject to

define a rotation of axes, we define a kind of composite (different from the arithmetic average) of the actual subjects, which is used to determine a more nearly optimal orientation. This seems to work well in cases of both real and artificial (errorful) data.

A third "phase" of the IDIOSCAL program assumes no individual differences whatever, forcing all individuals to have the same axis orientation and weights (except for a possible overall scale factor). This is tantamount to a scaling of the averaged data (but averaged in the more appropriate way outlined by Horan, 1969).

Thus the three phases of IDIOSCAL are very closely analogous to the first three phases of PREFMAP (for *PREF*erence *MAP*ping of stimulus spaces) which will be discussed at a later point in this paper. To carry this analogy further, approximate  $F$  tests have been incorporated, as in PREFMAP, which may be useful for distinguishing between models, and may even help in judging dimensionality.

We now describe, in fairly cursory mathematical notation, the hierarchy of models (and related "phases" of analysis) involved in IDIOSCAL. To those familiar with PREFMAP (See III.B for a description) this hierarchy can be seen to be closely analogous to the hierarchy of decreasingly general "unfolding" models in that approach to external unfolding analysis.

In each case we begin description of the relevant model, assuming we have already (via assumption and/or appropriate preprocessing) obtained data values we believe to be approximate squared Euclidean distances between stimuli (objects),  $j$  and  $k$ , for each subject (data source)  $i$ , which we shall call  $\left[\delta_{jk}^{(i)}\right]^2$ , and state the model assumption for these values.

*"Phase I" of IDIOSCAL – The general model.*

$$\left[\delta_{jk}^{(i)}\right]^2 = \left[d_{jk}^{(i)}\right]^2 = (\mathbf{y}_j^{(i)} - \mathbf{y}_k^{(i)})(\mathbf{y}_j^{(i)} - \mathbf{y}_k^{(i)})', \quad (\text{II.B.1})$$

where

$$\mathbf{y}_j^{(i)} = \mathbf{x}_j \mathbf{T}_i, \quad (\text{II.B.2})$$

so that

$$\left[d_{jk}^{(i)}\right]^2 = (\mathbf{x}_j - \mathbf{x}_k) \mathbf{T}_i \mathbf{T}_i' (\mathbf{x}_j - \mathbf{x}_k)' = (\mathbf{x}_j - \mathbf{x}_k) \mathbf{R}_i (\mathbf{x}_j - \mathbf{x}_k)', \quad (\text{II.B.3})$$

where

$$\mathbf{R}_i = \mathbf{T}_i \mathbf{T}_i' \equiv \frac{1}{m} \sum_{j=1}^m \left[ \delta_{jk}^{(i)} \right]^2 \equiv \left[ \delta_{jk}^{(\cdot)} \right]^2. \quad (\text{II.B.4})$$

Defining  $\left[ \delta_{jk}^{(\cdot)} \right]^2 \equiv \frac{1}{m} \sum_{i=1}^m \left[ \delta_{jk}^{(i)} \right]^2$ , we have

$$\left[ \delta_{jk}^{(\cdot)} \right]^2 \equiv \left[ d_{jk}^{(\cdot)} \right]^2 = \frac{1}{m} \sum_{i=1}^m \left[ d_{jk}^{(i)} \right]^2 = (\mathbf{x}_j - \mathbf{x}_k) \mathbf{R} (\mathbf{x}_j - \mathbf{x}_k)', \quad (\text{II.B.5})$$

with

$$\mathbf{R} = \frac{1}{m} \sum_{i=1}^m \mathbf{R}_i. \quad (\text{II.B.6})$$

Without loss of generality, we may assume

$$\mathbf{R} = \mathbf{I}, \quad (\text{II.B.7})$$

so that

$$\left[ \delta_{jk}^{(\cdot)} \right]^2 \equiv \left[ d_{jk}^{(\cdot)} \right]^2 = (\mathbf{x}_j - \mathbf{x}_k) (\mathbf{x}_j - \mathbf{x}_k)'. \quad (\text{II.B.8})$$

That is  $\left[ \delta_{jk}^{(\cdot)} \right]^2$  is approximately an ordinary squared Euclidean distance defined in terms of coordinates  $\mathbf{x}$ . This fact allows us to obtain an approximate solution for  $\mathbf{X} \equiv (x_{jt})$ , the matrix of  $x$  coordinates, using the ‘‘classical’’ metric MDS approach (see II.A for details). Writing Eq. (II.B.3) in summational notation, we have

$$\begin{aligned} \left[ \delta_{jk}^{(i)} \right]^2 &\equiv \left[ d_{jk}^{(i)} \right]^2 = \sum_t^r \sum_{t'}^r (x_{jt} - x_{kt}) r_{tt'}^{(i)} (x_{jt'} - x_{kt'}) \\ &= \sum_{(tt')}^{r(r+1)/2} r_{(tt')}^{*(i)} \Delta_{(jk)(tt')}, \end{aligned} \quad (\text{II.B.9})$$

where

$$r_{(tt')}^{*(i)} = (2 - \delta_{tt'}) r_{tt'}^{(i)} \quad (\text{II.B.10})$$

and

$$\Delta_{(jk)(tt')} = (x_{jt} - x_{kt}) (x_{jt'} - x_{kt'}) \quad (\text{II.B.11})$$

(while  $\delta_{tt'}$  is the ‘‘Kronecker delta,’’  $\delta_{tt'} = 1$  if  $t = t'$ , 0 otherwise). Let

$$\mathbf{r}^*_i = \left[ r^*_{(11)}^{(i)}, r^*_{(12)}^{(i)}, r^*_{(13)}^{(i)}, \dots, r^*_{rr}^{(i)} \right]. \quad (\text{II.B.12})$$

So  $\mathbf{r}^*_i$  is a row vector of  $\binom{r+1}{2}$  components and

$$\Delta \equiv (\Delta_{(jk)(tt')}), \quad (\text{II.B.13})$$

an  $\binom{n}{2} \times \binom{r+1}{2}$  matrix, while

$$\mathbf{d}_i^{[2]} = \left[ \left[ d_{12}^{(i)} \right]^2, \left[ d_{13}^{(i)} \right]^2, \left[ d_{23}^{(i)} \right]^2, \dots, \left[ d_{(n-1)n}^{(i)} \right]^2 \right]'. \quad (\text{II.B.14})$$

So,  $\mathbf{d}_i^{[2]}$  is a column vector of  $\binom{n}{2}$  components. Eq. (II.B.9) can be written in matrix form as:

$$\mathbf{d}_i^{[2]} = \Delta \mathbf{r}^*_i, \quad (\text{II.B.15})$$

where  $\mathbf{d}_i^{[2]}$  and  $\Delta$  are known, and  $\mathbf{r}^*_i$  is to be solved for. The least squares solution is

$$\hat{\mathbf{r}}^*_i = \mathbf{d}_i^{[2]'} \Delta (\Delta' \Delta)^{-1}. \quad (\text{II.B.16})$$

Having solved for  $\hat{\mathbf{r}}^*_i$ , the entries can be “unpacked” in the appropriate way into  $\hat{\mathbf{R}}_i$  (a square symmetric matrix).  $\hat{\mathbf{R}}_i$  can then be factored into  $\hat{\mathbf{T}}_i \hat{\mathbf{T}}_i'$ . One way is to factor

$$\hat{\mathbf{R}}_i = \mathbf{U}_i \boldsymbol{\beta}_i^2 \mathbf{U}_i', \quad (\text{II.B.17})$$

with  $\mathbf{U}_i$  orthogonal and  $\boldsymbol{\beta}_i^2$  diagonal, and define

$$\hat{\mathbf{T}}_i = \mathbf{U}_i \boldsymbol{\beta}_i. \quad (\text{II.B.18})$$

$\mathbf{U}_i$  can be interpreted as an orthogonal rotation to a new coordinate system, and  $\boldsymbol{\beta}_i$  as weights applied to that new coordinate system.

[Tucker (1972), Harshman (1972) and others prefer to factor  $\hat{\mathbf{R}}_i$  into

$$\hat{\mathbf{R}}_i = \mathbf{W}_i^{1/2} \mathbf{C}_i \mathbf{W}_i^{1/2} \quad (\text{II.B.19})$$

with  $\mathbf{W}_i^{1/2}$  diagonal and  $\mathbf{C}_i$  defined to have unit diagonals.  $\mathbf{C}_i$  is then interpreted as a matrix of cosines of angles between dimensions, and  $\mathbf{W}_i^{1/2}$  as weights applied to these oblique dimensions. Harshman's PARAFAC-2 assumes  $\mathbf{C}$  is constant over subjects, but  $\mathbf{W}_i$  differs, thereby guaranteeing a unique orientation of coordinate axes. Tucker's Three-Mode Scaling, however, has no such uniqueness of orientation property.]

*Phase II of IDIOSCAL: Modification of Schönemann's "Analytic Solution" to Determine Analytic Approximation to INDSCAL.* If the INDSCAL model holds (exactly) then

$$\mathbf{R}_i = \mathbf{T}\mathbf{W}_i\mathbf{T}', \quad (\text{II.B.20})$$

for some  $\mathbf{T}$  (in general, nonorthogonal), and with  $\mathbf{W}_i$  diagonal. The essence of Schönemann's (1972) analytic solution seems to be that, if Eq. (II.B.20) holds for any two  $i$  (say  $i = 1$  and  $2$ ) with  $\mathbf{W}_i$  nondegenerate (that is, all diagonals nonzero) for both, we can solve exactly for  $\mathbf{T}$  (that is guaranteed at least to "fit" those two). This is because two square symmetric matrices are always simultaneously diagonalizable by a matrix  $\mathbf{T}$ , which is not, however, orthogonal (in general). Since clearly  $\mathbf{T}$  is only defined up to post-multiplication by a diagonal, we may, without loss of generality, assume  $\mathbf{T}$  to be so defined that

$$\mathbf{W}_1 = \mathbf{I}. \quad (\text{II.B.21})$$

Thus

$$\mathbf{R}_1 = \mathbf{T}\mathbf{T}'. \quad (\text{II.B.22})$$

$\mathbf{T}$  can be decomposed as

$$\mathbf{T} = \mathbf{U}\boldsymbol{\beta}\mathbf{V}', \quad (\text{II.B.23})$$

with  $\mathbf{U}$ ,  $\mathbf{V}$  orthogonal and  $\boldsymbol{\beta}$  diagonal, so that

$$\mathbf{R}_1 = \mathbf{U}\boldsymbol{\beta}^2\mathbf{U}'. \quad (\text{II.B.24})$$

Thus, factoring  $\mathbf{R}_1$  yields  $\mathbf{U}$  and  $\boldsymbol{\beta}^2$  (and thus  $\boldsymbol{\beta}$ ). We may then define

$$\begin{aligned} \mathbf{R}^*_2 &= \boldsymbol{\beta}^{-1}\mathbf{U}'\mathbf{R}_2\mathbf{U}\boldsymbol{\beta}^{-1} = \boldsymbol{\beta}^{-1}\mathbf{U}'\mathbf{T}\mathbf{W}_2\mathbf{T}'\mathbf{U}\boldsymbol{\beta}^{-1} \\ &= \boldsymbol{\beta}^{-1}\mathbf{U}'\mathbf{U}\boldsymbol{\beta}\mathbf{V}\mathbf{W}_2\mathbf{V}'\boldsymbol{\beta}\mathbf{U}'\mathbf{U}\boldsymbol{\beta}^{-1} = \mathbf{V}\mathbf{W}_2\mathbf{V}' \end{aligned} \quad (\text{II.B.25})$$

(since  $\mathbf{U}'\mathbf{U}$ , and thus  $\boldsymbol{\beta}^{-1}\mathbf{U}'\mathbf{V}\boldsymbol{\beta}$ , =  $\mathbf{I}$ ). Thus, factoring  $\mathbf{R}^*_2$  yields  $\mathbf{V}$  (and, incidentally,  $\mathbf{W}_2$ , although that is of no real interest). Having thus obtained  $\mathbf{U}$ ,  $\boldsymbol{\beta}$ , and  $\mathbf{V}$ , they may be put together, according to Eq. (II.B.23), to define  $\mathbf{T}$  (which may be further post-multiplied by a diagonal matrix, if desired, for normalization purposes). Schönemann chooses the two subjects, in effect, to be the "average subject" whose  $\mathbf{R}$  matrix is the average of those for the real subjects, i.e.,



$$\mathbf{R} = \frac{1}{m} \sum_i^m \mathbf{R}_i, \quad (\text{II.B.26})$$

plus one of the “real” subjects (apparently arbitrarily chosen). Using the average subject is sensible, from a statistical point of view, and is also correct mathematically since it is easy to show that, if Eq. (II.B.20) holds for all  $i$ , then

$$\mathbf{R} = \mathbf{T}\mathbf{W}\mathbf{T}', \quad (\text{II.B.27})$$

showing that Eq. (II.B.20) also holds for this average subject, with  $\mathbf{W}$  replacing  $\mathbf{W}_i$ . The weakness of Schönemann’s (1972) solution from a statistical point of view is the choice of the second subject as some arbitrary real subject. Bad choice of this second subject could result in a very bad solution. Our modification of Schönemann’s procedure rests essentially on a more representative choice of the two subjects (or pseudosubjects, since both are composites of the “real” ones).

We have pursued two approaches to this. The first is to use a kind of crude clustering procedure to group the subjects into two groups of about equal size so that the profiles in the two groups are maximally different. “Average subjects” are then defined for each of the two groups, and those two group averages are used as the basis for finding the appropriate  $\mathbf{T}$ .

In the second approach the first subject is, as in Schönemann’s approach, the “average subject,” as defined in (II.B.26). Using the  $\mathbf{U}$  and  $\boldsymbol{\beta}$  found for that subject, we define the matrices  $\mathbf{R}^*_i$  as in (II.B.25). Since  $\mathbf{R}^*_i = \mathbf{V}\mathbf{W}_i\mathbf{V}'$ , (with  $\mathbf{V}$  orthogonal and  $\mathbf{W}_i$  diagonal), it follows that

$$(\mathbf{R}^*_i)^2 \equiv \mathbf{R}^*_i\mathbf{R}^*_i = \mathbf{V}\mathbf{W}_i^2\mathbf{V}', \quad (\text{II.B.28})$$

so that

$$\mathbf{Q} = \sum_i^m k_i (\mathbf{R}^*_i)^2 = \mathbf{V}\mathbf{W}^{(2)}\mathbf{V}', \quad (\text{II.B.29})$$

where

$$\mathbf{W}^{(2)} = \sum_i^m k_i \mathbf{W}_i^2 \quad (\text{II.B.30})$$

(the  $k_i$ ’s being weights for the different matrices).

$\mathbf{Q}$ , then, defines the second “pseudosubject”. Note that  $\mathbf{Q}$  is of the same general form as  $\mathbf{R}^*_i$ , so that factoring it should yield  $\mathbf{V}$  exactly (in the exact case) or approximately (in the more usual case of noisy data).  $\mathbf{Q}$ , however, provides a composite of all the subjects, but a different one than provided by  $\mathbf{R}$ .

We have tried two different ways of defining  $\mathbf{Q}$ , differing in definition of the weights. One is essentially the unweighted case, in which  $k_i = 1/m$  for all  $i$ . In the other case,  $k_i$  was defined as:

$$k_i = \frac{r_i^2}{\text{tr}(\mathbf{R}^*_i)^2},$$

where  $r_i$  is the correlation between  $d^2$ 's and predicted  $d^2$ 's in “Phase I” (in which the general IDIOSCAL model of Eq. (II.B.3) is fit).

*Finding the weights for the INDSCAL model.* Once the  $\mathbf{T}$  yielding the correct orientation of axes is found (or a hopefully reasonable approximation thereto, as described above), we may find the INDSCAL weights as outlined below. (The  $x$ 's below have presumably been defined by use of this  $\mathbf{T}$  and so correspond to the “correct” dimensions). Recalling the INDSCAL model:

$$\left[\delta_{jk}^{(i)}\right]^2 \cong \left[d_{jk}^{(i)}\right]^2 = \sum_i^r w_{it} (x_{jt} - x_{kt})^2 \quad (\text{II.B.32})$$

in matrix form, this can be written as:

$$\delta_i^{[2]} \cong \mathbf{d}_i^{[2]} = \Delta^* \mathbf{W}'_i, \quad (\text{II.B.33})$$

where  $\mathbf{d}_i^{[2]}$  is defined as before (in Eq. II.B.14),  $\delta_i^{[2]}$  is the analogous column vector with  $\left[\delta_{jk}^{(i)}\right]^2$  replacing  $\left[d_{jk}^{(i)}\right]^2$ , while  $\mathbf{W}_i$  is the row vector (of  $r$  components) with general entry  $w_{it}$ , and  $\Delta^*$  is the  $\begin{pmatrix} n \\ 2 \end{pmatrix}$  by  $r$  matrix

$$\Delta^* = (\Delta^*_{(jk)t}), \quad (\text{II.B.34})$$

with

$$\Delta^*_{(jk)t} = (x_{jt} - x_{kt})^2. \quad (\text{II.B.35})$$

Much as before, the least squares solution for  $\mathbf{W}_i$  is:

$$\hat{\mathbf{W}}_i = \delta_i^{[2]'} \Delta^* (\Delta^* \Delta^{*'})^{-1}, \quad (\text{II.B.36})$$

which immediately yields estimates of the weights.

*Estimation with and without constant term.* The estimation schemes above have involved no additive constant terms for the  $d^2$ 's. It is conceivable, however, that better fits could be obtained by adding such constant terms. This means that Eq. (II.B.3) is modified to become:

$$\left[ \delta_{jk}^{(i)} \right]^2 \cong (\mathbf{x}_j - \mathbf{x}_k) \mathbf{R}_i (\mathbf{x}_j - \mathbf{x}_k)' + c_i \quad (\text{II.B.37})$$

while Eq. (II.B.32) becomes:

$$\left[ \delta_{jk}^{(i)} \right]^2 \cong \sum_t w_{it} (x_{jt} - x_{kt})^2 + c_i \quad (\text{II.B.38})$$

It is straightforward to alter the regression schemes for estimating the  $\mathbf{R}_i$ 's or the  $w_{it}$ 's, as the case may be, to incorporate such a constant. This is done by simply adding an extra independent pseudovisible (whose values are all 1) to the regression scheme. This will, of course, change the estimates of the  $\mathbf{R}$ 's and  $w$ 's to some degree. Inclusion of this constant has advantages for interpretation of the  $F$  ratios to be described later. As will be seen subsequently, it also seems to improve the fit in Phase II (corresponding to the INDSCAL approximation).

*Phase III of IDIOSCAL.* Phase III corresponds to a model assuming essentially no individual differences, so that all subjects are assumed to be equivalent to the average subject. An overall scale factor is allowed for each subject, however, and (possibly) an additive constant.

Thus the models are either

$$\left[ \delta_{jk}^{(i)} \right]^2 \cong a_i (\mathbf{x}_j - \mathbf{x}_k) (\mathbf{x}_j - \mathbf{x}_k)', \quad (\text{II.B.39})$$

or

$$\left[ \delta_{jk}^{(i)} \right]^2 \cong a_i (\mathbf{x}_j - \mathbf{x}_k) (\mathbf{x}_j - \mathbf{x}_k)' + c_i, \quad (\text{II.B.40})$$

where the  $\mathbf{X}$  matrix is the one derived for the average subject.

*Approximate F tests for comparing the three phases.* Since the model in the three phases are fit by using least-squares linear regressions (with appropriately defined pseudo-variables), it is possible to define approximate  $F$  tests for comparing models in the three phases (as well as for assessing goodness of fit in each independently). This is very closely analogous to similar approximate  $F$  tests in the PREFMAP, PREFMAP-2

and PREFMAP-3 procedures, for those familiar with this (Sec. III.B.). This is most appropriate when the constant term has been included, since otherwise the residual mean square is not an unbiased estimate of error variance. The approximate  $F$ 's and their degrees of freedom are defined below. (See Table 2).

The " $F$ 's" must, of course, be taken with a large grain of salt since, first of all the required normality assumptions cannot be taken seriously, and, secondly, the configurations (which define the "independent" pseudovariates) have been fitted to the data. Since, however, these  $F$ 's are computed for each subject separately, and since each subject plays only a small part in determining the configuration in each case, this second objection can presumably be ignored as the number of subjects grows "large". Possibly some adjustment of degrees of freedom would correct for it when the number of subjects is small. Presumably a "jackknife" procedure could be used, even for small numbers of subjects, but this would be expensive computationally. Analogous approximate  $F$  ratios (called Pseudo- $F$ 's) could be calculated to test "significance" of added dimensions in INDSCAL or IDIOSCAL. This could conceivably lead to a way of objectively assessing dimensionality in individual differences scaling. (A somewhat related approach based on a "leave one out" procedure has recently been investigated by Weinberg and Carroll 1986.)

### **II.C Application of Three-Way MDS to Some Ecological Data on Seaworms**

We now consider applications of Three-Way MDS methods to ecological data, illustrating this with a specific application to the data in the article by Fresi *et al.* (1983). (Some multivariate analyses are reported by Fresi *et al.* in that paper.) The Fresi *et al.* data involve frequencies of observation of 88 varieties of seaworms in samples taken from five sites in the harbor of Ischia in the Bay of Naples, over four time periods (February 1975, July 1975, February 1976 and July 1976). There are many ways in which these three-way (or even "higher way") methods could be applied to these data. The data array, which is presented in the Fresi *et al.* paper as a rectangular data table, is more appropriately formulated as a general three-way array (seaworms  $\times$  sites  $\times$  time periods), or even possibly as a *four-way* array (seaworms  $\times$  sites  $\times$  months  $\times$  years). A direct analysis, for example by the general three-way CANDECOMP procedure (see equation II.A.16 and related discussion), or the closely related PARAFAC procedure proposed by Harshman (1970) would be

**Table 2. Pseudo-F's for assessing and comparing models fitted in the IDIOSCAL procedure.**

Effect	Pseudo-F	$df_1$	$df_2$
Phase I	$\left(\frac{df_2}{df_1}\right) r_I^2 / (1-r_I^2)$	$r(r+1)/2$	$\binom{n}{2} - r(r+1)/2 - 1$
Phase II	$\left(\frac{df_2}{df_1}\right) r_{II}^2 / (1-r_{II}^2)$	$r$	$\binom{n}{2} - r - 1$
Phase III	$\left(\frac{df_2}{df_1}\right) r_{III}^2 / (1-r_{III}^2)$	1	$\binom{n}{2} - 2$
Phase I-Phase II	$\left(\frac{df_2}{df_1}\right) (r_I^2 - r_{II}^2) / (1-r_I^2)$	$r(r-1)/2$	$\binom{n}{2} - r(r+1)/2 - 1$
Phase II-Phase III	$\left(\frac{df_2}{df_1}\right) (r_{II}^2 - r_{III}^2) / (1-r_{II}^2)$	$r-1$	$\binom{n}{2} - r - 1$
Phase I-Phase III	$\left(\frac{df_2}{df_1}\right) (r_I^2 - r_{III}^2) / (1-r_I^2)$	$r(r+1)/2 - 1$	$\binom{n}{2} - r(r+1)/2 - 1$

NOTE:  $r_I$ ,  $r_{II}$  and  $r_{III}$  represent correlations (between  $d^2$  and  $\hat{d}^2$ ) calculated by IDIOSCAL for a particular individual in phases I, II and III, respectively.

possible. Since CANDECOMP generalizes straightforwardly to the 4 or higher way case, an analysis of the 4-way table mentioned could be subjected to 4-way CANDECOMP analysis (see Carroll and Pruzansky 1984). A number of questions arise as to how best to normalize the data in this case, however; furthermore it is not at all clear that the rather strong and specific model(s) assumed in this CANDECOMP/PARAFAC type of analysis is (are) appropriate to these data. On the other hand, the general IDIOSCAL model/method seemed *too* general. For these and other reasons it was decided to pursue an exploratory INDSCAL analysis, using the Pruzansky (1975) SINDSCAL program, in a way to be described below. (This was based to some degree on the general approach used by Wish in his analysis of a large battery of data on perception and subjective ratings of nations, reported in Wish and Carroll 1974.) The Fresi *et al.* data were frequencies, as indicated earlier. The marginal frequencies of the 88 species of seaworms over the 20 sites  $\times$  time periods

ranged from 1 (for over a dozen of the species) to well over a thousand. Because of this, our first impulse was to normalize these data by converting to relative frequencies so that the normalized data for all species would sum to one. While this seemed to us like a wise first step in normalizing these data, it did not lead to readily interpretable results in any of the further analyses we attempted. We therefore abandoned this normalization of the data, and attempted instead an alternate transformation of these data, suggested by Pierre Legendre, which should have the effect of somewhat more nearly equalizing the total weight of resulting data values for the various species, as well as reducing the skewness of these distributions. This transformation was of the form:

$$z_{jlm} = \log(f_{jlm} + 1) \quad (\text{II.C.1})$$

where  $f_{jlm}$  is the frequency of seaworm species  $j$  at site  $l$  for month  $m$  in year  $p$ , and  $z_{jlm}$  is the corresponding transformed value. Data transformations are discussed in some detail in section 2 of Gower's paper in this volume. After this initial transformation, we then further normalized the data to have zero mean and unit variance within each site  $\times$  month  $\times$  year, so that the final "normalized" data were of the form

$$y_{jlm} = \frac{z_{jlm} - \bar{z}_{lm}}{s_{lm}} \quad (\text{II.C.2})$$

where  $\bar{z}_{lm}$  is the mean and  $s_{lm}$  is the standard deviation of the  $z$ 's over all 88 species for site  $l$  on month  $m$  in year  $p$ .

Given this normalized (four-way) data array  $Y \equiv (y_{jlm})$ , we then proceeded to compute a number of *derived* dissimilarity matrices as follows.

Number of matrices

1	(Overall)	$D_1: d_{jk}^{(1)} = \sqrt{\sum_l \sum_m \sum_p (d_{jlm_p} - d_{klm_p})^2}$
5	(Site $l'$ )	$D_{Sl'}: d_{jk}^{(l')} = \sqrt{\sum_m \sum_y (d_{jl'm_p} - d_{kl'm_p})^2}$
$(l' = 1, 2, \dots, 5)$		
2	(Month $m'$ )	$D_{Mm'}: d_{jk}^{(m')} = \sqrt{\sum_l \sum_p (d_{jlm'_p} - d_{klm'_p})^2}$
$(m' = F, J)$		
2	(Year $p'$ )	$D_{Yp'}: d_{jk}^{(p')} = \sqrt{\sum_l \sum_m (d_{jlm_{p'}} - d_{klm_{p'}})^2}$
$(p' = 75, 76)$		
4	(Time $t'$ )	$D_{Tt'}: d_{jk}^{(t')} = \sqrt{\sum_l (y_{jlm_{t'}, p_{t'}} - y_{klm_{t'}, p_{t'}})^2}$
$(t' = F75, J75, F76, J76)$		

(where  $m_{t'}$  and  $p_{t'}$  are the month and year, respectively, associated with time period  $t'$ ).

In the above, the sites are encoded as simply S1 through S5, the months as *F* (for February) and *J* (for July) the years as 75 (for 1975) and 76 (for 1976) and the 4 “times” as corresponding combinations of the month and year codes. Obviously, these various matrices are far from independent of one another. (In fact, just to take one example, the square of the overall dissimilarity for  $j$  and  $k$  is just the sum of the squares of the five site dissimilarities (or of the four “time” dissimilarities). However, as a “first start” on an exploratory data analysis for these data, we used the resulting 14 matrices [1 overall + 5 sites + 2 months + 2 years + 4 times (months  $\times$  years)] as input to an INDSCAL analysis, using the SINDSCAL program. In this case the seaworm species comprised the “stimuli,” and the 14 derived dissimilarity measures defined the data sources. Since each of these dissimilarity matrices was, in fact, defined as a Euclidean distance matrix, we used the option in SINDSCAL specifying that the data *were* Euclidean distances. Thus the total input to SINDSCAL comprised 14 matrices, each a symmetric half matrix of Euclidean distances among the 88 worms (so each of the 14 matrices had  $88 \cdot 87/2 = 3828$  entries, for a total of  $14 \cdot 3828 = 53592$  data values). Needless to say, this was a rather large data array, at least for so computationally intensive a procedure as SINDSCAL! Analyses were done in 1 through 6 dimensions. The fit measures (VAF in the derived scalar products) are given in Table 3. Based on

**Table 3.** Fit measures (Variance accounted for in derived scalar products) for SINDSCAL solutions in 6 down to 1 dimension(s) for 14 dissimilarity matrices derived from Fresi *et al.* data.

Dimensionality	Total Variance
1	.594
2	.778
3	.833
4	.872
5	.900
6	.922

**Table 4.** Variance accounted for in dimensions from KYST2-A and MDPREF analyses when mapped into four dimensional SINDSCAL solution.

Solution/Dimension Code	$R^2$ (VAF) by four SINDSCAL dimensions
k4-1	.988
k4-2	.990
k4-3	.809
k4-4	.603
k3-1	.985
k3-2	.985
k3-3	.718
k2-1	.976
k2-2	.972
k1-1	.840
md1	1.000
md2	.999
md3	.941
md4	.880

the pattern of these fit measures, and on inspection of the results, it was decided to report the 4-dimensional INDSCAL solution (although it was somewhat debatable whether the 4 or the 5-dimensional solution should be chosen).

Interpreting these results, since the present author knows little of the biology of the 88 species of seaworms, we focused on the pattern of the 14 different data sources in the "subject (or source) space." The weights for the 14 data sources on the four dimensions are displayed graphically in Figures 4 and 5. (As can be seen in the Figures, some of the source weights are slightly negative, a condition which should not occur in INDSCAL, since all subject or source weights should be zero or positive. These are only very slightly negative, however, and so can probably be plausibly interpreted as essentially zero weights, which have become slightly negative due to error in the data. We will henceforth interpret these, in fact, as though they are zero.)

Figure 4 shows the dimension one vs. two plane of the source space. Dimension one has very high weight for sites 4 and 5, and for the 1975 time periods. Since only sites 4 and 5 have very large weights on this dimension and the 1975 time periods have large weights on it, we may label dimension one (sites 4 and 5; 1975). The corresponding



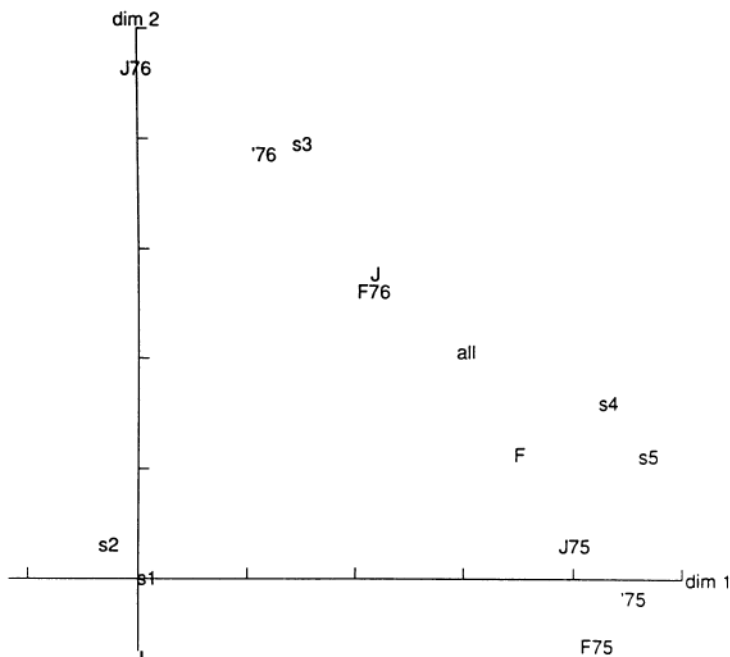


Figure 4. Dimension one-two plane of source space for SINDSCAL solution for Fresi *et al.* data.

variable seems to be one that was particularly prevalent in sites 4 and 5, somewhat in sites 3, and not at all in sites 1 and 2, and quite salient in 1975, but very weakly present in 1976. Dimension two, on the other hand, seems to be very strongly weighted in site 3, very slightly in sites 4 and 5, but not at all in sites 1 or 2.

Whatever dimension two taps was especially prevalent in 1976. Thus dimension two will be labeled (site 3; 1976). One interesting point in these results is that sites 4 and 5 seem to occupy essentially the same location in all four dimensions. Thus these two sites were, insofar as these analyses are concerned, virtually indistinguishable.

We now look at the plane defined by the remaining 2 dimensions, dimensions 3 and 4, in Figure 5. What "jumps out" at us in this plane is that site 2 has high weight on dimension 3 and virtually zero weight on dimension 4, while site 1 reverses this pattern, having almost identically zero weight on dimension three but a very large weight on dimension four. Sites 3, 4 and 5 have essentially zero weights on both these dimensions, while the matrices relating to the various time periods (as well as the "all" matrix corresponding to overall dissimilarities over all sites  $\times$  time periods) generally have moderate weights on both. Thus dimension three seems to correspond to whatever

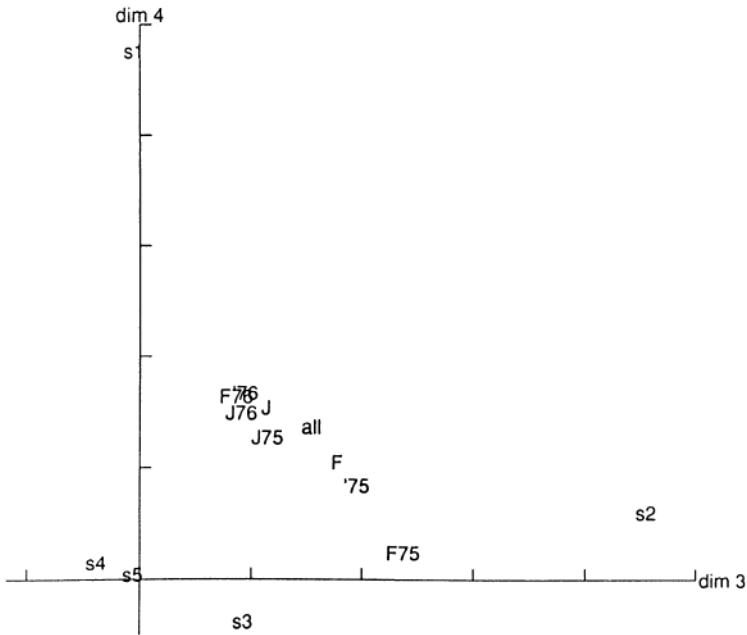


Figure 5. Dimension three-four plane of source space for SINDSCAL solution for Fresi *et al.* data.

distinguishes site 2 from the others, while dimension four corresponds to the variable most prevalent in site 1.

While they must be taken with a fairly large “grain of salt,” distances among the source points have a certain interpretation in these INDSCAL subject (source) space plots. Without actually doing the computation we can see from inspection of these two planes that sites 4 and 5 are exceedingly close, and in turn are relatively closer to site 3 than to either sites 1 or 2. Conversely, sites 1 and 2 are by far closer to one another than to any of the other three sites.

The variables corresponding to these dimensions are defined in a sense by the coordinates of the points corresponding to the species of seaworms on these dimensions. These can be seen graphically by inspection of Figures 6 and 7, which show these 88 points plotted also in the plane of dimension one versus two and that of three versus four. It should be kept in mind that, despite the “ $z = \log(f+1)$ ” transformation that was used, the relative size of the distribution of variables for these species were quite different, and so frequency effects are clearly affecting these results. Thus, the seaworm species “farthest out” on these dimensions tend to be those whose overall frequencies were greatest. Table 2 gives the actual species of seaworm corresponding to the 88

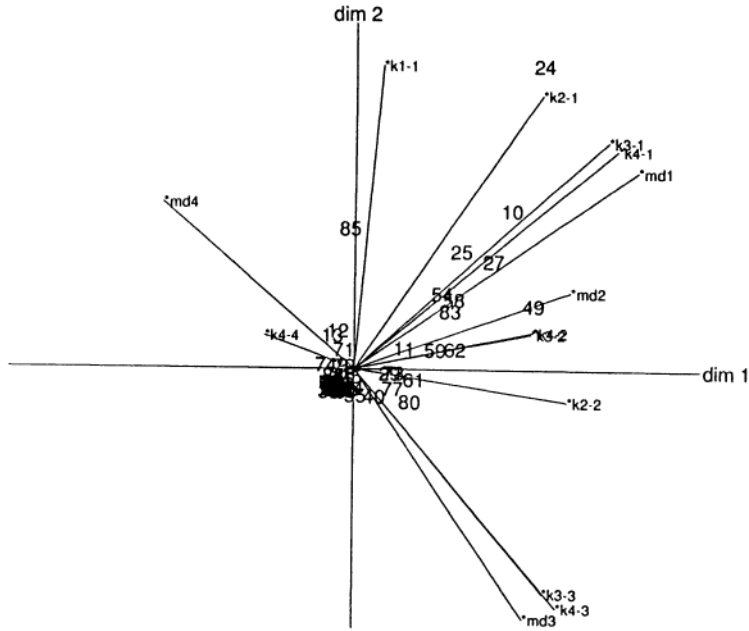


Figure 6. Dimension one-two plane of stimulus (species) space for SINDSCAL solution for Fresi *et al.* data. Vectors are from mapping of KYST and MDPREF dimensions into SINDSCAL space, using PREFMAP.

points shown in these figures, for the benefit of those knowledgeable about these species. (It should be commented that we have reflected some of these dimensions so that the positive values always tend to imply greater frequency.)

In Figures 6 and 7 vectors are shown indicating the direction best corresponding to the dimensions derived from the one through four dimensional KYST2-A solutions shown in Figures 2 and 3. Since there were a total of ten such dimensions (4 + 3 + 2 + 1 for the four through one dimensional solutions, respectively) there are a total of ten vectors indicated corresponding to these. These are encoded “ $kr-t$ ” where “ $kr$ ” stands for the KYST  $r$ -dimensional solution, while  $t$  indicates the  $t^{\text{th}}$  dimension in that solution. Since, in the case of KYST, the solutions for different dimensionalities do not have any *necessary* correspondence, these ten dimensions are all distinct, although it will be noted that the  $t^{\text{th}}$  dimension in solutions for different dimensionalities *do* tend to correspond fairly closely, though certainly not perfectly. In addition to these ten dimensions from the various KYST solutions, four other vectors, labelled md1 through md4 are also shown. These are four dimensions from another MDS analysis, called MDPREF, which will be described in section III.

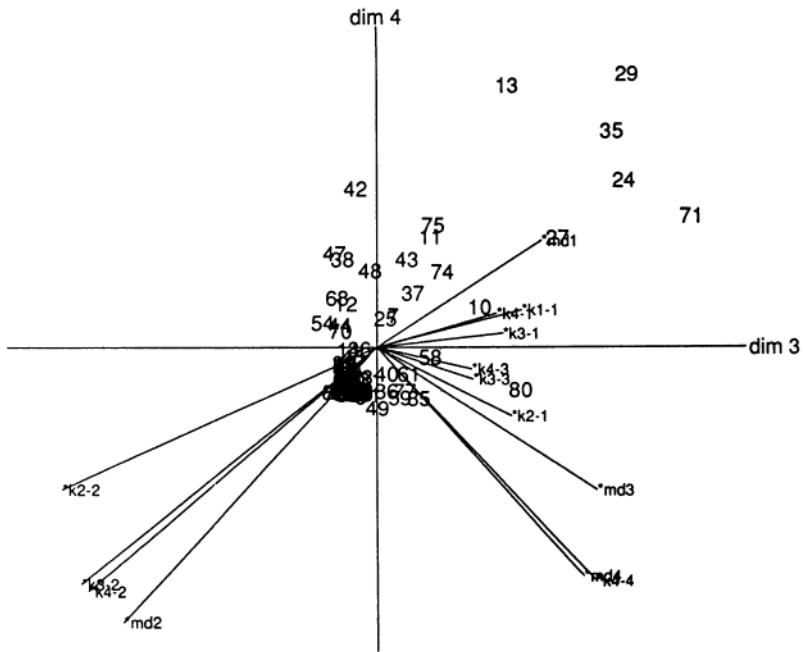


Figure 7. Dimension three-four plane of species space for SINDSCAL solution for Fresi *et al.* data. Vectors are as in Figure 6, except projected into three-four plane of SINDSCAL space.

Table 4 gives figures which indicate how well these fourteen dimensions (ten from the one through four dimensional KYST solution plus four from the four dimensional MDPREF analysis) from the other analyses “fit” into the four dimensional SINDSCAL space. The values in Table 4 are squared multiple correlations ( $R^2$ 's), which can be interpreted as proportions of variance accounted for in these fourteen dimensions from KYST and MDPREF via the four SINDSCAL dimensions. (Since the KYST analyses had to be done on only a subset of 55 of the seaworm species, these  $R^2$ 's were necessarily based only on this subset of 55 of the total 88 species, however.) In fact, the procedure used for determining these best fitting directions (or vectors) was the PREFMAP-3 procedure described also in section III. The particular analysis done in these cases involved fitting the vector model, with linear regression options. In the case of this particular set of options, PREFMAP is equivalent to the use of multiple linear regression. Thus we may view these vector directions as being defined by the regression coefficients from a multiple linear regression. In fact, the projections of these vectors onto the SINDSCAL coordinate axes are, in the present case, proportional to the Beta coefficients for these regressions.

The point to be drawn from these PREFMAP/multiple regression analyses is that the four SINDSCAL dimensions capture quite well essentially all the dimensions emerging from the other MDS analyses. The fact that the vector directions best representing these other dimensions do not coincide directly with the coordinate axes indicates, however, that the SINDSCAL dimensions do not correspond in a simple one-one fashion with these KYST and MDPREF dimensions. Rather, each of these alternative dimensions corresponds to a different *linear combination* of the SINDSCAL dimensions. Since the SINDSCAL dimensions have a unique orientation, while those in the other solutions are defined only up to arbitrary rotation (or linear transformation), we feel it appropriate to treat the SINDSCAL solution as defining the “reference space” in terms of which the others are defined. As already seen, the SINDSCAL dimensions do have a particularly simple association with the various derived dissimilarity matrices — particularly with those defined for the five different sites. This suggests that these dimensions may correspond to variables characterizing the species having especially meaningful relations to the geographic variables distinguishing sites (as well as, secondarily, to variables related to the four different time periods).

To correct for the frequency effect we present another pair of plots in which the following transformations have been effected.

- (1) The origin of the space was first transformed so all the coordinate values were non-negative (by subtracting the smallest algebraic coordinate value on each dimension from all the coordinates).
- (2) After this translation to a “more or less” rational origin (such that essentially all the very low frequency species are at or very close to that origin) we now multiply all coordinates of each species point by the reciprocal of its marginal value in the “ $z = \log(f + 1)$ ” scale. This tends to convert all values to something approximating a relative frequency scale. The coordinate value on each dimension after this transformation can be interpreted as the *relative* value of the species on that dimension (relative to its overall frequency in the samples taken from all 20 sites  $\times$  time periods comprising these data). While these plots, shown in Figures 8 and 9, are no more interpretable to us than were the earlier figures (6 and 7) we hope they may help ecologists or other biologists in interpreting these dimensions.

For general discussions of two- and three-way MDS and related models and methods for proximity data, we refer the reader to Carroll and Wish (1974a,b), Wish and Carroll (1974), Kruskal and Wish (1978), Carroll and Arabie (1980), Shepard (1980), Carroll and Pruzansky (1980, 1986) and Arabie, Carroll and DeSarbo (in press).

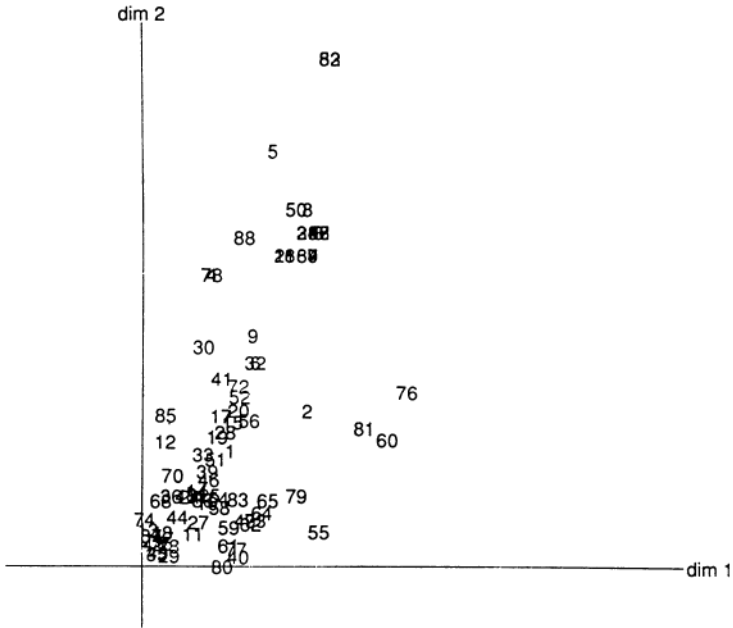


Figure 8. Transformed dimension one-two coordinates of SINDSCAL species space for Fresi *et al.* data.

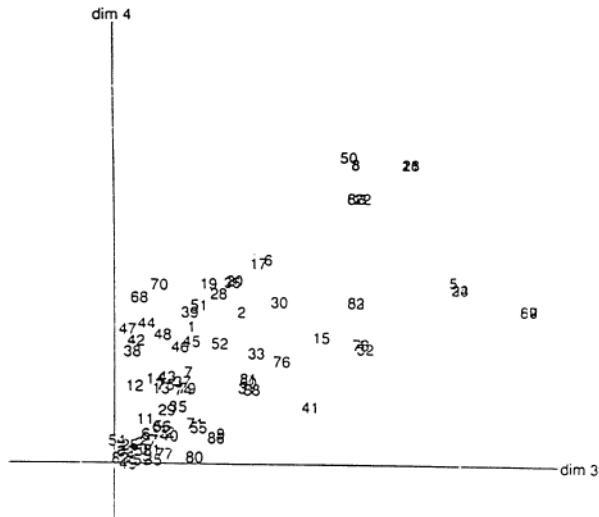


Figure 9. Transformed dimension three-four coordinates of SINDSCAL species space for Fresi *et al.* data.

### III. MDS (AND OTHER MULTIDIMENSIONAL ANALYSES) OF PREFERENCE (OR OTHER DOMINANCE) DATA

#### III.A MDPREF

MDPREF (Carroll and Chang 1964b), standing for MultiDimensional PReferences scaling, is a model and method implemented in a computer program by Chang and Carroll (1968) to perform an “internal” analysis of  $m$  subjects’ preferences (or any type of dominance) data. The program utilizes a vector preference model and develops simultaneously the vector directions for the subjects (or the site  $\times$  time “variables” in our “seaworms” example) and the configuration of stimuli (or objects “seaworms” in the present case) in a common space. A theoretical discussion of this and other methods of internal analysis of preference (or proximity) data is provided by Carroll (1972, 1980) or Heiser (1981). By an “internal” analysis, we mean that both stimulus (object) points and subject (variable) parameters (vectors in this case) are determined entirely from the preference (dominance) data.

#### *Theoretical Discussion*

Since MDPREF was originally developed for individual differences preference analysis, we will often refer to it as though it is an analysis of preference judgments by different subjects. However, it can be applied to any kind of “dominance” data, in which each of  $m$  “variables” measure the relative dominance of each of  $n$  “objects” in some respect. In MDPREF the dominance judgments can be either paired comparisons, rankings, or ratings of stimuli or other objects. The following discussion is for the latter case; the development is quite similar for the former two, except that in the case of rankings the ranks are substituted for preference (dominance) scale values, while if the data are paired comparisons, preference (dominance) scale values are derived by methods described by Chang and Carroll (1968). See also Carroll (1972, 1980).

The model assumes that stimulus (object) points are projected onto subject (variable) vectors, with preference (degree of dominance) being determined by the relative size of these projected values (the larger value being preferred). Let  $\mathbf{x}_j = (x_{j1}, \dots, x_{jr})$  represent an  $r$ -dimensional stimulus point for the  $j$ -th stimulus and  $\mathbf{y}_i = (y_{i1}, \dots, y_{ir})$  represent the vector for subject  $i$  in the same  $r$ -dimensional space. (For simplicity, we now speak simply of preference of subjects for stimuli; the reader can make the necessary substitution of terminology if desired.) Then  $\hat{s}_{ij}$ , the estimated preference

scale value of stimulus  $j$  for subject  $i$ , is defined by the scalar product:

$$\hat{s}_{ij} \equiv y_i \mathbf{x}'_j, \quad (\text{III.A.1})$$

(the expression on the right being the scalar product in matrix notation). This can be written, more generally, in matrix notation as follows:

Let  $\mathbf{X} \equiv (x_{ji})$  be the  $n \times r$  matrix of stimulus coordinate values and  $\mathbf{Y} \equiv (y_{it})$  be the  $m \times r$  matrix of the termini of subject vectors, then

$$\hat{\mathbf{S}} \equiv (\hat{s}_{ij}) = \mathbf{Y} \mathbf{X}'. \quad (\text{III.A.2})$$

The problem is to determine the matrices  $\mathbf{Y}$  and  $\mathbf{X}'$  from the set of paired comparison judgments such that  $\hat{\mathbf{S}}$  accounts for the paired comparisons data as well as possible in some statistically well-defined sense (realized by minimizing an "objective function" embodying the statistical criterion to be optimized). Carroll and Chang (1964b) describe procedures – one iterative and one utilizing an Eckart-Young (1936) decomposition – that accomplish this task. [In more modern terminology, the "Eckart-Young decomposition" is frequently called, or closely related to, the "singular value decomposition" (svd).] It is the latter that is implemented by MDPREF, and that is described below.

If the input data are already scale values of preference (this matrix  $\mathbf{S}$  is called the "first score matrix") the program proceeds to decompose  $\mathbf{S}$  by the Eckart-Young procedure, which involves computing eigenvalues and eigenvectors of the matrix  $\mathbf{S}'\mathbf{S}$  or  $\mathbf{S}\mathbf{S}'$  (whichever is smaller). If the input data are paired comparisons, they are first converted to a "first score matrix" of scale values by summing over rows and/or columns of each paired comparisons matrix. Monte Carlo analyses by Carroll and Chang have indicated that the simpler, Eckart-Young, procedure works as well with errorful data as the iterative one. This is the reason MDPREF utilizes only the Eckart-Young procedure. This overall procedure can be shown to have certain least squares properties. Among other properties, in the case in which the original data are paired comparisons, it provides a least squares fit in a certain sense to the original paired comparisons data, schematized as a matrix of plus and minus ones (and possibly some zeros). See Carroll (1972, 1980) for details.



### *Input Options*

As noted earlier, MDPREF has two input options, namely, paired comparisons and direct judgments of preference scale values. In the case of paired comparisons, options exist for reading in weight matrices specific to each subject and for handling missing data. In the case of direct preference judgments (e.g., rankings) two options exist for normalization – either: a) subtracting row means or b) subtracting row means and then dividing entries by the standard deviation of values for that row.

### *Output Details*

The following are the major output categories entailed in a typical run of MDPREF:

1. First score matrix normalized according to alternative chosen from above options.
2. Cross-products matrix of subjects.
3. Cross-products matrix of stimuli.
4. Eigenvalues of the first score matrix.
5. Estimates of the first score matrix after factorization. (This is sometimes called the “second score” matrix.)
6. Coordinates of stimuli and vector directions for subjects in the user-specified dimensionality.
7. Plots of some or all pairs of dimensions, including both stimuli and subjects. (Many different versions of MDPREF exist, with different details regarding this and other options). See Chang and Carroll (1968) for further details on the specific version of MDPREF available on the MDS-1 tape.

### **III.B PREFMAP, PREFMAP-2 and PREFMAP-3**

PREFMAP (PREference MAPping) is a procedure that analyzes preference (or other dominance) data in terms of a set of multidimensional preference models, developed by Carroll and Chang (1968), which include and generalize the “vector model” first proposed by Tucker (1960) and the basic Coombsian unfolding model of preference (Coombs, 1964). Collectively, these are called the linear-quadratic hierarchy

of models. PREFMAP utilizes a *known* configuration of stimuli and attempts to portray an individual's preference data via this hierarchy of models. PREFMAP is called an *external* analysis, since the stimulus (object) space is (externally) given, and only the subject (variable) parameters (e.g., ideal points or vectors) are to be determined. Specifically, PREFMAP consists of four phases, corresponding to analysis in terms of four models. The phases are referred to as Phases I, II, III and IV. As one goes from Phase I to Phase IV, the underlying assumptions become stronger and model complexity is therefore considerably reduced.

### *Theoretical Discussion*

PREFMAP starts out with the following assumptions:

1. A group of individuals share the same perceptual configuration of  $r$  dimensions for a set of  $n$  stimuli. Let  $\mathbf{X} = (x_{jt})$   $j = 1, 2, \dots, n; t = 1, 2, \dots, r$  represent the common perceptual space. Generally  $\mathbf{X}$  will be *externally* defined (i.e., given *a priori* as input to the PREFMAP procedure).

2. Further, the preference value for the  $j^{\text{th}}$  stimulus of any individual, say the  $i^{\text{th}}$ , is (at least) monotonically related to the squared "distance" between the individual's ideal point and the location of the stimulus in space. Let the matrix  $\mathbf{S} \equiv (s_{ij})$   $i = 1, 2, \dots, m; j = 1, 2, \dots, n$  represent the scale values of  $m$  individuals' preferences for the  $n$  stimuli. Each row of the  $\mathbf{S}$ -matrix represents the scale values for individual  $i$ 's preferences for the  $n$  stimuli. (For convenience, we assume that smaller values represent higher preferences.) In general, PREFMAP assumes  $F_i(s_{ij}) \cong d_{ij}^2$ . The models differ in definition of  $d_{ij}^2$ , and in that of  $F_i$ .

Two versions of PREFMAP models may be distinguished – metric and nonmetric. In the metric version the function  $F_i$  is assumed to be linear, while a general monotonic function, not specified *a priori*, is permitted in the nonmetric case. Thus, the preference scale values are assumed to be defined on at least an interval scale in the metric version while only their ordinal relationships are utilized in the nonmetric version. We discuss the metric version of the PREFMAP algorithm first and then describe the nonmetric case.

### *Metric Version of the PREFMAP Algorithm*

In the metric version of the PREFMAP algorithm, it is assumed that the scale values of preference are *linearly* related to squared distance, that is, that  $F_i$  is linear. Assuming  $F_i$  has nonzero slope, we may invert it and write:

$$s_{ij} \cong ad_{ij}^2 + b \quad (\text{III.B.1})$$

where  $a$  and  $b$  are constants ( $a > b$ ) and  $\cong$  denotes approximate equality (except for error terms not expressed).

Let  $\mathbf{x}_j \equiv (x_{j1}, \dots, x_{jr})$  represent the row vector of coordinates of the  $j^{\text{th}}$  stimulus ( $j = 1, 2, \dots, n$ ) and  $\mathbf{y}_i \equiv (y_{i1}, \dots, y_{ir})$  represent the vector of coordinates of the ideal point for the  $i^{\text{th}}$  individual ( $i = 1, 2, \dots, m$ ). Given the above relationship and input data for  $\mathbf{x}_j$  and  $s_{ij}$ , the PREFMAP method solves, for each individual, for estimates of the coordinate values of the vector  $\mathbf{y}_i$ , and, depending on the model, possibly for additional parameters associated with individuals.

In model IV the squared distances are defined in a special way which corresponds to the special case when the ideal point is infinitely distant from the stimuli, so that only its *direction* matters. In this special case, the squared distance is actually defined by a linear equation, and can also be viewed as equivalent to projection on a vector in the appropriate direction; thus the name "vector model". This equivalence of the linear, or vector, model to the unfolding model with ideal points at infinity is demonstrated in Carroll (1972, 1980).

Four alternative models for relating preference data to a given stimulus space, called models I, II, III and IV, are included in the hierarchy proposed by Carroll and Chang. The four models correspond, in the obvious fashion, to the four phases of PREFMAP, in a decreasing order of complexity. Phase I fits a highly generalized unfolding model of preference (model I); Phase II utilizes a more restrictive model assuming weighted Euclidean distances analogous to those assumed in the INDSCAL model discussed earlier; Phase III is the "simple" or Coombsian unfolding model in which ordinary (unweighted) Euclidean distances are assumed; and Phase IV is the linear, or "vector", model. Phases I, II and III differ in the way the term  $d_{ij}^2$  is formulated, i.e., in the definition of the metric, while Phase IV can be viewed as putting certain restrictions on ideal point locations, as discussed earlier.

All four phases utilize regression procedures (quadratic or linear) to estimate coefficients which are then reparametrized to provide estimates of parameters associated with the corresponding model. This is described in detail in Carroll (1972, 1980).

### Phase I

One way to describe the model assumed in Phase I is to assume that both  $\mathbf{x}_j$  and  $\mathbf{y}_i$  are operated on by an orthogonal transformation matrix  $\mathbf{T}_i$  – which is idiosyncratic for each subject – and *weighted* squared distances are then computed from the *transformed values*. Thus, one defines:

$$\mathbf{x}_j^* = \mathbf{x}_j \mathbf{T}_i \quad (\text{III.B.2})$$

and

$$\mathbf{y}_i^* = \mathbf{y}_i \mathbf{T}_i, \quad (\text{III.B.3})$$

and then computes the (weighted) Euclidean squared distances  $d_{ij}^2$  by the formula:

$$d_{ij}^2 = \sum_{t=1}^r w_{it} (x_{jt}^* - y_{it}^*)^2. \quad (\text{III.B.4})$$

Geometrically, this corresponds to an orthogonal, or rigid, rotation of the coordinate system, followed by differential stretching of the new (rotated) coordinate system. Different rotations and different patterns of weights are allowed for each individual.

### Phase II

Phase II differs from Phase I in that it does not assume a different orthogonal transformation for each individual, although it allows differential weighting of dimensions, so that squared distances are computed simply by

$$d_{ij}^2 = \sum_{t=1}^r w_{it} (x_{jt} - y_{it})^2. \quad (\text{III.B.5})$$

### Phase III

Phase III is the “simple” unfolding model, but it allows the possibility that some or all of the dimensions have negative weight, making Phase III equivalent to Phase II, with weights  $w_{it} = \pm 1$  for each individual. To be precise, the weights  $w_{it} = \pm \delta_t$ , where each  $\delta_t = \pm 1$ .

### Phase IV

Phase IV utilizes the vector model in which preference values are related to coordinates of the stimulus space by an equation (excluding the error term) of the form:

$$s_{ij} \cong \sum_{t=1}^r b_{it} x_{jt} + c_i. \quad (\text{III.B.6})$$

This equation contains only linear terms, so least squares estimates of the  $b_{it}$ 's can be derived immediately by multiple linear regression procedures. Having estimated the coefficients  $b_{i1}, b_{i2}, \dots, b_{ir}$ , the direction cosines of the vector for the  $i^{\text{th}}$  individual are obtained by normalizing the vector of estimated coefficients  $\hat{\mathbf{b}}_i = (\hat{b}_{it})$  to unit length by dividing each  $\hat{b}_{it}$  by  $\sqrt{\sum_t \hat{b}_{it}^2}$ . Parameters of the other models are also fit by regression procedures, although these are more complex. The reader is referred to Carroll (1972, 1980) for a more detailed exposition of this.

In Phase II, much as in INDSCAL, the orientation of coordinate axes is critical. Since the axis orientation of the *a priori* space may be essentially arbitrary, an approximate solution is provided for the appropriate orientation. This will automatically be provided in either PREFMAP or PREFMAP-2 if Phase I precedes Phase II. Otherwise, Phase II can be entered directly, but with an initial solution for what is called the "canonical rotation". In Phase III the problem is a little more involved still, since a general linear transformation may be required. This can be viewed as entailing an orthogonal transformation followed by a differential weighting of dimensions. This, called the "canonical rotation and canonical weights", can also be solved for. In PREFMAP-3 it is optional whether the "canonical rotation" and/or "canonical weights" will be solved for. In some cases the orientation may be assumed to be correct as given and only the canonical weights asked for. PREFMAP, PREFMAP-2 and PREFMAP-3 all differ in how the canonical orientation and/or canonical weights are solved for. In fact, in PREFMAP-3 it is possible to solve for "canonical weights" without necessarily solving for the "canonical rotation." See Chang and Carroll (1972) or Meulman, Heiser and Carroll (1986) for details.

### *Nonmetric Version of the PREFMAP Algorithm*

It may be recalled that the nonmetric version of PREFMAP fits monotonic functions relating the preference scale values and the squared Euclidean distances between a subject's ideal point and the stimulus points. This is accomplished by the procedure described below.

1. Solve for the parameters of the appropriate regression equation (quadratic or linear) to predict the  $s_{ij}$ 's. This step is essentially the metric version of PREFMAP. The "predicted" values (from the model) will be called  $s_{ij}^{(1)}$   $i = 1, 2, \dots, m$ ;  $j = 1, 2, \dots, n$ .

2. Estimate the monotone function  $M_i^{(1)}$  for subject  $i$  that best predicts the estimates (the  $\tilde{s}_{ij}^{(1)}$ 's) from the original  $s_{ij}$ 's, using the procedure described by Kruskal (1964b) for least squares monotone regression. Define  $\hat{s}_{ij}^{(1)} \equiv M_i^{(1)}(s_{ij})$ .

3. Replace  $s_{ij}$  with  $\hat{s}_{ij}$  to compute a new set of predicted values,  $\tilde{s}_{ij}^{(2)}$ .

4. Using the new set of  $\tilde{s}_{ij}$ 's, compute a new monotone function  $M_i^{(2)}$  and a new set of  $\hat{s}_{ij}$ 's, namely  $\hat{s}_{ij}^{(2)}$ .

5. Continue this iterative procedure until the process converges (i.e., until no more changes occur in the monotone function or regression coefficients). Specifically, the process is terminated by reference to a parameter called CRIT. If the sum of squares of differences in the predicted  $\hat{s}_{ij}$ 's for the  $I^{\text{th}}$  and  $(I-1)^{\text{st}}$  iterations is less than CRIT, the process stops at the  $I^{\text{th}}$  iteration.

### *Input Parameters*

In all the PREFMAP programs, the preference data can be expressed in one of two ways: a) smaller values indicating higher preferences or b) larger values indicating higher preferences. The programs can start with any prespecified phase and can work their way down to any model of lower complexity. PREFMAP-3 actually allows different models to be fit for different subjects in the same analysis.

Other options include: a) normalization of original scale values versus leaving them as initially defined and b) computing each subject's scale values for each new phase or, alternatively, using the estimates of the previous phase as the original values for the

following phase. There are also various options concerning whether or not the canonical rotation and/or weights are computed prior to entering a particular phase.

### *Output Details*

A typical run of PREFMAP produces some or all of the following output:

1. Listing of all input parameters selected and the original configuration of stimuli.
2. For each subject the printout of the original scale values, regression coefficients and estimates of  $d_{ij}^2$  (or  $\hat{s}_{ij}$ , where  $\hat{s}_{ij} = a_i d_{ij}^2 + b_i$ , or equals projection of stimulus  $j$  on vector for subject  $i$  in the case of the “vector model”) for each phase and for each iteration in the case of the monotone (or nonmetric) version.
3. For Phase I (only) the direction cosines of each subject’s idiosyncratic rotation.
4. Coordinates (or direction cosines for Phase IV) of ideal point and weights of the dimensions specific to each subject. In Phase I, the orthogonal rotation matrix may also be printed for each subject. Depending on options selected, the canonical rotation matrix and/or canonical weights may also be provided as output.
5. Plot showing the relationship between the monotone transform of the scale values and original scale values (optional).
6. Plot showing the positions for ideal points or vector directions of all subjects as well as stimulus positions.
7. A summary table showing the correlation coefficients for each subject by each phase and corresponding  $F$ -ratios, including  $F$ -ratios for testing the statistical significance of the improvement in fit associated with moving from a simple to a more complex method. Such an  $F$  is associated with every pair of models (IV versus III, II or I; III versus II and I; and II versus I). In each case, it can be taken as assessing whether the more complex model (with a *lower* Roman numeral) fits the data significantly better than the simpler (higher numbered) model. These tests are possible because of the hierarchical embeddedness (or nested structure) of these models; that is, the fact that each “simpler” model is a special case of each more complex one. In terms of the algebraic structure of the models, each more complex model includes all the parameters of any simpler model, plus additional parameters. The situation is formally

equivalent to testing significance of additional terms in a stepwise regression scheme.

*PREFMAP-2* has the additional feature of allowing definition of a so-called “internal” stimulus configuration directly from the preference data itself. For further details on *PREFMAP* and *PREFMAP-2* see Chang and Carroll (1972). *PREFMAP-3* does not allow generation of such an “internal” stimulus configuration, but does have many other options. *PREFMAP-3* is much more flexible in the mix of models fit to different subjects. In a single analysis different subjects may be fit by different models in the hierarchy of models described here. These models are simply called, in *PREFMAP-3*, *G* for *General Unfolding*), *W* (for *Weighted Unfolding*), *U* (for simple *Unfolding*) or *V* (for *Vector model*). Greater flexibility also exists in *PREFMAP-3* in “metric” vs. “nonmetric” fitting for different subjects. See Meulman, Heiser and Carroll (1986) for details on *PREFMAP-3*.

It would seem in principle to be very interesting to apply the entire family of models in the *PREFMAP* hierarchy to the Fresi *et al.* data. For example, it would seem quite appropriate to fit model II (the simple unfolding, or “ideal point” model), using each of the site  $\times$  time period variables as a pseudo-subject, seeking an ideal point in the four dimensional space of seaworm species determined by *INDSCAL/SINDSCAL* such that the frequency of species for that site  $\times$  time period is inversely related to distance from that ideal point. One could think of this “ideal point” as the species of seaworm most ideally suited to that particular site/time period combination. Time constraints did not allow for a thorough analysis of these data via the *PREFMAP* hierarchy of models, however. We therefore opted for an *internal* analysis of the site  $\times$  time period variables, using the *MDPREF* vector model approach. *MDPREF*, as discussed earlier, simultaneously determines a space for the “stimuli” (species in this case) and the “subjects” (sites  $\times$  time periods) in terms of a vector model. A vector model can actually be thought of as an unfolding or “ideal point” model with the ideal points all infinitely distant (or, in practice, very far) from the stimuli (species), so that the vector direction simply corresponds to the direction of the ideal point from the centroid of the stimuli (species). It is of interest both to see how well *MDPREF* accounts for these data, and also how the structure of the species space relates to that determined by the three-way *INDSCAL/SINDSCAL* analysis.



### III.C MDPREF Analysis of the Fresi *et al.* Seaworm Data

We attempted an analysis of the Fresi *et al.* data on seaworm species using MDPREF. As indicated earlier, dominance relationships can be attributed to variables much more general than preference judgments (narrowly construed). More generally, dominance data are any data indicating the tendency of objects to dominate one another in some respect or context. Thus the relative frequency of the various seaworms at the 5 sites and the 4 time periods can be taken as dominance data for these species at these sites  $\times$  time periods. (In fact, dominance data, broadly defined, can be viewed as encompassing essentially any variety of multivariate data.)

We thus applied MDPREF to these data, treating the seaworm species as "stimuli" and the 20 sites  $\times$  months  $\times$  years as "subjects." The "total and marginal" variance accounted for (VAF) for dimensionalities from 1 through 20 are displayed in Table 5.

**Table 5.** Variance accounted for (VAF) and commulative VAF for MDPREF solutions in dimensionalities 1 through 20, for Fresi *et al.* data.

Dimensions	Variance	Cumulative Variance Accounted For
1	0.490	0.490
2	0.201	0.691
3	0.053	0.745
4	0.048	0.794
5	0.036	0.830
6	0.028	0.858
7	0.025	0.884
8	0.020	0.904
9	0.019	0.924
10	0.014	0.938
11	0.013	0.952
12	0.010	0.962
13	0.008	0.970
14	0.008	0.978
15	0.004	0.983
16	0.004	0.988
17	0.003	0.992
18	0.003	0.996
19	0.002	0.998
20	0.001	1.000

(In MDPREF, as mentioned earlier, the unrotated  $r-1$  dimensional principal axis solution is simply the  $r$  dimensional one with the least important – in VAF terms –

dimension dropped. Because of this “nesting” property, this calculation is straightforward). Based on the VAF figures, and on interpretability criteria, once again it was decided to report the four dimensional solution.

Since we are focusing, in our attempt to interpret these solutions, on the structure of the variables (sites  $\times$  time periods), we present the positions of the vectors for these 20 variables separately from the species points in Figures 8 and 10. In these Figures we use the same coding for these variables as in the Fresi *et al.* paper; a three symbol (number, letter, number) code. The first number (1-5) denotes the site, the letter denotes the month (F = February, L = July), while the third number denotes the year (5 = 1975, 6 = 1976). (We used an “L” rather than a “J” here to encode “July” to maintain consistency with the coding used by Fresi *et al.*). MDPREF does not, like INDSCAL/SINDSCAL, produce unique dimensions, so that rotation of coordinate axes is usually necessary to attain an optimally interpretable set of dimensions. In the present case, however, perhaps fortuitously, the orientation of axes originally obtained appears to lead to a quite interpretable structure (without rotation) for these 20 variables. (This is not entirely a happenstance, no doubt; the principal axis orientation in which MDPREF dimensions emerge is certainly more likely than a purely random orientation to yield interpretable structure.)

In the interest of grouping the dimensions in a fashion enhancing interpretability, we did permute their order. Thus Figure 10 shows the plane defined by dimensions one and three. Dimension one can be seen, from the fact that all variables have positive projections on that dimension, to be a “consensus” dimension, reflecting whatever factor is most nearly shared in common by all sites  $\times$  time periods. Figure 11 shows the 88 seaworm species in the same plane. The projections of the seaworms onto the dimension one axis would probably correspond very closely to the mean value of the twenty variables (i.e., with the mean of the log of the frequencies +1). This dimension could be interpreted, then, as overall “abundance” of the species, and the loading of a variable on that dimension simply indicates the extent to which that variable reflects this overall “abundance.” (As in factor analysis, the size of that loading can be viewed as a direct measure of the correlation of that variable with this first dimension.) Except that sites 1 and 2 seem to have very slightly lower weights on this dimension than do sites 3, 4 and 5, however, there seems to be nothing “interpretable” about this dimension vis à vis these

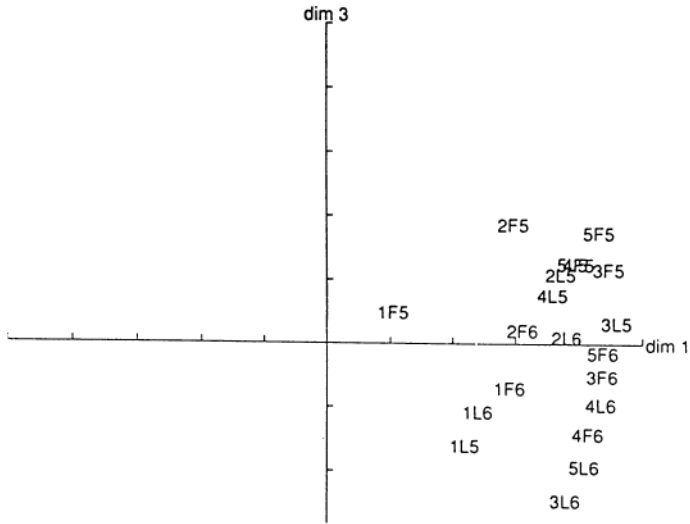


Figure 10. Termini of vectors projected into one-three plane for 20 site  $\times$  time period variables for unrotated MDPREF solution for Fresi *et al.* data.

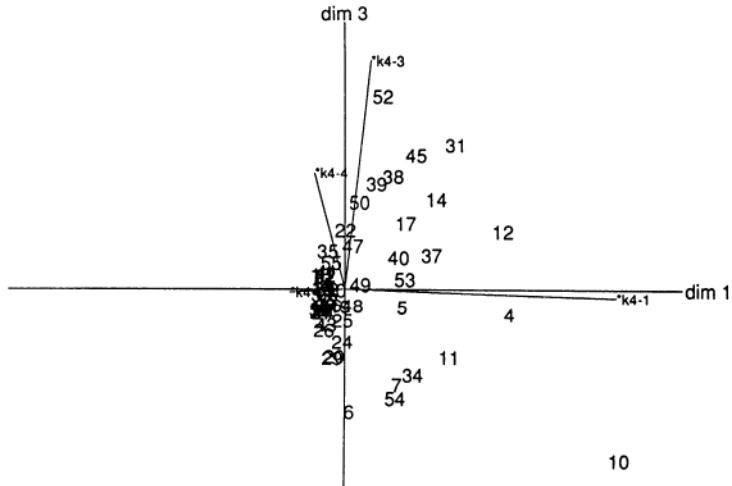


Figure 11. One-three plane of unrotated MDPREF stimulus (species) space for Fresi *et al.* data. Four vectors show result from mapping of dimensions from four dimensional KYST-2A solution into unrotated MDPREF space.

site  $\times$  time variables. Dimension three is more interesting, however. Note that almost all the variables involving the year 1975 (those whose code ends with "5") weight positively on that dimension, while those involving 1976 tend to exhibit negative weights.

In fact almost all the variables with a final "5" are in the upper right quadrant, and almost all those with a final "6" in the lower right quadrant. The most glaring exception is "1L5" (site 1, in July 1975) which appears just below "1L6" in the lower right hand quadrant. We have no definite explanation for this anomaly, although a partial explanation may be that there is something special about site 1 as a whole on this dimension. We note that, in general, the variables involving site 1 for a given time period seem to have systematically lower values on this dimension than do those for the other four sites. For example, 1F5 has a much lower value than do 2F5, 3F5, 4F5 and 5F5 all of which are at the extreme positive end of dimension 3, while 1F5 is almost at the zero point. Whatever dimension three corresponds to in its effect on the 88 species of seaworms, it is a factor that was positive (tended to increase the abundance of those species at the positive pole of that dimension) in 1975, and negative in 1976. A more explicitly descriptive way of stating the same thing is that those species at the positive end tended to be relatively more abundant in 1975, those at the negative end to be relatively more so in 1976.

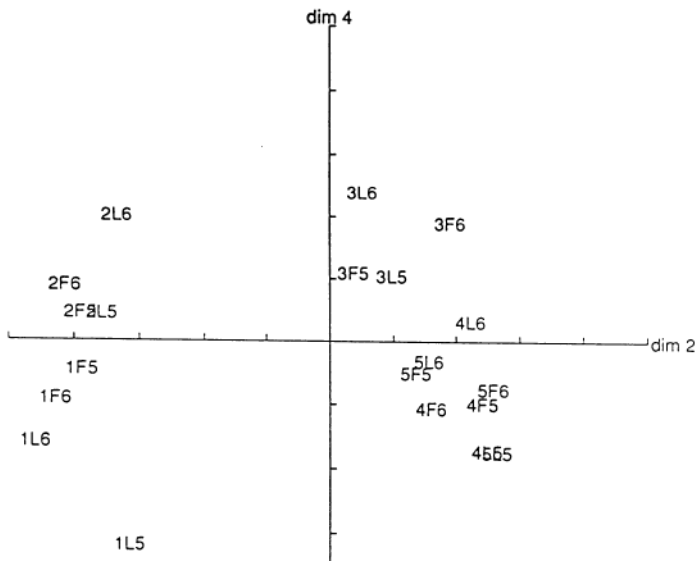


Figure 12. Termini of 20 variable vectors projected into two-four plane of unrotated MDPREF solution for Fresi *et al.* data.

We now shift to the remaining plane of this four-dimensional MDPREF solution, shown, for the sites  $\times$  times, in Figure 12. This is the plane defined by dimensions two and four. This plane distinguishes among the five sites to a remarkable degree. (It is dubious that a technique such as discriminant analysis, specifically geared to doing this, could do a significantly better job of separating these five groups.) As it is, we see that dimension two makes the most clearcut separation; that between sites 1 and 2 at the left (negative) end and sites 3, 4 and 5 at the right (positive) end. Then dimension four separates site 1 from 2 on the one hand, and site 3 from an amalgam of sites 4 and 5 on the other, so that site 1, 2, 3 and (4,5) wind up following neatly in a clockwise fashion (more-or-less) in the lower left, upper left, upper right and lower right quadrants, respectively. A map of the harbor of Ischia is given in the Fresi *et al.* paper. One can see from inspection of this map the reason why sites 1 and 2, located in open sea and separated by the harbor entrance from the other three sites, might be so clearly distinguished from those other sites, both in this representation and in the INDESCAL/SINDESCAL (source space) representation. This map also suggests some hypotheses as to why sites 4 and 5 may be so nearly indistinguishable. Site 3 is closer to the strait providing the harbor entrance, and separating sites 3, 4 and 5 from sites 1 and 2, so that it may be more affected by water flowing through that strait, while its ecology

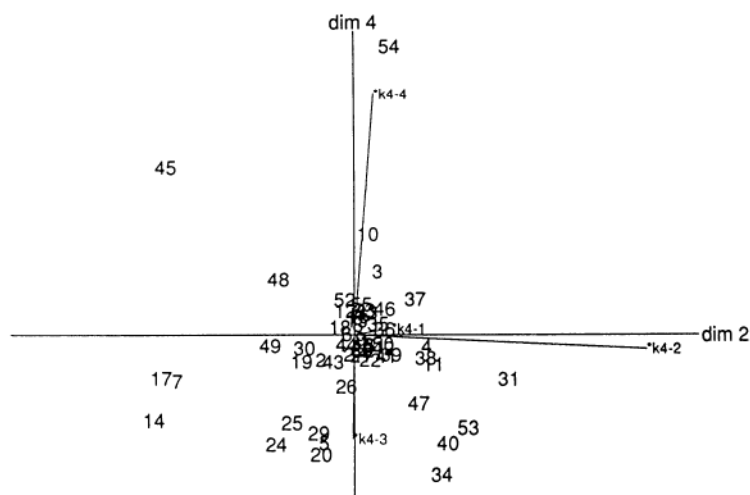


Figure 13. Two-four plane of unrotated MDPREF species space for Fresi *et al.* data. Vectors are as in Figure 11, except projected into two-four plane of unrotated MDPREF space.

may also more closely resemble that of 1 and 2 than does that of sites 4 and 5, which lie more distinctly in the harbor area. Figure 13 shows the dimension two-four plane of the stimulus (species) space, indicating how the seaworm species array themselves on these dimensions separating the various sites. (Again, it should be noted that overall frequency of the species has not been normalized here.)

It might be noted, by comparing Figures 1 and 2 to Figures 11 and 13, that the dimensions emerging from the KYST-2A analysis of the "Overall" dissimilarity matrix are essentially the same as those (for the seaworm species) in the unrotated MDPREF analyses. This is true despite the fact that the KYST-2A analysis omitted 33 of the 88 species, and also despite the marked difference in types of analysis. KYST-2A is a nonmetric technique aimed at accounting for rank orders of these derived dissimilarities, while MDPREF is a metric technique aimed at accounting for the values of the 88 species on the 20 site  $\times$  time variables. This congruence of the dimensions in these two analyses is shown directly by using PREFMAP-3, in a manner essentially identical to that described in section II.C, to "map" the dimension from the four dimensional KYST-2A solution into this MDPREF species space. The four vectors representing these four KYST dimensions (k4-1, k4-2, k4-3 and k4-4), respectively correspond very closely, as can be seen, to the corresponding dimensions (one through four, respectively) of the MDPREF solutions. The VAF's (or squared multiple correlations) were: .989, .991, .806 and .854 respectively. It is not unusual, however, for these two quite different analyses to produce highly comparable results. The reasons for this probably are twofold:

(1) The theoretically *nonmetric* KYST analysis is, in fact, essentially equivalent to a *metric* one, since the function relating input dissimilarities (distances) to recovered distances is almost perfectly linear and, in fact, goes very nearly through the origin, indicating that the input distances are very nearly ratio scale estimates of the derived distances. It should be emphasized, as spelled out in more detail below, that this might not have happened!

(2) The KYST-2A solution is rotated to principal components orientation, while the MDPREF solution is essentially a principal components solution.

The only seemingly important difference between these two solutions vis à vis the “worm” stimuli is in the scaling of these dimensions. Even this is not of any real significance however. It merely reflects the fact that, in MDPREF the stimulus (seaworm species) space is arbitrarily scaled to unit variance on all dimensions (and zero covariance – i.e., a “spherical” distribution), while the differential VAF (Variance accounted for) is absorbed in the vectors, while in KYST the differential VAF is reflected in the scaling of the stimulus (worm) dimensions. Thus, in *this* case at least, the simple metric MDPREF analysis has recovered essentially the same structure for the seaworm species as did the more complex and sophisticated KYST-2A procedure, while MDPREF has also extracted information about the “subjects” (sites  $\times$  times) in the form of the 20 vector locations, such that projection of stimulus points onto subject vectors yields approximations to the original dominance data.

It should be stressed, however, that this simple relationship between these two types of analysis will not always be exhibited. Particularly in the case of strong nonlinearities in the data, KYST-2A can yield a lower dimensional, more parsimonious representation of the stimuli (or other objects) than MDPREF (or other principal components/factor analytic type models and methods).

#### *Rotation of the MDPREF Solution to Congruence with SINDSCAL*

As mentioned, MDPREF does not yield unique dimensions, but rather is subject to rotational indeterminacies. In fact, more generally, a linear transformation of the stimulus space can be effected, as long as the appropriate companion transformation, given by the “inverse adjoint” transformation, is applied to the subject vectors. However, we shall restrict ourselves in the present case to orthogonal transformations, with possible overall dilations, or scale transformations. Since the inverse adjoint of an orthogonal transformation is the same orthogonal transformation, this leads to a particularly simple form (which has other advantages as well). Since the stimulus spaces in both MDPREF and SINDSCAL are scaled to have equal variances of projections of stimuli (species) on coordinate axes, restricting the class of transformations to be orthogonal seems appropriate in this case.

Figure 14 shows the dimension one versus two plane of the transformed species space superimposed on the same planes of the SINDSCAL space. In this representation the

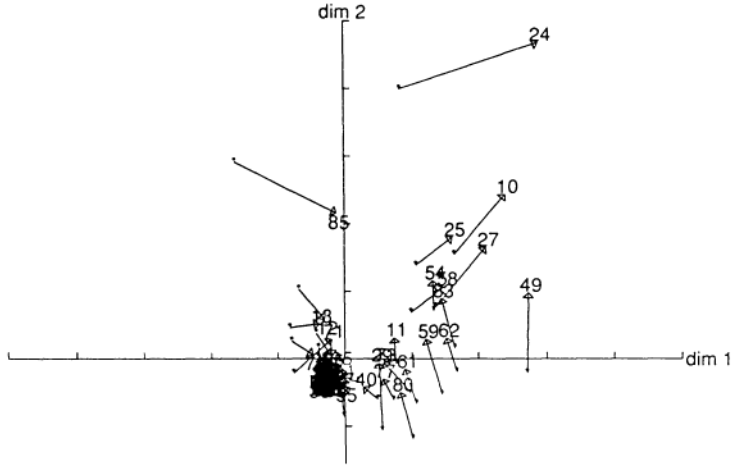


Figure 14. One-two plane of the MDPREF species space rotated to optimal congruence with one-two plane of SINDSCAL species space (two solutions superimposed, with "arrows" connecting corresponding points).

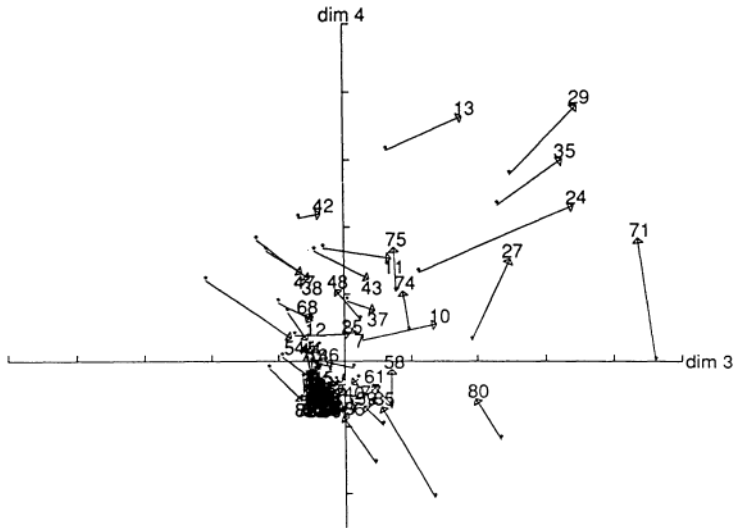


Figure 15. Same as Figure 14, but three-four plane of rotated MDPREF species space superimposed on same plane of SINDSCAL solution.

two points representing the same species are connected with one arrow. The terminus (arrowhead) of the arrow shows the position of the species point in the SINDSCAL representation, while the origin (shown by an asterisk) shows the point in the MDSCAL representation after rotation to optimal congruence with the SINDSCAL representation. In this case, the SINDSCAL configuration provides the "target," and the MDPREF



solution is rotated to best congruence in a certain least squares sense (specifically, so that the sum of squares of the arrow length is minimized). The specific procedure used was a variant of one originally proposed by Cliff (1966), which is closely related to the “orthogonal procrustes” approach described by Gower in section 9.1 of his paper in this volume. Figure 15 shows a similar plot for the dimension three-four plane. (It should be kept in mind that the dimensions referred to here are those from the SINDSCAL solution, so the one-two plane should be taken as corresponding to those dimensions from the SINDSCAL analysis, *not* from the MDPREF solution first described. Since the MDPREF coordinate system has been completely transformed in this process, there is no necessary one-one correspondence with those dimensions.) Figures 16 and 17 show the rotated MDPREF solution in those same two planar projections, but this time with the

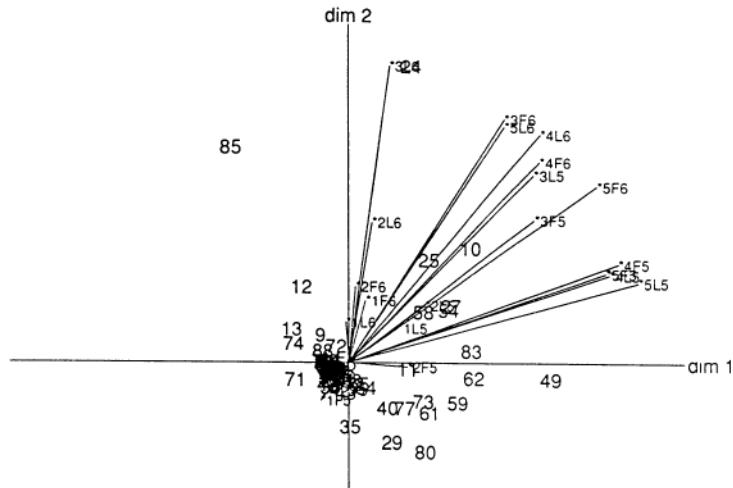


Figure 16. One-two plane of MDPREF space rotated to optimal congruence with SINDSCAL solution. Site  $\times$  time period vectors as well as 88 seaworm species points are shown, in joint space representation.

(rotated) vectors shown simultaneously (and, in fact, with lines connecting them to the origin to make their vectorial nature more evident). In Figure 16, showing the dimension one-two plane of this rotated MDPREF joint representation, we see that all the vectors for sites 3, 4 and 5 when projected into that plane have substantial lengths, while those for sites 1 and 2 have lengths, when projected into this plane, that are very near zero. Thus these two dimensions are accounting for virtually all the reliable variance for sites 3, 4 and 5, and essentially none for site 1 and 2. This is consistent with the fact that, in the SINDSCAL representation, the derived dissimilarity matrices

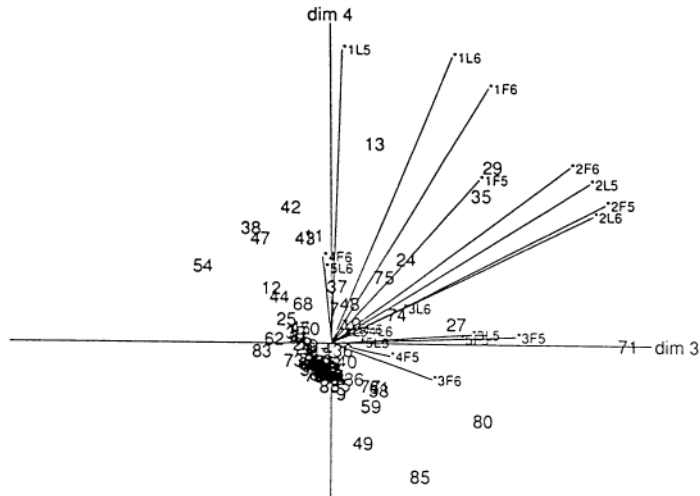


Figure 17. Same as Figure 16, except that three-four plane of joint MDPREF representation, after rotation to optimal congruence with SINDSCAL solution, is plotted showing both species points and variable vectors.

for sites 3, 4 and 5 had high, clearly non-zero, weights on the corresponding INDSCAL dimensions, while those derived for sites 1 and 2 had near zero weights. The opposite pattern shows up in the plane for dimension three and four of this rotated MDPREF representation shown in Figure 17; the lengths of the vectors for sites 1 and 2 projected in this plane are substantial, while those for sites 3, 4 and 5 are near zero. Again, this is consistent with the INDSCAL results. We can also see in this three, four plane a clear separation between sites 1 and 2, with site 1 having higher weights on dimension four than three, and site 2 the opposite pattern. In the one, two plane we can see some, but not as clear, differentiation of site 3 from site 4 and 5. These three sites are much more “mixed up” in this representation than in others we have seen. There is some hint of the differentiation based on year (1975 versus 1976) in the vectors for sites 3, 4 and 5 in this plane, however.

In a sense this representation provided by the MDPREF solution after rotation to optimal congruence with SINDSCAL provides the most cogent and succinct representation of all for these data. What it shows is an overall four dimensional representation, but one neatly partitioned into two two-dimensional subspaces. The one-two subspace seems to account for most of the variance in the variables related to sites 3, 4 and 5 while the three-four subspace accounts for most of that in variables related to sites 1 and 2. A further “nice” aspect of this representation is that almost all the vectors

lie in the positive quadrants of these two planes, so the weights of these four dimensions are almost all positive or zero. This suggests that the use we have made of SINDSCAL in this case may provide a very effective basis for *rotation* of an MDPREF type representation to a special kind of generalized "simple structure."

It now only remains for ecologists and biologists to "interpret" the dimensions in terms of their effects on the seaworm species. We happily defer that privilege to these experts. To aid such experts in this creative endeavor, however, we provide a final table, Table 6, in which the coordinates for the 88 seaworm species on the dimensions of the four different configurations discussed in this paper are presented.

*Acknowledgments.* Invaluable help in conducting the data analyses reported and other technical help in preparing this paper were provided by Rhoda T. Iosso and Barbara B. Hollister. Thanks are also due to Martina Bose and to Karen Golday for word processing and other technical assistance. Finally, I am greatly indebted to Pierre Legendre and Joseph B. Kruskal, plus two anonymous reviewers, for careful readings of the paper at various stages of its development, leading to enormous improvements in its content.

## REFERENCES

- Arabie, P., & J. D. Carroll. 1980a. MAPCLUS: A mathematical programming approach to fitting the ADCLUS model. *Psychometrika* 45: 211-35.
- Arabie, P., & J. D. Carroll. 1980b. How to use MAPCLUS: A computer program for fitting the ADCLUS model. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Arabie, P., J. D. Carroll & W. S. DeSarbo. 1987. Three-Way multidimensional scaling and related techniques. Sage University Paper. In press.
- Bloxom, B. 1968. Individual differences in multidimensional scaling. *Research Bulletin*, 68-45. Educational Testing Service, Princeton, NJ.
- Carroll, J. D. 1965. Parametric mapping of similarity data as a scaling method. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Carroll, J. D. 1968. Generalization of canonical correlation analysis to three or more sets of variables. *Proceedings of the 76th Annual Convention of the American Psychological Association* 3: 227-228.
- Carroll, J. D. 1969. Polynomial factor analysis. *Proceedings of the 77th Annual Convention of the American Psychological Association* 4: 103-104.

Table 6. Coordinates of points for 88 species of swarms on dimensions of four different multidimensional representations described in text.

Numerical Code	Unrotated				Rotated				SINDSCAL COORDINATES				Transformed				Abundance Measure ABUND = $\sum \log (j + 1)$
	MP1	MP2	MP3	MP4	RM1	RM2	RM3	RM4	SI1	SI2	SI3	SI4	TS1	TS2	TS3	TS4	
1	-0.55	.014	-.022	-.011	-.028	-.016	-.043	.009	-.046	-.041	-.056	.033	.010	.014	.014	.027	2.07
2	-.054	.023	.004	-.022	-.028	-.016	-.043	-.009	-.040	-.044	-.052	-.048	.020	.019	.025	.031	1.38
3	-.061	.020	-.018	.016	-.029	-.016	-.019	.040	-.053	-.042	-.052	-.053	.021	.041	.050	.054	.69
4	-.057	.014	-.002	.071	-.040	-.010	-.034	.030	-.059	-.031	-.032	.064	.008	.036	.050	.024	1.09
5	-.059	.019	-.003	.061	-.057	.003	.001	.058	-.057	-.035	-.039	.065	.015	.051	.069	.037	.69
6	-.060	.016	-.027	.001	-.052	-.000	-.006	.056	-.053	-.043	-.053	.045	.013	.025	.030	.041	1.09
7	-.029	-.070	0	-.047	-.037	-.011	-.042	.017	-.052	-.049	-.020	-.048	.002	.002	.014	.018	7.49
8	-.060	.016	-.011	.034	-.025	-.047	.000	-.058	-.054	-.040	-.053	-.048	.020	.044	.049	.062	.69
9	-.030	.039	-.024	.122	-.045	-.007	-.021	.038	-.031	.009	-.029	.074	.013	.028	.020	.006	2.77
10	.335	.133	-.053	-.021	-.060	.053	.012	.085	.340	.346	.151	-.061	.007	.007	.004	.002	52.78
11	.113	-.064	-.037	-.203	.233	.219	.028	-.033	.104	.045	.077	-.167	.006	.004	.005	.009	28.65
12	.003	-.019	-.256	.002	-.109	-.012	-.034	-.183	.043	.084	-.052	-.066	.002	.015	.003	.015	9.97
13	.102	-.344	-.200	-.088	-.097	.146	.108	-.094	-.055	.074	.194	-.396	.000	.004	-.008	.015	31.82
14	-.036	.001	-.031	.019	-.116	.065	.069	-.340	.035	-.023	-.048	.003	.006	.009	.007	.017	5.12
15	-.057	.009	.004	.022	-.038	.007	-.020	.009	-.053	-.046	-.029	.055	.010	.018	.041	.025	1.38
16	-.061	.009	.005	.034	-.036	-.021	-.014	.032	-.056	-.044	-.045	.048	.017	.038	.060	.062	.69
17	-.059	.002	-.008	.008	-.043	-.019	-.009	.040	-.055	-.044	-.047	.034	.009	.018	.028	.041	1.38
18	-.061	.009	.005	.034	-.039	-.021	-.024	.018	-.056	-.044	-.045	.045	.017	.038	.060	.062	.69
19	-.057	.006	-.034	-.006	-.043	-.019	-.009	.040	-.052	-.042	-.054	.025	.008	.016	.018	.036	1.79
20	-.059	.013	-.032	-.008	-.038	-.010	-.043	.003	-.052	-.044	-.054	.039	.011	.019	.023	.037	1.38
21	-.061	.009	.005	.034	-.035	-.012	-.048	.008	-.056	-.044	-.045	.048	.017	.038	.060	.062	.69
22	-.061	.020	-.018	.016	-.043	-.019	-.009	.040	-.053	-.042	-.052	.053	.021	.041	.050	.054	.69
23	-.060	.019	.005	.046	-.040	-.010	-.034	.030	-.054	-.042	-.038	-.066	.020	.041	.070	.036	.69
24	.559	.021	-.352	.193	-.043	-.011	-.009	.052	.409	.666	.371	-.252	.005	.008	-.005	.004	84.99
25	.211	.146	-.143	-.055	.117	.567	.121	-.143	.229	.256	.008	-.044	.008	.008	.002	.003	37.03
26	-.061	.009	.005	.034	-.150	.197	.081	-.041	-.056	-.044	-.045	.048	.017	.038	.060	.062	.69
27	.322	-.023	.117	.045	-.043	-.019	-.009	.040	.300	.235	.270	-.166	.006	.005	.006	.004	55.29
28	-.058	.011	-.037	-.017	.194	.109	.207	-.033	-.052	-.044	-.035	.035	.009	.016	.020	.034	1.60
29	.185	-.390	.183	-.162	-.034	-.012	-.053	.000	.073	-.009	.376	-.413	.003	.001	.009	.010	46.32
30	-.057	.008	-.014	.046	.078	-.151	.270	-.301	-.058	-.031	.041	.045	.007	.027	.033	.033	1.38
31	-.050	0	.010	.027	-.053	-.001	-.012	.038	-.050	-.046	-.014	-.048	.006	.009	.026	.015	2.77
32	-.058	.015	.010	.048	-.036	-.021	-.001	.029	-.053	-.043	-.031	.065	.013	.025	.050	.023	1.09
33	-.054	.002	.003	.035	-.043	-.013	-.002	.051	-.053	-.042	-.028	.044	.007	.013	.028	.022	2.07
34	-.060	.019	.005	.046	-.042	-.015	-.004	.034	-.054	-.042	-.038	-.066	.020	.041	.070	.036	.69
35	.121	-.375	.133	-.083	-.043	-.011	-.009	.052	-.012	-.016	.353	-.327	.001	.001	.012	.011	36.08
36	-.055	-.033	.029	.014	-.002	-.121	.250	-.252	.060	-.050	-.031	.002	.003	.008	.023	.037	2.39
37	-.016	-.107	-.036	-.050	-.040	-.046	.014	.011	-.059	-.031	.050	-.082	.000	.003	.013	.016	10.44
38	-.027	-.069	-.142	.225	-.044	-.028	.005	-.097	-.038	-.046	-.057	-.133	.003	.002	.003	.022	9.83
39	-.057	.006	-.049	-.039	.000	-.033	-.143	.195	-.051	-.045	-.055	.023	.007	.011	.014	.030	2.19
40	-.004	.024	.120	-.047	-.030	-.014	-.065	.010	.039	-.060	.010	-.039	.011	.001	.010	.005	9.21
41	-.050	.014	.002	.046	.069	-.085	.021	.030	.051	-.029	.016	.070	.009	.023	.039	.011	1.79
42	-.011	-.155	-.112	-.206	-.041	-.005	-.004	.045	-.037	-.034	-.037	-.239	.001	.002	.003	.025	13.19
43	-.010	-.125	-.066	-.166	-.022	-.052	-.073	-.231	-.043	-.050	.041	-.133	.001	.001	.010	.017	12.73
44	-.049	-.021	-.087	-.098	-.010	-.056	-.048	-.178	-.050	-.044	-.061	-.034	.004	.006	.005	.028	4.38

45	-.052	.006	-.051	-.036	-.025	-.017	-.092	-.078	-.052	-.044	-.041	-.014	.005	.008	-.014	.024	3.09
46	-.052	.009	-.044	-.025	-.034	-.012	-.057	-.025	-.046	-.041	-.052	.025	-.007	.010	-.012	-.023	2.77
47	-.032	-.074	-.146	-.185	-.030	-.009	-.056	-.008	-.048	-.041	-.069	-.142	.002	.003	.002	.027	8.54
48	-.025	-.101	-.002	-.024	-.022	-.018	-.126	-.178	-.051	-.038	-.015	-.116	.002	.004	-.009	.026	7.87
49	.226	.290	.296	.084	-.046	-.041	.026	-.067	.388	.139	-.004	.091	.012	.005	.002	.000	36.62
50	.061	.016	-.015	.028	.385	.030	.049	.168	.055	.040	-.054	.047	.018	.044	.047	.063	.69
51	-.057	.008	-.044	-.029	-.045	-.008	-.025	.034	-.051	-.045	-.055	.028	-.008	.013	.016	.032	1.94
52	-.053	.019	-.010	.042	-.031	-.013	-.060	-.009	-.047	-.033	-.050	.047	.011	.021	.020	.024	1.79
53	-.058	.032	-.015	.046	-.043	-.000	-.015	.043	-.052	-.027	-.053	.068	.023	.063	.049	.033	.69
54	-.136	.162	-.179	-.263	-.044	.003	.024	.052	.187	.165	-.087	-.037	.009	.008	.000	.004	27.90
55	-.041	.040	.077	.021	.189	.102	.225	-.130	.003	-.058	-.038	.069	.021	.004	-.016	.007	2.99
56	-.042	.052	-.010	.011	.011	-.049	.010	.070	-.027	-.013	-.058	.067	.012	.018	.009	.007	3.17
57	-.061	.020	-.018	.016	-.010	-.000	-.040	.044	-.053	-.042	-.053	.053	.021	.041	.049	.054	.69
58	.175	.112	.069	-.067	-.040	-.010	-.034	.030	.213	.152	.076	.014	.009	.007	.005	.002	30.31
59	.095	.131	.230	-.043	.140	.097	.075	.079	.172	.042	.030	.075	.010	.004	.005	.000	23.04
60	-.052	.040	.027	.019	.206	-.075	.062	.106	-.026	.040	-.050	.068	.030	.015	-.026	.016	1.38
61	.060	.064	.215	-.037	-.011	-.027	.017	.056	.127	-.022	.044	.039	.010	.002	.007	.002	18.71
62	.106	.173	.064	-.204	.151	-.094	.076	.073	.216	.044	-.046	.038	.013	.005	.001	.002	21.60
63	-.059	.015	.018	-.048	-.238	-.028	-.102	-.007	-.054	-.044	-.028	.069	.020	.038	.085	.031	.69
64	-.043	.055	.035	-.040	-.040	-.019	-.000	.054	-.007	-.043	-.059	.066	.014	.006	.006	.005	4.18
65	-.046	.050	.028	-.024	.028	-.047	-.043	.034	-.016	-.043	-.057	.066	.015	.008	.008	.007	3.40
66	-.061	.020	-.018	.016	.014	-.040	-.039	.037	-.053	-.042	-.053	.053	.021	.041	-.049	.054	.69
67	-.059	.015	.018	-.048	-.040	-.010	-.034	.030	-.054	-.044	-.028	.069	.020	.038	.085	.031	.69
68	-.044	-.046	-.075	-.051	-.040	-.019	-.000	.054	-.058	-.031	-.065	-.075	.002	.008	.004	.033	4.90
69	-.059	.015	.018	-.048	-.048	-.007	-.052	-.066	-.054	-.044	-.028	.069	.020	.038	.085	.031	.69
70	-.051	-.020	-.050	-.018	-.040	-.019	-.000	.054	-.057	-.035	-.060	-.025	.003	.011	-.008	.036	3.17
71	.143	-.369	.274	.324	-.047	-.008	-.041	-.025	-.029	.042	.473	.199	.001	.003	.015	.008	35.64
72	-.034	.053	-.034	-.048	-.108	-.031	.500	-.000	-.028	.007	-.053	.069	.011	.022	.009	.006	3.46
73	.010	.118	.087	-.130	-.030	.031	-.032	.054	.082	-.008	-.067	.067	.013	.005	.001	.002	10.89
74	.010	-.151	-.035	.105	-.140	-.072	-.068	.029	-.068	.009	.095	-.114	-.000	.005	.013	.014	13.68
75	-.026	-.166	.008	-.020	-.112	.036	.105	-.049	-.030	-.021	.081	-.185	.002	.002	-.009	.016	17.23
76	-.054	.037	.020	.024	-.041	-.032	.084	-.113	-.032	-.047	-.050	.068	.032	.021	-.033	.020	1.09
77	.024	.049	.177	-.016	-.018	-.023	.017	.055	.080	-.043	.038	.065	.011	.002	.009	.001	13.03
78	-.057	.014	-.002	.071	.103	.086	.061	.072	-.059	-.031	-.032	.064	.008	.036	.050	.024	1.09
79	-.047	.035	.025	-.020	-.057	.003	.001	.058	-.021	-.049	-.052	.053	.018	.008	.014	.015	2.48
80	.073	-.024	.395	.065	.004	-.040	.031	.030	.117	-.071	.213	.063	.009	.000	.015	.001	19.26
81	-.053	.037	.030	.017	.143	-.169	.253	.128	-.030	-.047	-.050	.067	.027	.017	-.026	.017	1.38
82	-.058	.032	-.015	.046	-.011	-.030	.015	.054	-.052	-.027	-.053	.068	.023	.063	.049	.033	.69
83	.119	.213	.019	-.175	-.044	.003	-.024	.052	.205	.127	-.070	.067	.011	.008	.000	.001	23.58
84	-.059	.015	.018	-.048	-.234	.022	-.124	.013	-.054	-.044	-.028	.069	.020	.038	.085	.031	.69
85	.099	.064	-.228	.558	-.040	-.019	-.000	.054	-.017	.306	.059	.076	.002	.018	.007	.000	20.00
86	-.034	.008	.053	.057	-.233	.412	.145	.224	-.033	-.042	-.010	-.065	.007	.008	-.019	.005	4.85
87	-.061	.020	-.018	.016	-.021	-.024	.033	.060	-.053	-.042	-.053	.053	.021	.041	-.049	.054	.69
88	-.052	.036	-.016	.085	-.040	-.010	-.034	.030	-.051	-.014	-.049	.070	.012	.041	-.027	.015	1.38

- Carroll, J. D. 1972. Individual differences and multidimensional scaling, p. 105-155. In R. N. Shepard, A. K. Romney & S. Nerlove (Eds.) *Multidimensional scaling: Theory and applications in the behavioral sciences* (Vol. 1). Seminar Press, New York and London.
- Carroll, J. D. 1980. Models and methods for multidimensional analysis of preferential choice (or other dominance) data, p. 234-289. In E. D. Lantermann & H. Feger (Eds.) *Similarity and choice*. Hans Huber Publishers, Bern, Stuttgart and Vienna.
- Carroll, J. D., & P. Arabie. 1980. Multidimensional Scaling. *Annual Review of Psychology* 31: 607-49.
- Carroll, J. D., & P. Arabie. 1982. How to use INDCLUS, A computer program for fitting the individual differences generalization of the ADCLUS model. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Carroll, J. D., & P. Arabie. 1983. INDCLUS: An individual differences generalization of the ADCLUS model and the MAPCLUS algorithm. *Psychometrika* 48: 157-169.
- Carroll, J. D., & J. J. Chang. 1964a. A general index of nonlinear correlation and its application to the problem of relating physical and psychological dimensions. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Carroll, J. D., & J. J. Chang. 1964b. Nonparametric multidimensional analysis of paired-comparisons data. Paper presented at the joint meeting of the Psychometric and Psychonomic Societies, Niagara Falls, October 1964. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Carroll, J. D., & J. J. Chang. 1967. Relating preference data to multidimensional scaling solutions via a generalization of Coombs' unfolding model. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Carroll, J. D., & J. J. Chang. 1969. A new method for dealing with individual differences in multidimensional scaling. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Carroll, J. D., & J. J. Chang. 1970. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika* 35: 283-319.
- Carroll, J. D., & J. J. Chang. 1972. IDIOSCAL (Individual Differences In Orientation SCALing): A generalization of INDSCAL allowing *IDIO*syncratic reference systems as well as an analytic approximation to INDSCAL. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Carroll, J. D., & J. J. Chang. 1972. SIMULES: *SIMU*ltaneous Linear Equation Scaling. *Proceedings of the 80th Annual Convention of the American Psychological Association* 7: 11-12.
- Carroll, J. D., & S. Pruzansky. 1980. Discrete and hybrid scaling models, p. 108-139. In E. D. Lantermann & H. Feger (Eds.) *Similarity and Choice*. Hans Huber Publishers, Bern, Stuttgart, Vienna.

- Carroll, J. D., & S. Pruzansky. 1984. The CANDECOMP-CANDELINC family of models and methods for multidimensional data analysis, p. 372-402. In H. G. Law, W. Snyder, J. Hattie, & R. P. McDonald (Eds.) *Research methods for multimode data analysis*. Praeger, New York.
- Carroll, J. D., & S. Pruzansky. 1986. Discrete and hybrid models for proximity data, p. 47-59. In W. Gaul and M. Schader (Eds.) *Proceedings of the German Classification Society Annual Meeting, (Classification as a tool of research)*, Elsevier Science Publishers B.V. (North Holland), Amsterdam.
- Carroll, J. D., S. Pruzansky & J. B. Kruskal. 1980. CANDELINC: A general approach to multidimensional analysis of many-way arrays with linear constraints on parameters. *Psychometrika* 45: 3-24.
- Carroll, J. D., & M. Wish. 1974a. Models and methods for three-way multidimensional scaling, p. 57-105. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.) *Contemporary developments in mathematical psychology* (Vol. II). W. H. Freeman, San Francisco.
- Carroll, J. D., & M. Wish. 1974b. Multidimensional perceptual models and measurement methods, p. 391-447. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. II). Academic Press, New York.
- Chang, J. J. 1968. How to use PARAMAP: A computer program which performs parametric mapping. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Chang, J. J. 1971. Multidimensional scaling program library. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Chang, J. J., & J. D. Carroll. 1968a. How to use MDPREF: A computer program for multidimensional analysis of preference data. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Chang, J. J., & J. D. Carroll. 1968b. How to use PROFIT: A computer program for property fitting by optimizing nonlinear or linear correlation. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Chang, J. J., & J. D. Carroll. 1969. How to use INDSCAL: A computer program for canonical decomposition of N-way tables and individual differences in multidimensional scaling. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Chang, J. J., & J. D. Carroll. 1972a. How to use IDIOSCAL: A computer program for individual differences in orientation scaling. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Chang, J. J., & J. D. Carroll. 1972b. How to use PREFMAP and PREFMAP 2: Programs which relate preference data to multidimensional scaling solution. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Chang, J. J., & J. D. Carroll. 1972c. How to use SIMULES: A computer program which does simultaneous linear equation scaling. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.

- Cliff, N. 1966. Orthogonal rotation to congruence. *Psychometrika* 31: 33-42.
- Coombs, C. H. 1964. *A Theory of Data*. Wiley, New York.
- Eckart, C., & G. Young. 1936. The approximation of one matrix by another of lower rank. *Psychometrika* 1: 211-218.
- Fresi, E., R. Colognola, M. C. Gambi, A. Giangrande & M. Scardi. 1983. Ricerche sui popolamenti bentonici di substrato duro del porto di Ischia. Infralitorale fotofilo: Policheti. *Cahiers de Biologie Marine* XXIV: 1-19.
- Gower, J. C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* 53: 325-38.
- Harshman, R. A. 1970. *Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multimodal factor analysis*. UCLA Work Pap. Phonetics 16, 84pp.
- Harshman, R. A. 1972. Determination and proof of minimum uniqueness conditions for PARAFAC II. UCLA: Work. Pap. Phonetics 22: 111-17.
- Heiser, W. J. 1981. *Unfolding analysis of proximity data*. Doctoral Dissertation, University of Leiden, Leiden, The Netherlands.
- Horan, C. B. 1969. Multidimensional scaling: Combining observations when individuals have different perceptual structures. *Psychometrika* 34: 139-65.
- Johnson, S. C. 1967. Hierarchical clustering schemes. *Psychometrika* 32: 241-54.
- Kruskal, J. B. 1964a. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29: 1-27.
- Kruskal, J. B. 1964b. Nonmetric multidimensional scaling: A numerical method. *Psychometrika* 29: 115-29.
- Kruskal, J. B. 1965. Analysis of factorial experiments by estimating monotone transformations of the data. *Journal of the Royal Statistical Society Series B* 27: 251-263.
- Kruskal, J. B. 1977. Multidimensional scaling and other methods for discovering structure, p. 296-339. In K. Enslein, Ralston, & Wilf (Eds.) *Statistical methods for digital computers (Vol. III of Mathematical Methods for Digital Computers)*. Wiley, New York.
- Kruskal, J. B., & F. Carmone. 1969. How to use M-D-SCAL (version 5M) and other useful information. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Kruskal, J. B., & F. Carmone. 1968. Use and theory of MONANOVA: A program to analyze factorial experiments by estimating monotone transformation of the data. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Kruskal, J. B., & M. Wish. 1978. *Multidimensional Scaling*. Sage, Beverly Hills.
- Kruskal, J. B., F. W. Young, & J. B. Seery. 1973. How to use KYST-2A: A very flexible program to do multidimensional scaling and unfolding. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Meulman, J., W. J. Heiser, & J. D. Carroll. 1986. PREFMAP-3 User's Guide. Unpublished manuscript, AT&T Bell Laboratories, Murray Hill, NJ.



- Meredith, W. 1964. Rotation to achieve factorial invariance. *Psychometrika* 29: 187-206.
- Pruzansky, S. 1975. How to use SINDSCAL: A computer program for individual differences in multidimensional scaling. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Ramsay, J. O. 1977. Maximum likelihood estimation in multidimensional scaling. *Psychometrika* 42: 241-66.
- Sands, R., & F. W. Young. 1980. Component models for three-way data: an alternating least squares algorithm with optimal scaling features. *Psychometrika* 45: 39-67.
- Schonemann, P. H. 1972. An algebraic solution for a class of subjective metrics models. *Psychometrika* 37: 441-51.
- Shepard, R. N. 1962. The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika* 27: 125-140.
- Shepard, R. N. 1962. The analysis of proximities: Multidimensional scaling with an unknown distance function. II. *Psychometrika* 27: 219-246.
- Shepard, R. N. 1980. Multidimensional scaling, tree-fitting and clustering. *Science* (210): 390-98.
- Shepard, R. N., & J. D. Carroll. 1966. Parametric representation of nonlinear data structures, p. 561-592. In P. R. Krishnaiah (Ed.) *Multivariate Analysis*. Academic Press, New York.
- Takane, Y., & J. D. Carroll. 1980. How to use MAXSCAL-4.1: A program to fit the simple (unweighted) Euclidean distance model to various types of nonmetric dissimilarity data by maximum likelihood. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Takane, Y., & J. D. Carroll. 1981. Nonmetric maximum likelihood multidimensional scaling from directional rankings of similarities. *Psychometrika* 46: 389-405.
- Takane, Y., F. W. Young, & J. de Leeuw. 1977. Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features. *Psychometrika* 42: 7-67.
- Torgerson, W. S. 1958. *Theory and Methods of Scaling*. Wiley, New York.
- Tucker, L. R. 1960. Intra-individual and inter-individual multidimensionality. In *Psychological Scaling: Theory and Applications*, ed. H. Gulliksen, S. Messick. Wiley, New York.
- Tucker, L. R. 1972. Relations between multidimensional scaling and three-mode factor analysis. *Psychometrika* 37: 3-27.
- Weinberg, S. L., & J. D. Carroll. 1986. Choosing the dimensionality of an INDSCAL-derived space by using a method of resampling. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April 16-20. Unpublished manuscript, ATT Bell Laboratories, Murray Hill, NJ.
- Wish, M., & J. D. Carroll. 1974. Applications of individual differences scaling to studies of human perception and judgment, p. 449-491. In E. C. Carterette & M. P. Friedman (Eds.) *Handbook of perception*, Chapter 13 (Vol. II). Academic Press,

New York.

Young, F. W. 1968. TORSCA-9: A FORTRAN IV program for nonmetric multidimensional scaling. *Behavioral Science* 13: 343-344.

Young, F. W., & W. S. Torgerson. 1967. TORSCA, a FORTRAN IV program for Shepard- Kruskal multidimensional scaling analysis. *Behavioral Science* 12: 498.