Made in United States of America
Reprinted from BRITTONIA
Vol. 27, No. 4, October–December, 1975
pp. 305–316

STATISTICAL CHARACTER WEIGHTING AND SIMILARITY STABILITY[1]

ROBERT P. ADAMS[2]

Abstract. Terpenoid data from seven species of *Juniperus* were used to examine: *1*) the effect of using different character weights upon the same set of OTU's; *2*) the effect of the organization of the initial data base by populations, individuals, and averages on statistical weights and the resulting classification; and *3*) the effect of the use of exemplars on statistical weights and similarity measures. Cophenetic correlation and numerical taxonomy, along with certain relationships well documented for these species, were used to examine these questions. Equal weighting of characters, advocated by most numerical taxonomists, gave the most distorted results. This was followed by the 1/coefficient of variation, then the weighting of $1 - S_w/S_t$, $\sqrt{F-1}$, and finally F and $F - 1$ weighting producing the highest fidelity to the known similarities. The use of either all individuals or random individuals appeared to be better than the use of population averages for some OTU's and individuals for other OTU's and much better than the use of population averages only. The use of exemplars has small effects, mostly causing the OTU's to cluster more loosely when exemplars were included in the computation of the statistical weights. The use of equal weighting is discussed and strongly discouraged in numerical taxonomy in favor of statistically derived weights.

Equal weighting of characters has been generally accepted in numerical taxonomy and has been actively advocated (Sokal & Sneath, 1963; Sneath & Sokal, 1973). This practice has become axiomatic to phenetics although recently controversy has developed. Colless (1971) concluded that certain kinds of "nonphylogenetic" (quotation marks his) weighting may be allowed and the methods still be phenetic. The early objections of numerical taxonomists to unequal weighting were due to the subjective *a priori* weights assigned to "primitive" or "conservative" characters. This point has now been adequately emphasized to classical taxonomists, and we should now re-examine objective methods for *a posteriori* weighting. Although several weighting methods have been proposed, I shall focus on those which are statistical in nature and particularly applicable to continuous characters.

The purpose of this paper is to investigate the following questions which apply directly or indirectly to statistical character weighting and similarity stability: *1*) Which statistical weighting produces the highest fidelity of known similarities for the same set of Operational Taxonomic Units (OTU's)? *2*) How does the organization of the initial data base by populations, individuals, and averages affect the statistical weights and the resulting classifications? *3*) How does the use of exemplars affect statistical weights and similarity measures?

Farris (1966) suggested that characters might be weighted according to the reciprocal of the sample variance of a random sample drawn from the population (for the construction of evolutionary trees). Although this would give more weight to those characters which vary little within the populations (or OTU's), the character still might not be of any value if it did not vary from one OTU to another. To overcome this difficulty, Flake, von Rudloff, and Turner (1969) proposed to weight characters by w, where: $w = 1 - S_w/S_t$. Here, S_w = average within OTU standard deviation (for a given character) and S_t = pooled sample or total standard deviation (for the same character). When $S_w > S_t$, set $w = 0$. Thus characters which vary

[1] In part supported by grants GB-24320 and GB-37315X from the National Science Foundation. Computer time was furnished by Colorado State University.

[2] Department of Botany and Plant Pathology, Colorado State University, Fort Collins, CO 80521.

little within OTU's and vary considerably among OTU's have lower S_w/S_t ratios and relatively high weights (w).

A similar weighting was used by Adams and Turner (1970) and Adams (1972*a*, 1972*b*) when the F values of the characters were used as weights and where the F ratio is defined as: variance among OTU's/variance within OTU's, as commonly used in analysis of variance (ANOVA). This weighting seems logical to taxonomists since if four species have leaf lengths which varied considerably within each species and differed little among the species, one would decide that leaf length was of little value in classification. The F statistic does essentially that very same job. A more intermediate weighting was used by Zanoni (1974), who used $\sqrt{F-1}$. To my knowledge the only comparison made of the weighting schemes and of the resulting stability of the similarity measures was mine (Adams, 1972*b*), which showed that in the analysis of the effects of misidentified compounds errors of up to 20% of the total weights might be tolerated.

The following weighting schemes will be compared: all characters equally weighted (1.0); $1/CV_i$, where CV_i is the average coefficient of variation within OTU's for character i (this weights characters according to variability within groups as proposed by Farris (1966)); $1-S_w/S_t$, the weighting used by Flake et al. (1969); $\sqrt{F-1}$, used by Zanoni (1974); F, used by Adams and Turner (1970) and Adams (1972*a,b*); and $F-1$ (to compare with F weighting).

When one uses statistically derived weighting coefficients, the organization of the data by population averages (*versus* individuals) may present difficulties. For instance, suppose one has ten individuals from each of 20 populations of OTU A, ten individuals of OTU B, and one individual of OTU C. Should one consider all 200 individuals of OTU A as one population set and the 10 individuals of OTU B as another sample set or use the average of each of the 20 populations of OTU A to obtain the average of 20 "specimens" of OTU A to compare with the average of 10 individuals of OTU B? When statistical weights are used, the F's may vary tremendously according to manner the data are structured for analysis. Ten OTU's will be compared in which weights are based on the variation among: *1*) individuals within each OTU; *2*) population averages within OTU's A and P and among individuals within the other OTU's; *3*) randomly chosen individuals from populations within OTU's A and P and among individuals within the other OTU's; and *4*) population averages within each of the ten OTU's.

The adding of individuals or population averages to numerical taxonomy is often of value when individuals or OTU's of uncertain affinities are analyzed. The usual procedure is to analyze the OTU's at hand and then to add a few exemplars in a subsequent run to see where the exemplars cluster. When no character weighting is used this appears to work satisfactorily. The use of exemplars, when the weights are determined by computing estimates of variability using a set of OTU's that does not include the variability of the exemplars, may or may not be valid. This question will be examined by comparisons of weights and the resulting numerical taxonomy of ten OTU's analyzed together *versus* seven OTU's analyzed together with three OTU's added to the numerical taxonomic analysis after the weights have been obtained.

Analyses of different methods need some standard with which to compare the various results. This has always presented a problem in taxonomy and particularly in numerical taxonomy. In this study I have chosen seven species of *Juniperus* that I have analyzed for several years (Adams, 1969, 1970, 1972*a,b*, 1973; Adams & Turner, 1970; Powell & Adams, 1973; Zanoni & Adams, 1973); namely, *J. ashei*, *J. deppeana*, *J. flaccida*, *J. monosperma*, *J. pinchotii*, *J. scopulorum*, and *J. virginiana*. In addition, three other OTU's are added as exemplars: *J. deppeana* f. *sperryi*

(Correll) Adams, *J. pinchotii* (two divergent populations from the trans-Pecos region of Texas denoted *PE*), and a population of *J. scopulorum* from Palo Duro Canyon in the Texas Panhandle. These ten OTU's will be denoted respectively by *A*, *D*, *F*, *M*, *P*, *S*, *V*, *SP*, *PE*, and *ST*. The purpose of this paper is to compare different weighting methods and not to resolve relationships in *Juniperus*. Within the aforementioned ten OTU's, four relationships, as certain as modern objective methods can quantify them, are obvious from the studies cited: *1*) *P* (*J. pinchotii*) and *PE* (*J. pinchotti* from the trans-Pecos) are more similar to each other than either is to any other OTU in this study (Adams, 1972*a*); *2*) *ST* (a population of *J. scopulorum* from the Texas Panhandle) is most similar to *S*, a population of *J. scopulorum* about 150 miles west of *ST* (Adams, 1969); *3*) *V* (*J. virginiana*) is more similar to *S* or *ST* (*J. scopulorum*) than to any other OTU in this study (Adams, 1969); *4*) *SP* (*J. deppeana* f. *sperryi*) is more similar to *D* (*J. deppeana*) than to any other OTU, and *D* and *SP* are very similar (Adams, 1973). Other similarities are known among the OTU's, but these four are so obvious and incontrovertible that only these will be used to argue the points of various weighting methods. Notice that in each of the four cases cited we will be concerned with the *most similar* neighbor, therefore single linkage clustering will give a good representation of the four critical similarities used as base-line datum. The criterion for judging the relative merits of various weighting schemes will be the ability of the various methods to accurately describe these four similarities.

Materials and Methods

The terpenoids of the OTU's have been extracted, identified, and quantified (Adams, 1969). The approximately 136 terpenoid characters were analyzed by ANOVA to determine average CV's and F ratios for use in this study. According to the manner by which the initial data were organized, from 78 to 83 characters had F ratios larger than 1.0 and were thus included in the various analyses. In the case of the comparison of various weighting coefficients, the same 83 characters were used in each comparison.

The similarity measure used is basically a weighted matching coefficient (Sokal & Sneath, 1963).

Let:

x_i = value of character i in OTU (population) x
y_i = value of character i in OTU (population) y
Rd_{xy} = relative dissimilarity between OTU's x and y
Sr_{xy} = relative similarity between OTU's x and y
$= 1 - Rd_{xy}$
W_i = weight of character i in the summation
Rg_i = range of character i encountered in all the data
n = number of comparisons between OTU's x and y (excluding negative matches)

Then:

1. $$Rd_{xy} = \frac{\sum_i^n \frac{W_i|x_i - y_i|}{Rg_i}}{\sum_i^n W_i}; \qquad 0 \leqslant Rd \leqslant 1$$

2. $$Sr_{xy} = 1 - Rd_{xy}; \qquad 0 \leqslant Sr_{xy} \leqslant 1$$

TABLE 1. Weighting used in the six comparisons of different methods of character weighting. Eighty-three terpenoid characters were found with F ratios greater than 1.0. The various weights have been normalized to a percentage of the total weight to facilitate comparisons across the table. There is a trend toward equal weighting from left to right. Only the first 16 characters are shown; the complete table for 83 characters may be obtained from the author.

Character	$F-1$	F	$F-1$	$1-S_w/S_t$	$1/CV$	1.0 equal
1	.23	.25	.66	.64	1.47	1.20
2A	20.60	20.31	6.20	3.19	3.02	1.20
2B	1.44	1.44	1.64	1.84	4.10	1.20
2C	.19	.21	.59	.55	.46	1.20
3	.48	.49	.95	1.07	1.86	1.20
4	7.29	7.18	3.68	2.81	2.71	1.20
5	3.36	3.32	2.50	2.41	2.01	1.20
6	.56	.57	1.02	1.17	.69	1.20
7	1.63	1.62	1.74	1.93	2.48	1.20
8	1.40	1.40	1.62	1.83	2.09	1.20
9	.18	.20	.59	.54	1.86	1.20
10	.89	.90	1.29	1.50	.54	1.20
11	1.30	1.30	1.55	1.77	1.78	1.20
12	2.22	2.20	2.03	2.14	.38	1.20
13	1.05	1.05	1.40	1.62	1.47	1.20
14	.97	.98	1.35	1.56	2.09	1.20

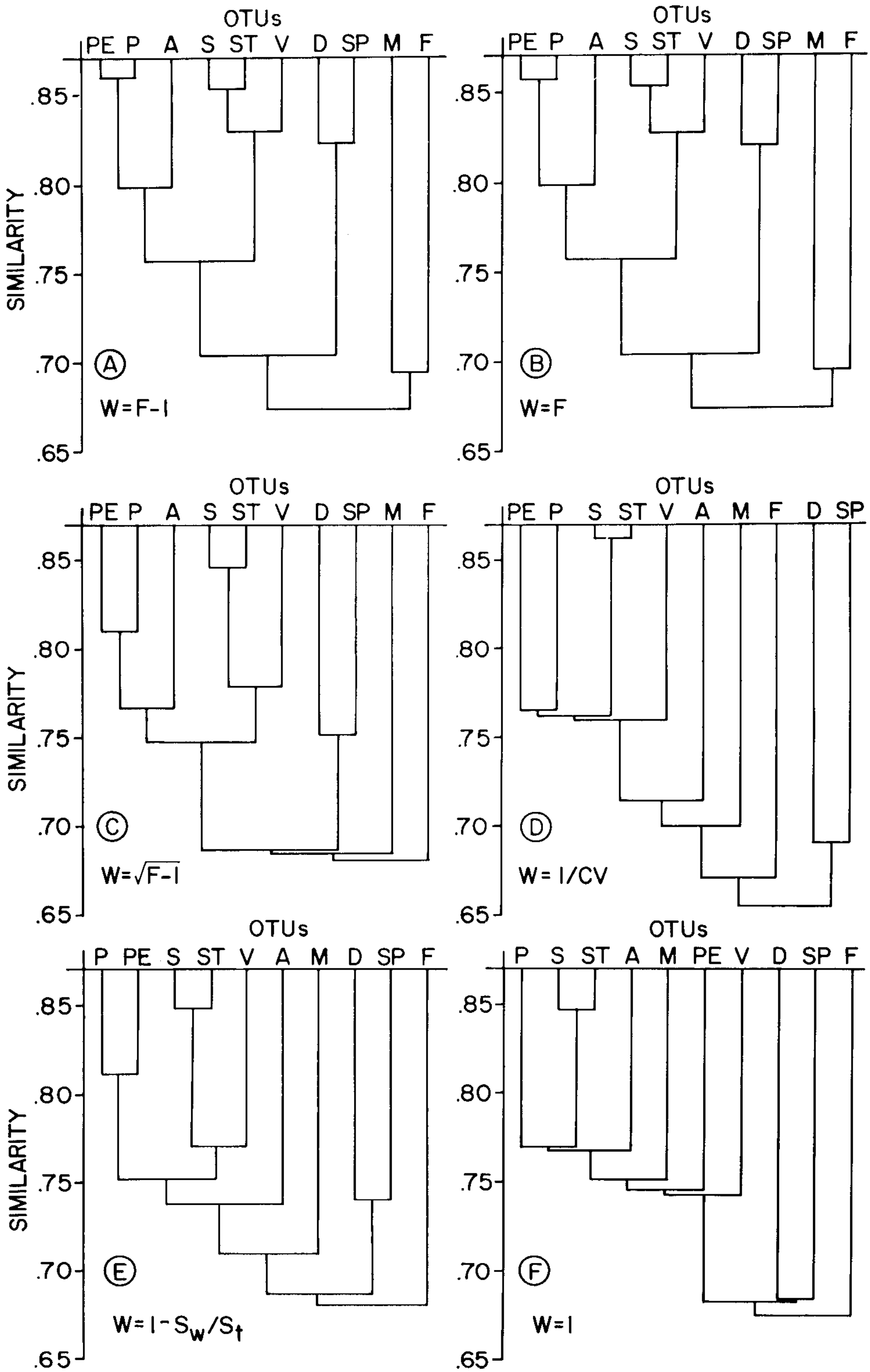
The effects of using different weighting were examined by substituting the following weightings (W) into Formula 1: 1.0; $1/CV$; $1-S_w/S_t$; $\sqrt{F-1}$; F; and $F-1$. The ten aforementioned OTU's were used with data composed of individuals only (i.e., $A = 64$ trees, $D = 16$ trees, $F = 15$ trees, $M = 15$ trees, $P = 88$ trees, $S = 8$ trees, $V = 10$ trees, $SP = 1$ tree, $PE = 11$ trees, and $ST = 5$ trees).

The effects of different organizations of the data (into various sets) were examined by comparing the ten OTU's using data sets composed of individuals only, *versus* data sets of population averages for *J. ashei* (13 populations) and *J. pinchotii* (17 populations) with individual tree data for the other eight OTU's, *versus* the use of one randomly chosen tree from each population of *J. ashei* (13 trees) and each population of *J. pinchotii* (17 trees) with individual tree data for the other eight OTU's, *versus* the use of population averages only. When population averages (pop. av.) were used the following sets were constructed: $A = 13$ pop. av., $D = 3$ pop. av., $F = 1$ pop. av., $M = 3$ pop. av., $P = 17$ pop. av., $S = 1$ pop. av., $V = 2$ pop. av., $SP = 1$ pop. av., $PE = 2$ pop. av., and $ST = 1$ pop. av. Obviously the error degrees of freedom in ANOVA was greatly diminished in the last case, but this sort of treatment is common in the literature. Thus, it seemed important to compare the use of population averages only, in the computation of statistical weights.

The effect of adding exemplars, before and after weights have been calculated, was examined by comparison of the results using all ten OTU's in the computation of weights *versus* the computation of weights using only seven OTU's and then adding the three exemplar OTU's (PE, ST, SP) during the computation of similarity measures.

→

FIG. 1. Phenograms of ten OTU's of *Juniperus* using various weighting methods, based on 83 terpenoid characters with F ratios greater than 1.0. W indicates the weighting method used. Phenograms A, B, C, and E adequately portray the known facts with D and F (equal weighting) being most distorted.


OTUs
PE P A S ST V D SP M F
SIMILARITY
.85 .80 .75 .70 .65
Ⓐ
W=F-1
OTUs
PE P A S ST V D SP M F
Ⓑ
W=F
OTUs
PE P A S ST V D SP M F
Ⓒ
W=√F-1
OTUs
PE P S ST V A M F D SP
Ⓓ
W=1/CV
OTUs
P PE S ST V A M D SP F
Ⓔ
W=1-Sw/St
OTUs
P S ST A M PE V D SP F
Ⓕ
W=1

Results

Effects of Different Statistical Weights on Similarities.—Table 1 shows the first 16 of the 83 character weightings used in examining the first question. The highest weighting went to compound 2A using the $F - 1$ weights, with this character accounting for 20.60% of the character weights. The decrease in weight for 2A across the table is evident in the following order: $F - 1$; F; $\sqrt{F-1}$; $1 - S_w/S_t$; $1/CV$ and finally equal weighting which gave 1/83 or 1.2% weight to this (and every) character. Another highly weighted character is compound 4, which varies from 7.29% ($F - 1$) to 2.71% in the weighting of Farris (i.e., $1/CV$) and of course 1.2% when equally weighted. It is obvious that a progressive series of weights ranges from the more dissimilar weights to the equal weights in the order of $F - 1$, F, $\sqrt{F - 1}$, $1 - S_w/S_t$, to 1.0 (equal weights). The weighting of Farris (inverse of variation within OTU's or 1/coefficient of variation in this paper) is near Flake's ($1 - S_w/S_t$) in many cases. Significant exceptions are characters 2B, 3, 8, 14, etc. Notice particularly compound 2B in Table 1, which has one of the largest weights for the method of Farris (i.e., 4.10%). Yet this character has 1.44%, 1.44%, 1.64%, and 1.84% weights using the other methods. It seems obvious that the use of the inverse of variation within OTU's may or may not agree with methods that take inter-OTU variation into consideration. Compound 2B appears to show little variation within the OTU's of this study but also fails to show variation among the OTU's. Thus it has very little weight in the methods that utilize F weighting.

The phenograms resulting from single-linkage clustering are shown in Figure 1. The method of obtaining character weights is shown at the lower left of each phenogram. Figures 1*A*, 1*B*, 1*C*, and 1*E* each satisfy the basic facts known about these OTU's, namely that *P* (*J. pinchotii*) and *PE* (*J. pinchotii*) are closely related (Adams, 1972*a*), *ST* (*J. scopulorum*) should be closely related to *S* (*J. scopulorum*), *ST* and *V* (*J. virginiana*) are closely related (Adams, 1969), and *D* (*J. deppeana*) and *SP* (*J. deppeana* f. *sperryi*) are closely related (Adams, 1973). The similarities among the other OTU's are not so well known although *A* (*J. ashei*) bears considerable affinity to *P* and *PE*, and *M* (*J. monosperma*) and *F* (*J. flaccida*) have somewhat similar oil patterns. To choose among weighting of $F - 1$, F, $\sqrt{F - 1}$, and $1 - S_w/S_t$ would be difficult, but the divergent populations of *P* (i.e., *PE*) are more similar to *P* when $F - 1$ and F-weighting are used than with $\sqrt{F - 1}$ or $1 - S_w/S_t$ weightings and there is a closer relationship between the OTU's *S* and *V* with F and $F - 1$ weightings. In addition we find that *SP* (which is a form of *D*) is more closely related to *D* using $F - 1$ and F weights. There does not appear to be much difference between using $F - 1$ and F weights (Figs. 1*A*, 1*B*), but on philosophical grounds there is good reason to subtract one from the F ratios since an F ratio of 1.0 implies that there is just as much variation within OTU's as among OTU's, and thus this character really contains zero amount of usable information for the classification.

The weighting of Farris ($1/CV$) results in the phenogram in Figure 1*D*. Notice that *P* and *PE* (*J. pinchotii*) are now only slightly more similar to each other than either one is to *J. scopulorum* (*S* and *ST*). *V* (*J. virginiana*) has now joined the cluster at a lower level of similarity than the *J. pinchotii* OTU's (*P* and *PE*). The OTU's *A*, *M*, and *F* are rather strung out in Fig. 1*D*. The two forms of *J. deppeana* (*D* and *SP*) are considerably less similar than in Figs. 1*A*, 1*B*, and 1*C*.

The sixth method, equal weighting, yields the most distorted results on these data. In Fig. 1*F* *P* and *PE* do not cluster together nor do *S*, *ST*, and *V*, nor do *D* and *SP*. The phenogram shows considerable chaining, which has often been attributed to single-linkage clustering (with some validity). This does indicate that of the four basic facts known about these ten OTU's, three are violated in the equally weighted

TABLE 2. Correlation coefficients between similarity matrices resulting from various weighting methods. Data in parentheses are the upper and lower 95% confidence limits for the correlation coefficient.

	$F-1$	F	$\sqrt{F-1}$	$1-S_w/S_t$	$1/CV$	1.0
$F-1$	----	.9999 (.9999, .9998)	.854 (.918, .749)	.761 (.862, .602)	.709 (.830, .526)	.608 (.765, .383)
F	----	----	.859 (.921, .757)	.767 (.866, .611)	.716 (.835, .536)	.616 (.770, .394)
$\sqrt{F-1}$	----	----	----	.986 (.993, .975)	.953 (.974, .916)	.913 (.951, .846)
$1-S_w/S_t$	----	----	----	----	.969 (.983, .944)	.944 (.969, .899)
$1/CV$	----	----	----	----	----	.970 (.983, .945)

phenogram (i.e., *P* and *PE* are not most similar; *S*, *ST*, and *V* are not most similar; and *D* and *SP* are barely most similar). In fact, the progression of character weighting is clearly correlated with those phenograms that describe the known facts about these OTU's! The heavier the character weighting, the better the phenograms appear to fit the known facts about these taxa.

Pearson product-moment correlation coefficients were calculated among the six similarity matrices (Table 2). F − 1 weighting is practically identical to *F* weighting and the correlation coefficient supports that relationship. Weighting by $\sqrt{F-1}$ is most highly correlated with $1 - S_w/S_t$ and then with $1/CV$ and equal weights. This is somewhat surprising since the $\sqrt{F-1}$ weights are so different from those obtained with $1/CV$ and 1.0. The weights obtained by $1 - S_w/S_t$, $1/CV$, and 1.0 are each highly correlated. Thus the correlation coefficients seem to divide the six similarity matrices into two groups, one group of the higher-weighted methods ($F-1$, F) and the other four, although $\sqrt{F-1}$ weighting is somewhat intermediate.

TABLE 3. Comparisons of the weights obtained using four different sample sets of data from the ten OTU's in this study. The weights have been normalized on a percent basis for comparison. Data for only the first 16 characters are given; a complete listing for all 83 characters may be obtained from the author.

Character	All OTU's (trees)	*A* and *P* (pop. av.)	*A* and *P* (13 and 17 trees)	All OTU's (pop. av. only)
1	.23	.17	.66	.31
2A	20.60	17.38	6.20	5.04
2B	1.44	1.58	1.64	1.31
2C	.19	.11	.59	.62
3	.48	.45	.95	.45
4	7.29	6.48	3.68	3.79
5	3.36	5.16	2.50	1.96
6	.56	.51	1.02	1.23
7	1.63	1.42	1.74	1.21
8	1.40	1.85	1.62	1.30
9	.18	.20	.59	.48
10	.89	.79	1.29	1.19
11	1.30	1.05	1.55	1.44
12	2.22	2.08	2.03	1.60
13	1.05	1.21	1.40	1.55
14	.97	.57	1.35	1.48

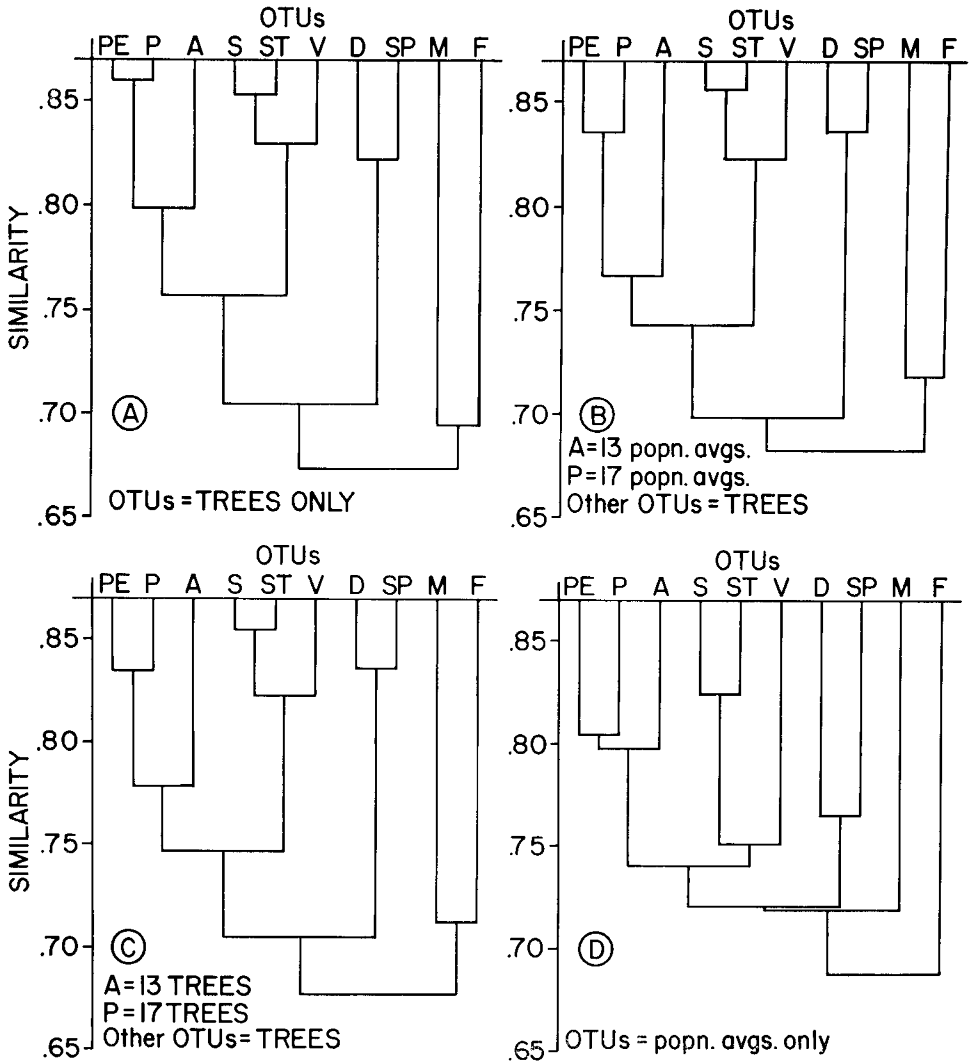


FIG. 2. Phenograms of ten OTU's of *Juniperus* using various organizations of the data sets to test the effects on statistically derived weights. See text for discussion.

Effects of Data Organization on Weights and Similarities.—The effect of different organizations of the data before the computation of the $F - 1$ weights is shown in Table 3. The greatest differences in weighting occurred when all OTU's were based on individual trees, and slightly more similar weights were obtained when the OTU's *A* and *P* were based on population averages. The use of one randomly chosen tree from each population of *A* and *P* resulted in more similar weights, but the most nearly equal weights were obtained when all OTU's were based on population averages only. In fact, the $F - 1$ weights obtained using population averages are very similar to those obtained by $\sqrt{F - 1}$ weighting using individual trees only (cf. Table 1).

Figure 2 shows the phenograms obtained by using these four data organizations. Figures 2*A*, 2*B*, and 2*C* are very similar, and we have very little basis to choose

TABLE 4. Correlation coefficients between similarity matrices resulting from the use of different methods to structure the data during the computation of the *F* ratios used for weighting. *F* − 1 weights were computed with the data sets arranged as: all OTU's composed of sets of individual trees only (trees only); OTU's *A* and *P* composed of 13 and 17 population averages ($A = 13\ p$, $P = 17\ p$); OTU's *A* and *P* composed of 13 and 17 individuals, each randomly chosen from a population ($A = 13$ trees, $P = 17$ trees); and all OTU's composed of population averages only (pop. av.).

	Trees only	$A = 13$ populations $P = 17$ populations	$A = 13$ trees $P = 17$ trees	Pop. av. only
Trees only	----	.973 (.985, .951)	.987 (.993, .977)	.791 (.880, .647)
A = 13 populations P = 17 populations	----	----	.995 (.997, .991)	.787 (.878, .642)
A = 13 trees P = 17 trees	----	----	----	.795 (.882, .653)

which is best. When population averages only are used for all OTU's, the error degrees of freedom goes down drastically and so does our confidence in the *F* ratios obtained from ANOVA. This starts to appear in Fig. 2*D* in that *P* and *PE* are more separated; *V* has almost left its group with *S* and *ST*; *S* and *ST* are less closely related; and *D* and *SP* are much less similar than before. Thus from a theoretical statistical view and from our basic facts, the weighting coefficients derived from the use of population averages only (Fig. 2*D*) appear to be somewhat poorer than those derived from the other three methods. It is interesting to compare the phenograms in Figs. 1 and 2 in that the use of population averages only (but *F* − 1 weighted) yields a phenogram very similar to that obtained using individual trees only but $\sqrt{F - 1}$ weighted (Fig. 1*C*).

Correlation coefficients were computed among the four similarity matrices (Table 4). The first three methods are each highly correlated whereas the use of population averages only to compute the weighting coefficients (*F* − 1) is not highly correlated with any of the other methods.

The use of more samples or more averages would be expected to lead to a better classification, and thus I would expect that Figs. 2*A* and 2*B* probably are a little

Table 5. Comparison of the weights obtained using ten OTU's or seven OTU's in computing *F* ratios. *F* ratios have been normalized to a percentage of the total weight for comparison. The three OTU's omitted from the second analysis are the exemplars: *ST* (population of *J. scopulorum* from Texas Panhandle): *SP* (*J. deppeana* f. *sperryi*) and *PE* (divergent populations *J. pinchotii* from trans-Pecos Texas). Data for only the first 16 characters are given; a complete listing for all 83 characters may be obtained from the author.

Character	10 OTU's	7 OTU's	Character	10 OTU's	7 OTU's
1	.23	.21	7	1.63	1.74
2A	20.60	21.20	8	1.40	1.39
2B	1.44	2.03	9	.18	.19
2C	.19	.20	10	.89	.90
3	.48	.39	11	1.30	1.20
4	7.29	7.40	12	2.22	2.43
5	3.36	3.56	13	1.05	.65
6	.56	.59	14	.97	1.02

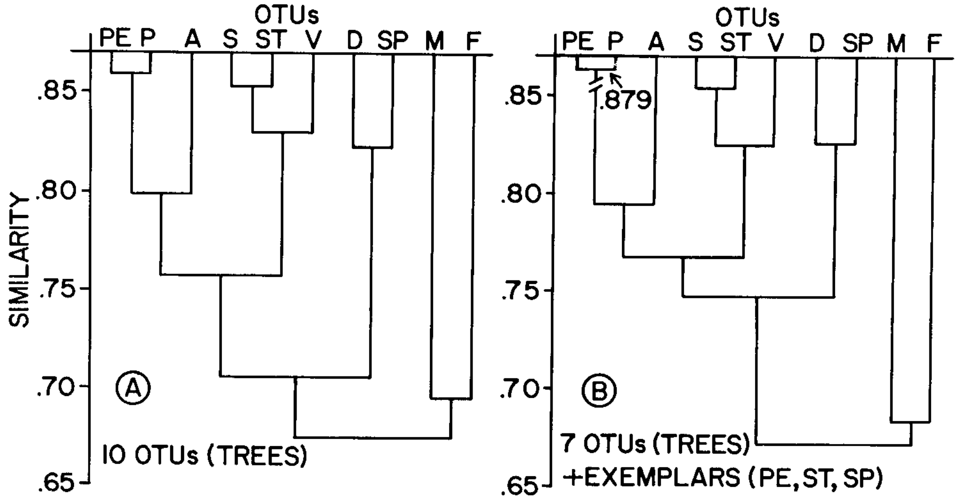


FIG. 3. Phenograms of ten OTU's of *Juniperus* to test the effect of inclusion and exclusion of exemplars (*PE*, *ST*, *SP*) in the computation of statistical weights. These effects are fairly small.

better than *2C* (and of course *2D*). The use of so many individuals for *P* and *A* but few individuals for the other OTU's could lead to peculiar *F* ratios favoring the separation of *A* and *P*. More study is needed to examine these effects, although it appears that unless the data sets are extremely unbalanced (in numbers), the resulting *F* ratios in the ANOVA and the subsequent classification will not be unduely affected.

Effects of Adding Exemplars.—Table 5 shows the weights ($F - 1$) computed using all OTU's in the ANOVA *versus* the use of seven OTU's in the ANOVA. The weights are very similar. The phenograms in Fig. 3 are very similar, and one can not say which fits the facts better. The similarity matrices are highly correlated (.993). It appears that the omission of OTU's *PE*, *ST*, and *SP* from the ANOVA (i.e., using these as exemplars after the weights have been computed) appears to favor a higher similarity between *P* and *PE*, *ST* and *S*, and *SP* and *D* (cf. Fig. 3*B*). This is not unexpected since the *F* weighting obtained from ANOVA favors those characters which divide the taxa. Thus when the three exemplars are included as OTU's in the ANOVA, the weighting favors a looser relationship among these exemplars and the correspondingly paired OTU.

It appears that if one wishes merely to identify some unknown OTU, it might be better to add it to the previous set after the ANOVA has been done. On the other hand, if one is interested in the relationship of a particular OTU to the other OTU's, one should include the exemplar(s) in the ANOVA or other weight-determining procedure.

CONCLUSIONS

The use of equally weighted characters as proposed by most numerical taxonomists for the past decade must be re-examined in light of the data presented in this paper. In fact, with the data set examined in this study, equal weighting gave the most distorted results followed, progressively, by 1/coefficient of variation within OTU's, $1 - S_w/S_t$, $\sqrt{F - 1}$, with $F - 1$ and F weighting having the highest fidelity to the known similarity. Are these results atypical? Further studies will undoubtedly be needed. Very similar results were obtained in studies of population differentiation of *J. ashei* (Adams, 1969, Adams & Turner, 1970) and *J. pinchotii* (Adams, 1969, 1972*a*). In both of these species the use of *F* weighted characters eliminated spurious similarities across geographical areas. I believe that the general case made for statistically derived weights will be re-emphasized as more examples are examined.

The criterion for weighting should satisfy both numerical and classical taxonomists since on the one hand the procedure is quite objective and repeatable yet entirely logical to the traditional taxonomists. Where multiple samples of the OTU's are not possible and many OTU's of the (presumedly) same taxon are used, one could follow the suggestion of Mayr (1964) and do a preliminary classification using equal weights and then put those OTU's together that are obviously closely related for use in ANOVA. The newly weighted characters could then be used for the second classification. Those OTU's of uncertain affinities in the original classification should probably be run as exemplars in the second classification using the weighted characters.

It should be noted that classifications in which characters just happen to have approximately equal statistical weights will of course not be affected by these methods. Flake (pers. comm.) found that in populational data for *J. virginiana* (Flake et al., 1969) it made little difference if the weights were $1 - S_w/S_t$ or *F* ratios. Examination of the *F* ratios (normalized) revealed that the largest weight was only 6.3% of the total weights and that many *F* values fell in the range of 2–4% of the total weights. This appears to be that special case mentioned above when the characters just happen to be of approximately equal *F* ratios and thus the differences between the various weighting methods are minimized. It is interesting to note in their paper Flake et al. (1969) tried equal weighting and found it inferior to the $1 - S_w/S_t$ weighting. In general, the probability that each character chosen for a study will in fact be of equal use in distinguishing taxa should be very small, however.

It appears to be better to compute statistical weights using a number of individuals or population averages per OTU than to have only a few population averages per OTU. This should be fairly evident when one considers the changes in error degrees of freedom in ANOVA and the loss of confidence in the data as the number of samples decrease.

The use and omission of exemplars in the computation of the weights in the ANOVA had only small effects in this study. Exemplars should probably be included in the ANOVA if one is primarily concerned with the relationships of the exemplar(s) to the other OTU's. If one is more interested in the identity of an unknown OTU (exemplar), then perhaps one should add the exemplar after the weights have been computed.

LITERATURE CITED

Adams, R. P. 1969. Chemosystematic and numerical studies in natural populations of *Juniperus*. Ph.D. Thesis. University of Texas, Austin.

———— 1970. Seasonal variation of terpenoid constituents in natural populations of *Juniperus pinchotii* Sudw. Phytochemistry **9**: 397–402.

———— 1972*a*. Chemosystematic and numerical studies of natural populations of *Juniperus pinchotii* Sudw. Taxon **21**: 407–427.

———— 1972*b*. Numerical analyses of some common errors in chemosystematics. Brittonia **24**: 9–21.

———— 1973. Reevaluation of the biological status of *Juniperus deppeana* var. *sperryi* Correll. Brittonia **25**: 284–289.

———— **& B. L. Turner** 1970. Chemosystematic and numerical studies of natural populations of *Juniperus ashei* Buch. Taxon **19**: 728–751.

Colless, D. H. 1971. "Phenetic," "phylogenetic," and "weighting." Syst. Zool. **20**: 73–76.

Farris, J. S. 1966. Estimation of conservatism of characters by constancy within biological populations. Evolution **20:** 587–591.

Flake, R. H., E. von Rudloff & B. L. Turner 1969. Quantitative study of clinal variation in *Juniperus virginiana* using terpenoid data. Proc. Natl. Acad. U.S.A. **64:** 487–494.

Mayr, E. 1964. The new systematics. *In* C. A. Leone (ed.), Taxonomic Biochemistry and Serology. New York.

Powell, R. A. & R. P. Adams 1973. Seasonal variation in the volatile terpenoids of *Juniperus scopulorum* (Cupressaceae). Amer. J. Bot. **60:** 1041–1050.

Sneath, P. N. A. & R. R. Sokal 1973. Numerical Taxonomy. San Francisco.

Sokal, R. R. & P. N. A. Sneath 1963. Principles of Numerical Taxonomy. San Francisco and London.

Zanoni, T. A. 1974. The genus *Juniperus* (Cupressaceae) in Mexico and Guatemala: A Taxonomic Revision. M.S. Thesis. Colorado State University, Fort Collins.

——————— **& R. P. Adams** 1973. Distribution and synonymy of *Juniperus californica* Carr. (Cupressaceae) in Baja California, Mexico. Bull. Torrey Bot. Club **100:** 364–367.