The Truss: Body Form Reconstructions in Morphometrics

Richard E. Strauss; Fred L. Bookstein

Systematic Zoology is currently published by Society of Systematic Biologists.

THE TRUSS: BODY FORM RECONSTRUCTIONS IN MORPHOMETRICS[1,2]

RICHARD E. STRAUSS AND FRED L. BOOKSTEIN

Abstract

Strauss, R. E., and F. L. Bookstein (Museum of Zoology, The University of Michigan, Ann Arbor, Michigan 48109) 1982. The truss: body form reconstruction in morphometrics. Syst. Zool., 31:113–135.—In principle, any measured distances between landmarks of a form may serve as characters for morphometric analyses. Systematic studies typically are based on a highly biased and repetitious sample of these. But collections of landmarks and the distances among them must be homologous from form to form for comparisons to be meaningful, and an adequate character set should at least permit the full reconstruction of the original configuration of landmarks.

We describe a geometric protocol for character selection, the truss network, which enforces systematic coverage of the form and which exhaustively and redundantly archives the landmark configuration. Reconstruction of the form from truss measures provides Cartesian coordinates for landmarks and allows estimation of, and compensation for, measurement error. Samples of forms may be averaged and standardized to one or more common reference sizes by regression of measured distances on a composite measure of body size, followed by reconstruction of the form using distance values predicted by the regression functions at some standard body size.

Principal component loadings of distance measures may be indicated directly on the truss network to display patterns of within-group allometry or between-group shape differences. Because the truss enforces use of cross measurements, discrimination among groups may be enhanced. Composite mapped forms are useful in biorthogonal analyses of differences in shape because they allow the comparison of averaged forms among samples. Certain patterns of principal component loadings are concordant with, and provide an initial sampling of, the biorthogonal grids for these deformations. [Allometry; biorthogonal analysis; discriminant analysis; fishes; morphometrics; multivariate analysis; principal components; triangulation; truss.]

Systematists are often interested in quantifying differences in form among different species, conspecific populations, or ontogenetic stages. Customarily, morphometric data are taken without regard for allometry or its variations among populations or growth stages. It has been common to study growth, for example, by analysis of body length or height or weight only; to describe the shapes of bones by measures of their lengths and widths; to characterize entire forms by relative lengths and breadths of head, trunk, tail, and appendages. Although such measures are deeply entrenched in the methodology of systematics, their usefulness in solving real biological problems may be limited. There are far more homologous measures on biological forms than are used in typical multivariate data sets (Humphries et al., 1981), and results of morphometric analyses can depend upon the particular set of measurements chosen. If the selection of distance measures does not correspond by accident or design to the principal directions of shape difference, the resulting descriptions of the differences between forms will be inadequate.

There are several biases and weaknesses inherent in traditional character sets such as those in our Figure 1 or in
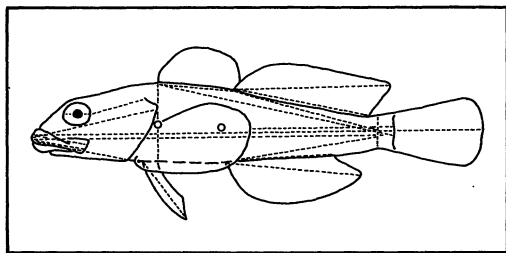
FIG. 1.—A "traditional" set of distance measures, one used by Strauss (1980), which is typical of character sets used in studies of fishes and other organisms. Dotted lines represent projections of the measurements onto the midsagittal plane. Small circles represent measures of body width.

Humphries et al. (1981; Fig. 1). They may be summarized as follows. (1) Most characters tend to be aligned with one of a very few axes, such as the "longitudinal," with only scant sampling of depth and breadth. Thus a large amount of information in the data is repetitious while other information, variation in oblique directions, is lacking. (2) Coverage of the form is highly uneven by region as well as by orientation: dense in some areas of the body and sparse in others. (3) Some morphological landmarks, such as the tip of the snout and the posterior end of the vertebral column, are used repeatedly. Any uncertainty in the positions of these morphological features will be propagated through series of measurements. (4) Many landmarks are "extremal" rather than "anatomical" (*sensu* Moyers and Bookstein, 1979). Anatomical landmarks are true homologous points identified by some consistent feature of the local morphology (Jardine, 1969; Schaeffer, 1976). Extremal landmarks are defined in terms of minimum or maximum distances (e.g., greatest body depth), and therefore their placement may not be homologous from form to form. (5) Many measurements extend over much of the body. Because short distances contain more localized information than long ones, covariances among long characters express average covariation and are less informative. (6) When measurements are taken on soft-bodied

organisms, the amount of distortion due to preservation cannot be easily estimated. The repeatability of individual measurements can be checked, but one cannot determine what the measure would have been if the specimen had been preserved differently.

Multivariate morphometrics offers no system for selecting characters to be studied other than that they be numerous and sample the whole form. Consequently, traditional data sets often turn out to be highly biased in many of the ways described above. Success in selecting effective characters has been largely a matter of chance (Bookstein, 1982). In this paper we propose a geometric protocol for character selection which largely overcomes the disadvantages of traditional data sets and leads logically to certain styles of analysis. The method allows us to: (1) systematically detect shape differences in oblique as well as horizontal and vertical directions, using a system of measures that ensures generally even coverage of the landmark configuration; (2) archive the configuration of landmarks so that the form may be reconstructed (mapped) from the set of distances among landmarks; (3) recognize and compensate for random measurement error; (4) average the forms of a sample of individuals, permitting us to compute and illustrate average shapes; (5) succinctly characterize and visualize multivariate trends of growth and allometry within populations; and (6) standardize forms for intergroup comparison, particularly for biorthogonal analyses of shape differences.

ARCHIVING THE FORM

For any set of homologous landmarks on an outline (as of the midsagittal plane of a fish), there are several possible patterns of distances among the landmarks that may be used to reconstruct the form (Mikhail, 1976). The simplest is triangulation. When the measured distances among landmarks are chosen to form a series of contiguous triangles (Fig. 2A), the set of measurements can be used to map the coordinates of the landmarks

A. TRIANGULATION $2n-3$ distances

B. TRUSS $5n/2 - 4$ distances

C. GLOBAL REDUNDANCY $3(n-2)$ distances
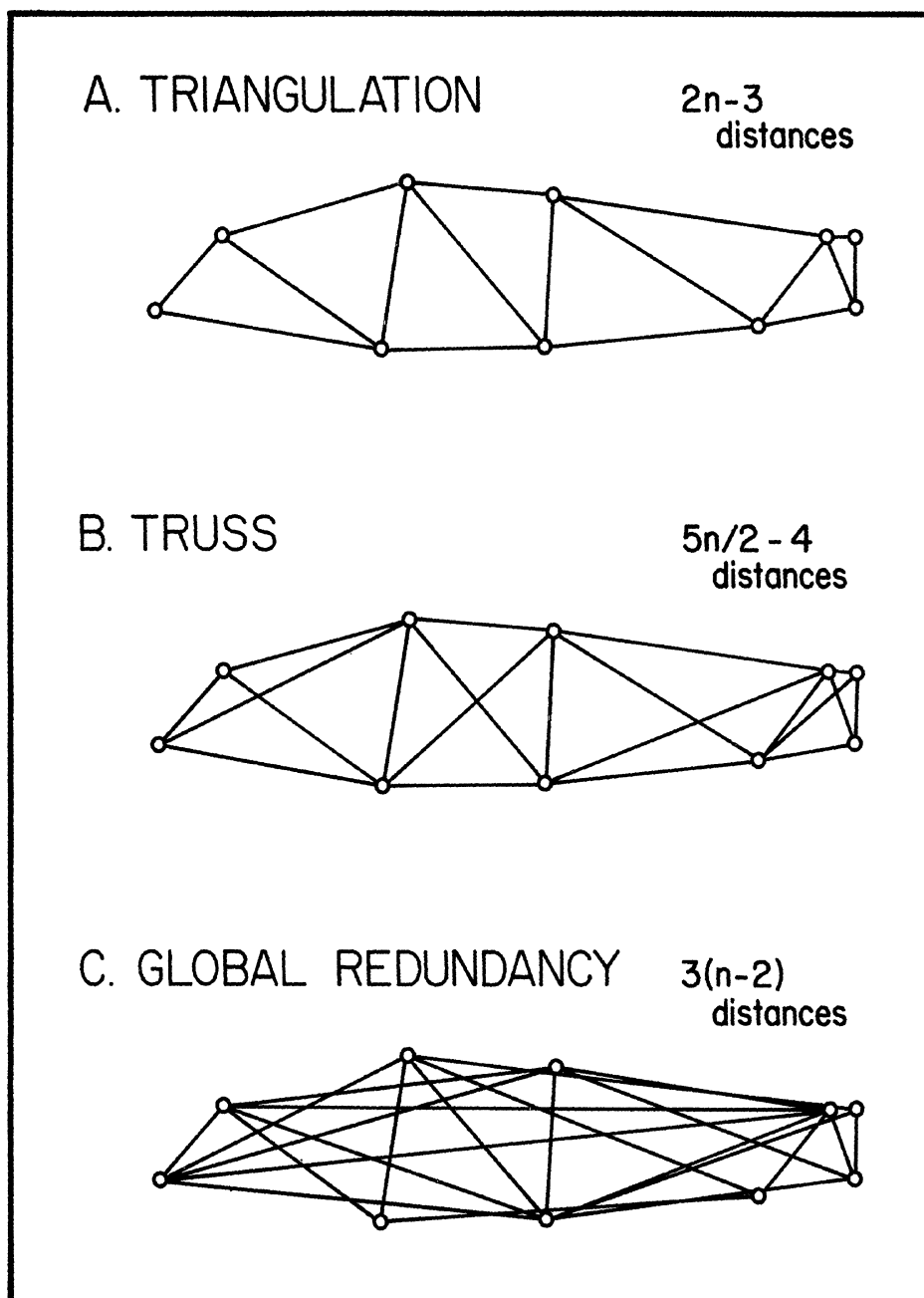
FIG. 2.—Possible patterns of distances, among 10 $(=n)$ coplanar landmarks, which completely archive the landmark configuration. (A) A triangulation network. (B) The box truss. (C) A "globally redundant" design patterned after Rohlf and Archie (1978).

FIG. 3.—Mapping Cartesian coordinates of landmarks by triangulation. (A) The first landmark is desig-nated to be the origin of the coordinate system, and the line connecting the first two landmarks to be the abscissa. The third landmark is positioned at the intersection of two circles, centered on the first two

(Fig. 3). The number of measures to be taken for a triangulation network increases as $2n$, twice the number of landmarks. Triangulated patterns are useful for descriptive measurements whether or n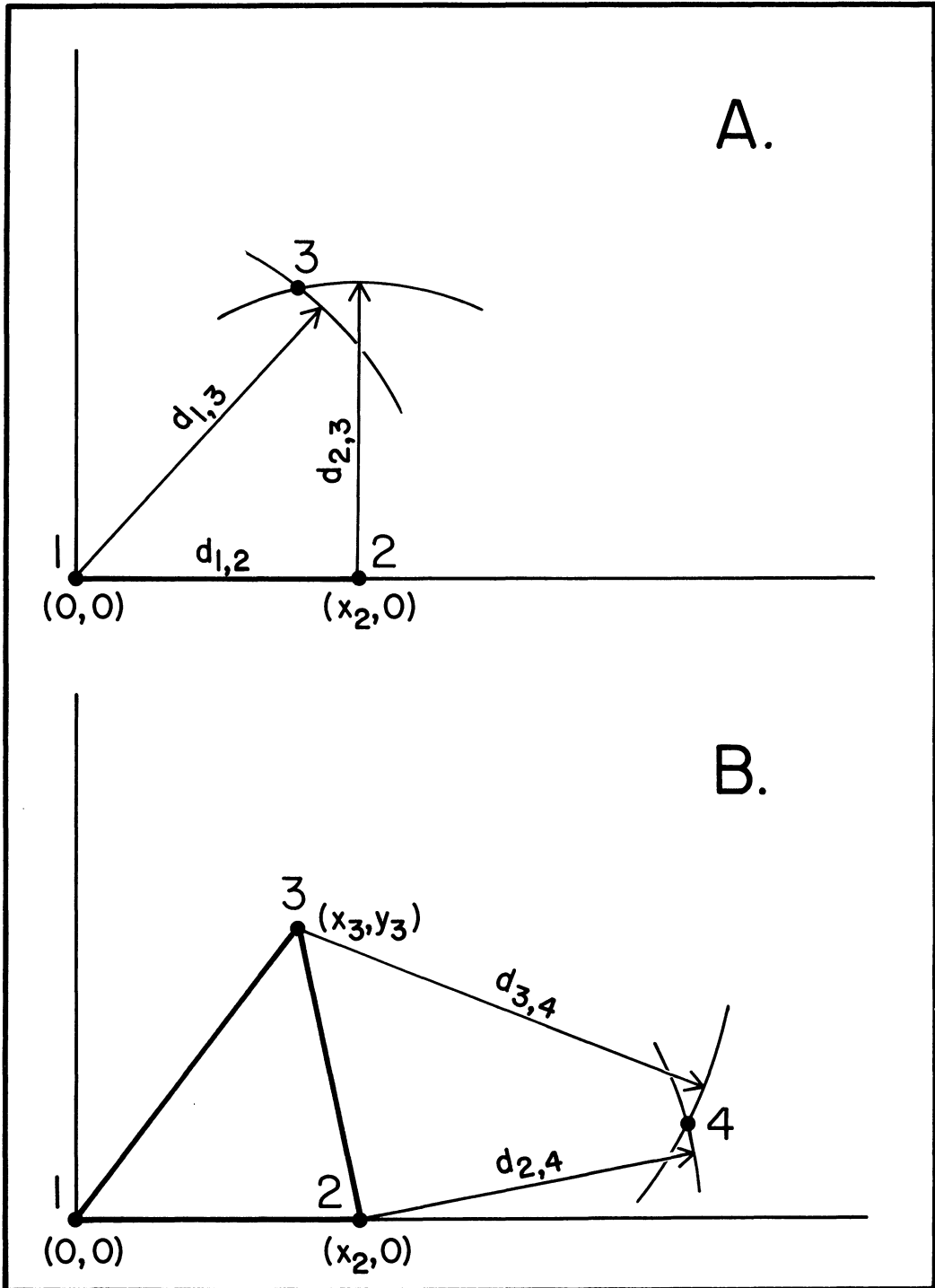ot figures are meant to be reconstructed (Olson and Miller, 1958, fig. 55; Skalak et al., 1982). However, an important problem attends the sequential mapping of coordinates by triangulation: small measurement errors are propagated as distortion throughout the form, so that the positions of the final landmarks can be unreliable.

The accuracy of the mapping could be improved if redundant distance measurements for each landmark were used to average the effects of random measurement error. Such a system was described by Rohlf and Archie (1978) to map landmarks (in their case, the positions of trees in a forest) using repeated measurements of distances among them. In their system any three landmarks are chosen for a basal reference triangle. Distances are then measured from each additional landmark back to any three previous ones (Fig. 2C). From these data, initial estimates of landmark coordinates can be established by triangulation. However, because each landmark sits at the apex of one or more extra triangles, the redundant measurements can be used to calculate an average set of coordinates by means of iterative least-squares. The technique can map landmark coordinates with a precision greater than that inherent in the original distance measures. The number of measurements needed increases as $3n$, which is still much less than the $n(n - 1)/2$ distances distributed among all possible pairs of landmarks (which increases as n^2).

A systematic pattern of measurements intermediate between these two is the box *truss* (Fig. 2B). In this system, homologous landmarks on the boundary of the form are divided into two tiers and paired. [There may be a single landmark left over at one end or the other.] The distance measures connect these landmarks into an overdeterminate truss network, a series of contiguous quadrilaterals each having both internal diagonals. Each quadrilateral shares one edge with the preceding quadrilateral and another with the succeeding one. In this way we add approximately one extra distance measure to a triangulated network for each four measures previously present. Paired landmarks at the ends of the truss network lie at the apex of three triangles; other paired landmarks lie at the apex of six triangles. This modest but systematic redundancy allows for checks on the consistency of the measurements (as described below). The extra distances permit the positions of mapped landmarks to be averaged, limiting the accumulation of measurement error. If landmarks are spaced suitably on the outline, the pattern of measurements will approach an idealized square truss (Fig. 4A), for which the expected confidence region for each landmark is smallest. The number of measurements required for the truss without unpaired landmarks is $5n/2$, midway between the $2n$ characters of the triangulation and the $3n$ characters needed for Rohlf and Archie's "globally redundant" system. The pattern ensures balanced coverage across the form and can be applied to various projections which together account for 3-dimensional aspects of shape (Fig. 4B, 4C).

ADJUSTMENT FOR MEASUREMENT ERROR: FLATTENING THE TRUSS

For each quadrilateral of the truss network there are six distances (four edges and two diagonals) among four landmarks (Fig. 5A). For the landmarks to be precisely coplanar, these distances must sat-

---

←

landmarks, whose radii are its distances to these landmarks. (B) The fourth landmark is then mapped in relation to the second and third landmarks. The procedure is continued across the form until coordinates have been assigned to all landmarks.
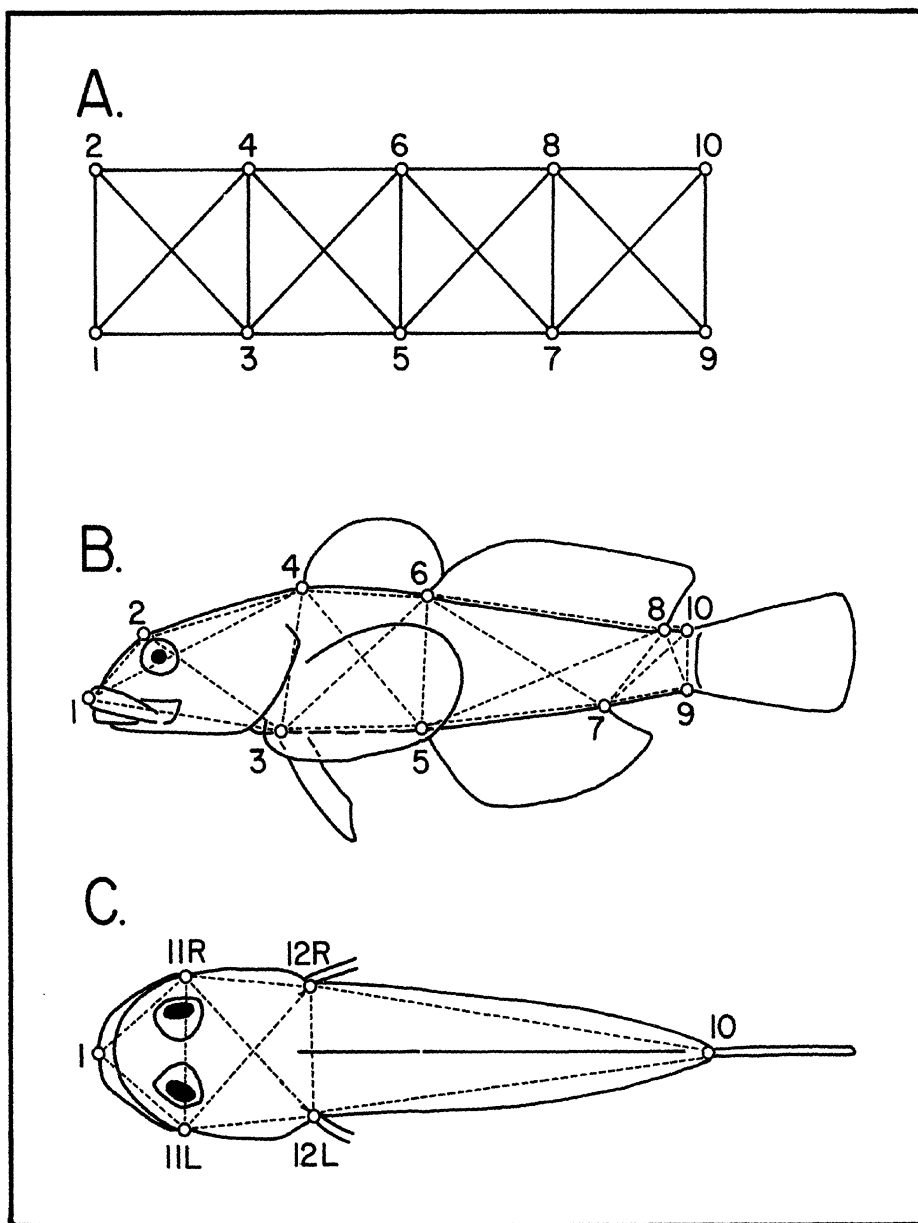
FIG. 4.—(A) An idealized square truss for 10 landmarks. (B) A truss network of distance measures applied to 10 midsagittal landmarks of a sculpin (*Cottus*). (C) A single truss cell with two appended triangles, applied to the dorsal projection of the same organism.
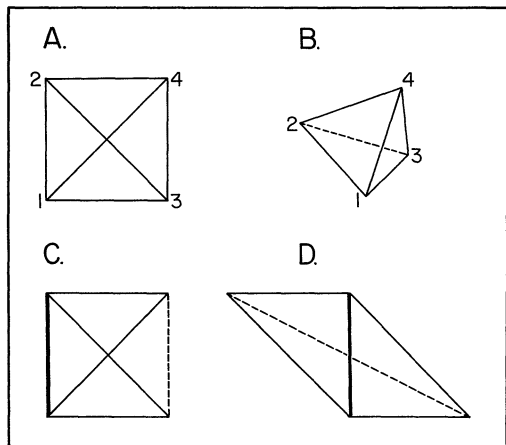
FIG. 5.—(A) A single truss cell consisting of four landmarks connected by six distance measures. (B) If the six distances do not satisfy Salmon's criterion of planarity, they will instead describe a three-dimensional tetrahedron. (C) For any set of five distances, the minimum permissible value of the sixth (dotted line) is the distance between the apices of two coplanar triangles lying on the same side of their shared edge (heavy line). (D) For the same set of five distances, the maximum permissible value of the sixth (dotted line) is the distance between the apices of two coplanar triangles lying on opposite sides of their shared edge (heavy line). If the length of the sixth distance is intermediate between these two extremes, the six distances describe a tetrahedron. If the sixth distance is beyond these extremes, the tetrahedron is imaginary.

isfy a single equation—the determinant of the following matrix must be exactly zero (Salmon, 1914, vol. 1:47):

$$\begin{vmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & d_{12}{}^2 & d_{13}{}^2 & d_{14}{}^2 \\ 1 & d_{12}{}^2 & 0 & d_{23}{}^2 & d_{24}{}^2 \\ 1 & d_{13}{}^2 & d_{23}{}^2 & 0 & d_{34}{}^2 \\ 1 & d_{14}{}^2 & d_{24}{}^2 & d_{34}{}^2 & 0 \end{vmatrix} = 288V^2$$

Here $d_{ij}{}^2$ is the squared distance between landmarks i and j, and V is the volume of the tetrahedron they determine. Because these measurements will often be taken with calipers on soft-bodied, hand-held specimens, it is unlikely that the six dis-

tances will exactly satisfy Salmon's criterion of planarity even though the landmarks would ideally lie in a single plane (e.g., the midsagittal). The distances will instead correspond to the six edges of a tetrahedron with some measurable volume (Fig. 5B). The mapping of landmarks then becomes a matter of flattening the tetrahedron by adjusting the six distances until the landmarks become coplanar.

Any five of the six distances exactly specify two coplanar triangles sharing an edge. We need not identify any particular one of the six distances (edge or diagonal) in a quadrilateral as the "extra" one. Given any five distances, the sixth must take one of two values, depending on whether the triangles are on the same side of the shared edge (Fig. 5C) or on different sides (Fig. 5D). For each of the six sets of five distances, the sixth distance will vary by measurement error (its own plus that due to error in the other five distances) from the exact value forced upon it by this planar construction. If the sixth length as measured is between the extremes of the two lengths indicated in Figures 5C and 5D, one can imagine this distance to be that of the real edge of a tetrahedron made up of the two trianges hinged along their common edge (Fig. 5B); the tetrahedron is the same for all choices of five edges out of six. The tetrahedron cannot be drawn if the length of the sixth edge is beyond either extreme, just as no triangle can be drawn with one edge-length greater than the sum, or less than the difference, of the other two; the "angle" at the hinge in this case is imaginary, and the squared volume according to the preceding determinant is negative.

Cell by cell, whether or not there is a real tetrahedron to be flattened, we may compute the exactly planar configuration that minimizes the sum-of-squares $\Sigma_1^6 (d - \hat{d})^2$ between the measured distances and the edges and diagonals of the flat reconstruction. This optimization is managed for us by the subroutine ZXSSQ from the International Mathematical and Statisti-

cal Library of numerical software (IMSL, 1977). The fitted planar quadrilateral will preserve Σ_1^6 d, the sum of the lengths in each cell, while balancing error-of-fit about all six distances. If one edge is grossly mismeasured, the fitted quadrilateral will show a pattern of compensating alterations in all the other distances.

Each pair of landmarks at which two quadrilaterals abut will likely be adjusted to different distances in the two separate quadrilateral adjustments involving them. In the assembly of the final configuration this discrepancy, which is usually quite small, can be compromised in the following way. The first quadrilateral is laid down arbitrarily, with its lower left landmark positioned at coordinate (0,0). Each succeeding quadrilateral is then sequentially adjoined to its predecessor such that the midpoints of the shared edges are superimposed and the edge directions aligned. This superimposition continues to approximately preserve the sum of the truss distances, and adds least error to the distances computed for those edges separately (at slight cost in net accuracy for the horizontal and diagonal segments).

The reconstruction of an entire truss in this way closely approximates a global best-fit that is feasible but much more expensive to compute. Propagation of errors from cell to cell (e.g., Fig. 6:4) is much less extensive than for triangulation without redundancy. The cross-measures we take tend to be not only shortest but also most nearly at 45° to the anteroposterior and dorsoventral measures, and therefore bear the most additional information. The longer distances that might be used to augment a globally redundant best-fit, such as diagonals 1–6, 2–5, or 3–10 (Fig. 4), neither add much further stiffness to

the reconstruction of forms as elongated as the fishes we study, nor contribute much additional information to subsequent morphometric analysis. Their loadings are extended averages of coefficients already quantified locally.

Figure 6 shows five specimens of the freshwater sculpin *Cottus cognatus* for which the landmarks have been mapped by this reconstruction. All except the fourth are reasonable approximations to the form of a fish. The apparent distortion of this individual is the result of a gross measurement error which was immediately obvious once the figure was drawn. Many such errors, especially those resulting from misreadings of calipers, ocular micrometers and other measuring devices, may go undetected in morphometric studies even when they result in outliers on scatter plots or histograms.

Substantial measurement errors will be exposed when the form is drawn, while more subtle errors may be revealed by measures of the mutual inconsistency of the data. A useful measure of such "strain," adjusted for the varying sizes of the cells, is the relative discrepancy between measured and reconstructed distances, $(d - \hat{d})/d$. (Note that this is not quite the quantity which is minimized in the flattening.) This statistic, computed for each distance measure, may be aggregated to describe the strain for each quadrilateral, each specimen, or for a sample of forms. One satisfactory net measure of strain is the root sum-of-squares of these, $\sqrt{\Sigma_1^{21} (d - \hat{d})/d]^2}$, for all the distances of a truss. A second useful measure of strain, at the level of the cell rather than the edge, might be the volume V of the tetrahedron before it is flattened, scaled by the 1.5 power of the flattened area. Many others are possible.

$\longrightarrow$

FIG. 6.—Mapped landmark configurations for five specimens of the sculpin *Cottus cognatus*. Landmarks are those indicated in Figure 4B. The body form of the fish in relation to these landmarks is indicated on the fifth figure. Total strain is a measure of the mutual lack of fit of the original distance measurements; it is the square root of the sum of the squared relative deviations of the distances in the mapped form from those originally measured.
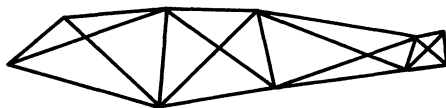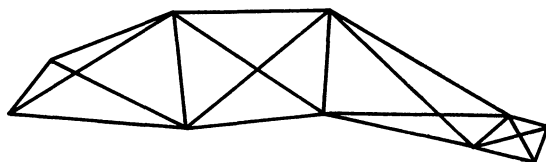
1. Total strain = 0.124
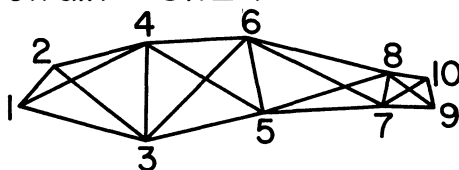
2. Total strain = 0.066
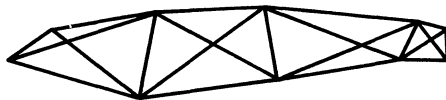
3. Total strain = 0.043

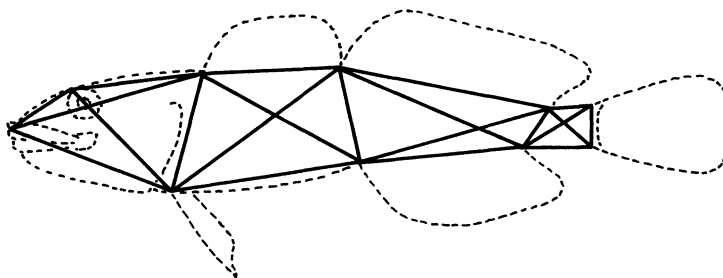4. Total strain = 0.180

5. Total strain = 0.051

TABLE 1. STRAIN STATISTICS FOR THE FIVE TRUSSES OF FIGURE 6.

		Distance measures[1]						
Specimen	Cell	1-2 3-4 5-6 7-8	1-3 3-5 5-7 7-9	2-3 4-5 6-7 8-9	1-4 3-6 5-8 7-10	2-4 4-6 6-8 8-10	3-4 5-6 7-8 9-10	RSS[2]
1	1	−.007	−.001	.004	.005	−.004	.005	.012
	2	.005	.012	−.018	−.018	.018	.010	.036
	3	.010	.005	−.002	−.005	.004	−.016	.020
	1-3							.041
	4	−.016	−.034	.048	.048	−.067	−.059	.118
	1-4							.124
2	1	.003	.001	−.001	−.001	.002	.001	.004
	2	.001	.002	−.002	−.002	.003	.004	.006
	3	.004	.019	−.012	−.017	.008	.005	.030
	1-3							.031
	4	.005	−.007	.025	.025	−.038	−.026	.058
	1-4							.066
3	1	.008	.002	−.003	−.005	.008	.004	.013
	2	.004	.006	−.010	−.009	.011	.005	.019
	3	.005	−.002	.002	.002	−.001	−.010	.011
	1-3							.025
	4	−.010	−.008	.014	.015	−.024	−.013	.036
	1-4							.043
4	1	.011	.001	−.011	−.010	.004	−.033	.038
	2	−.033	−.090	.078 ·	.078	−.062	−.042	.164
	3	−.042	.007	−.009	−.005	−.004	−.010	.045
	1-3							.166
	4	−.010	−.013	.027	.029	−.048	−.027	.072
	1-4							.180
5	1	.015	.003	−.005	−.008	.010	−.002	.021
	2	−.002	−.004	.008	.008	−.010	−.004	.016
	3	−.004	.000	.000	.000	−.001	−.010	.010
	1-3							.027
	4	−.010	−.008	.016	.018	−.032	−.015	.044
	1-4							.051

[1] The value for each distance measure (identified by the landmarks it connects) is the difference between the original measure and the corresponding distance on the final mapped configuration, scaled by dividing by the original measure.

[2] Net strain for each cell is the root sum-of-squares (RSS) of the six strain values. Net strain per form is the RSS of 16 or 21 strain values.

For example, the distance deviations (Table 1) for the trusses of Figure 6 indicate that the distortion of the fourth specimen may be the result of a measurement error for the lower horizontal element (between landmarks 3 and 5) of the second truss cell. In fact, this distance was mismeasured by approximately −9%; the lengths of the other distances in the cell were adjusted to compensate for this. The values of strain also reveal that for all five specimens the positions of landmarks 9 and 10 are somewhat uncertain, resulting in inconsistencies among the measures involving them. In such cir-

cumstances the affected truss cell might be excluded from subsequent analyses.

As a second example of the use of such statistics, we have plotted against a composite measure of body size the total strain per individual for data from 43 specimens of another sculpin, Cottus klamathensis (Fig. 7). Although only a slight relationship of imprecision of measurement to body size is evident,[3] there are obviously

[3] In the absence of measurement imprecision, strain is expressing nonplanarity of the original configuration of landmarks, and so would not be expected to correlate with specimen size.
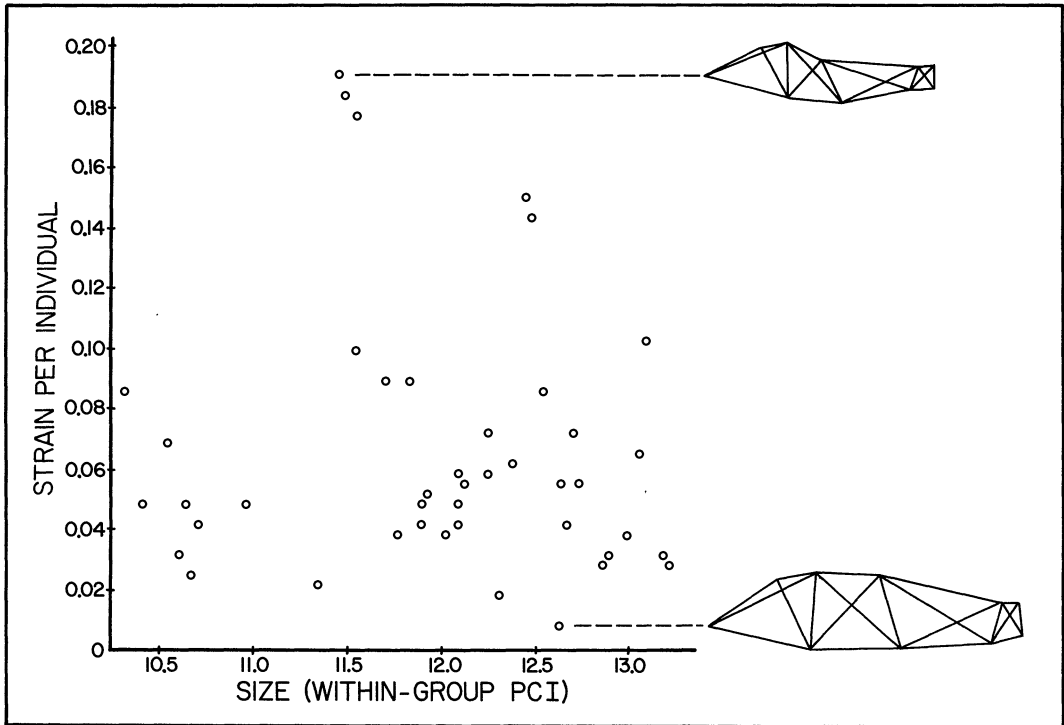
FIG. 7.—A scatter plot of strain (mutual lack of fit of the data) per individual against a composite measure of body size, for 43 specimens of *Cottus klamathensis*. Body size is the score on the first within-group principal component of the covariance matrix of log-transformed distances. Reconstructed forms are shown for the two specimens having the best and worst fit.

several individuals for which the data fit poorly. The three data points of highest strain corresponded to specimens for which either caliper settings were mis-read or keying errors were made. Values of strain for each distance measure within the forms again indicated the particular measures which were in error.

To test the degree to which the least-squares mapping procedure corrects for imprecision of measurement, we took data on a single specimen of *Cottus cognatus* to the nearest 20th of a millimeter, the usual resolution of measurement, and mapped them to reconstruct the form. The data were then remapped after being rounded to the nearest millimeter (Fig. 8). The strain estimate for the less precise data is twice that for the data as originally measured, 0.121 *vs.* 0.073, but the result-ing forms are virtually identical. We then

repeated the procedure for nine more specimens. Figure 9 is a scatter of the Cartesian coordinates of the ten individuals mapped at the lower precision (mean strain, 0.103) against the same coordinates mapped at the higher precision (mean strain, 0.058). If the mapped forms were identical over the refinement, the points would lie exactly along the 45° lines of equality. In fact, even though the strain is consistently greater for the less precise data, the points lie very close to this line; the slight distortion created by rounding off the data is corrected by the least-squares algorithm.

COMPUTING AVERAGE SHAPES AND CHARACTERIZING GROWTH TRENDS

Thus, the truss network can be used to map landmarks effectively by "relaxing" nonplanarities. We propose using this
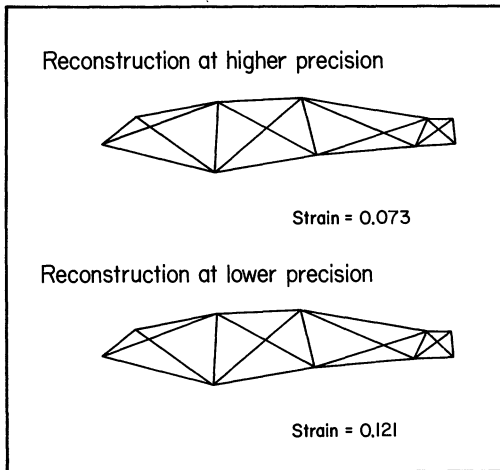
FIG. 8.—Reconstructed truss networks for a single specimen of *Cottus cognatus*. Measurements were originally taken to the nearest 0.05 mm (higher precision), and then rounded to the nearest 1 mm (lower precision).

truss system to produce single representative forms for subsequent descriptive analyses by averaging within-sample size and shape variation. There are five steps to this procedure.

(1) Choose landmarks on the body outline and connect them appropriately with line segments to form a truss network.
(2) Measure these distances on a sample of individuals (and check them by individual truss reconstructions).
(3) Diagnose the effects of body size and allometry on the measured truss characters by log-linear regressions of the measured distances upon some composite measure of body size (Fig. 10).
(4) Choose a "standard" body size and compute the predicted values of all the distances, as estimated by the log-linear regression functions, at that size.
(5) Map the coordinates of the landmarks by using the predicted distances in the mapping procedure described in the previous section. When the standard size is a population average, we have produced an average form by reconstruction from the average measurements.

The composite size measure used in steps 3 and 4 could be the first within-group principal component of the covariance matrix of the log-transformed distances (Jolicoeur, 1963). This is the latent (unmeasured) variable which optimally describes the joint log-linear covariation
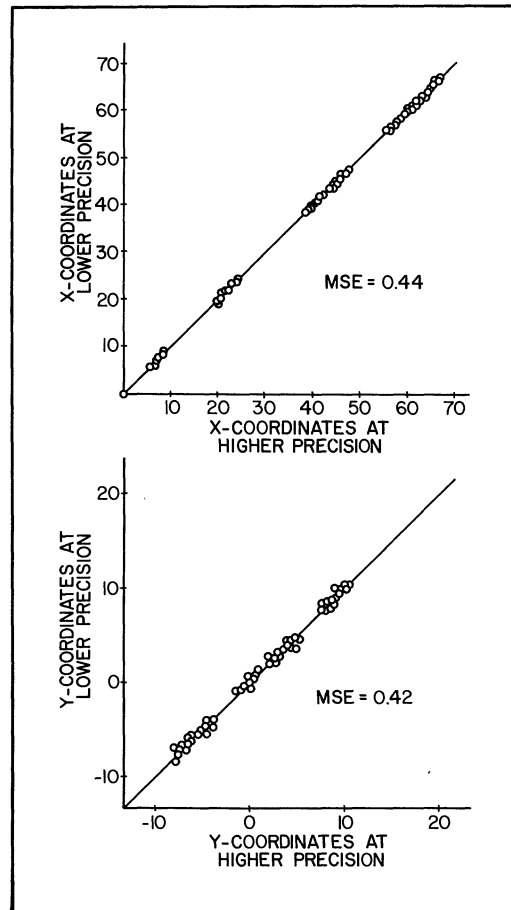

FIG. 9.—Scatter plots of the Cartesian coordinates of reconstructed truss networks for 10 specimens of *Cottus cognatus*. In all the reconstructions, landmark 1 was set to the origin and landmark 9 placed along the x-axis. Measurements were originally taken to the nearest 0.05 mm (higher precision), and then rounded to the nearest 1 mm (lower precision). If the mapped forms were identical over the change in precision, the points would lie exactly along the 45° lines of equality. Axis scales differ for the two plots. The mean-square errors (MSE) of the regressions are approximately equal.

in all distance measures simultaneously (Bookstein, 1982).[4] Removal of size vari-

---

[4] One might also use the log mean truss-element length, log(Σd/21), as it is approximately invariant under the operation of flattening the truss. It is very highly correlated with the usual first principal component, but may be calculated for each individual without considering the rest of the population.
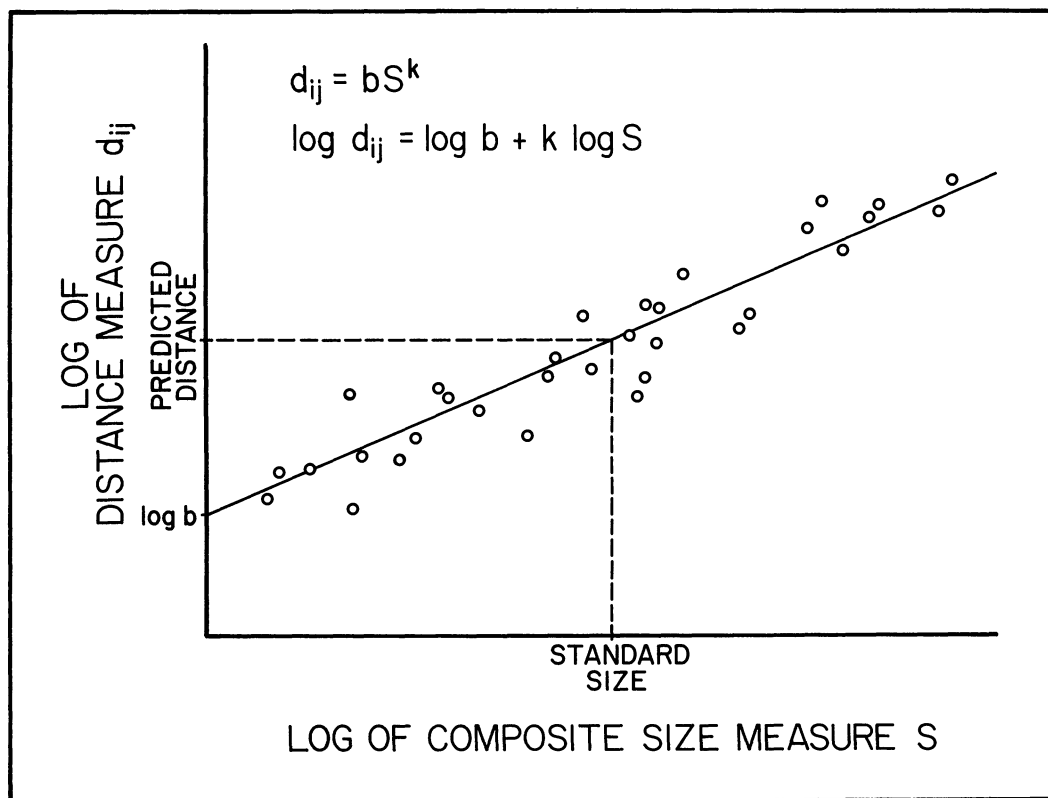
$$d_{ij} = bS^k$$

$$\log d_{ij} = \log b + k \log S$$

FIG. 10.—Illustration of a regression method for standardizing size. The logarithm of each truss measure (connecting landmarks i and j) is regressed on some composite log-scale measure of body size S. The predicted length at a standard size is transformed back to the original scale of measurement. The entire set of predicted truss-element lengths is then used to reconstruct the averaged form.

ation by univariate regression has some disadvantages (Humphries et al., 1981), primarily because the resulting character adjustments are not necessarily optimal for, or even concordant with, trends of character covariation. Nevertheless, because the mapping algorithm will effectively recover the most reasonable average configuration and the strain statistics provide measures of the discord, univariate regression is appropriate in this context. The procedure results in a *composite geometric form* which represents the average form of a sample of individuals, standardized to some arbitrary body size, which can be drawn, examined, and compared with others.

Because the standard size chosen for the composite form is arbitrary, it may be varied to allow a direct comparison of body shapes at different sizes. Illustrated in Figure 11 are averaged forms for two closely related species of western North American sculpins, *Cottus pitensis* and *Cottus klamathensis*, at three different composite body sizes scaled in terms of the original unit of measurement (millimeters). A change in standard size alters all distance measures simultaneously, each at its own allometric rate. By "growing" the form of each species we display the composite change in form resulting from allometric influences of body size on shape.

We may depict the same shape changes in a different style of diagram by indicating for each truss element its within-group allometric coefficient with respect to
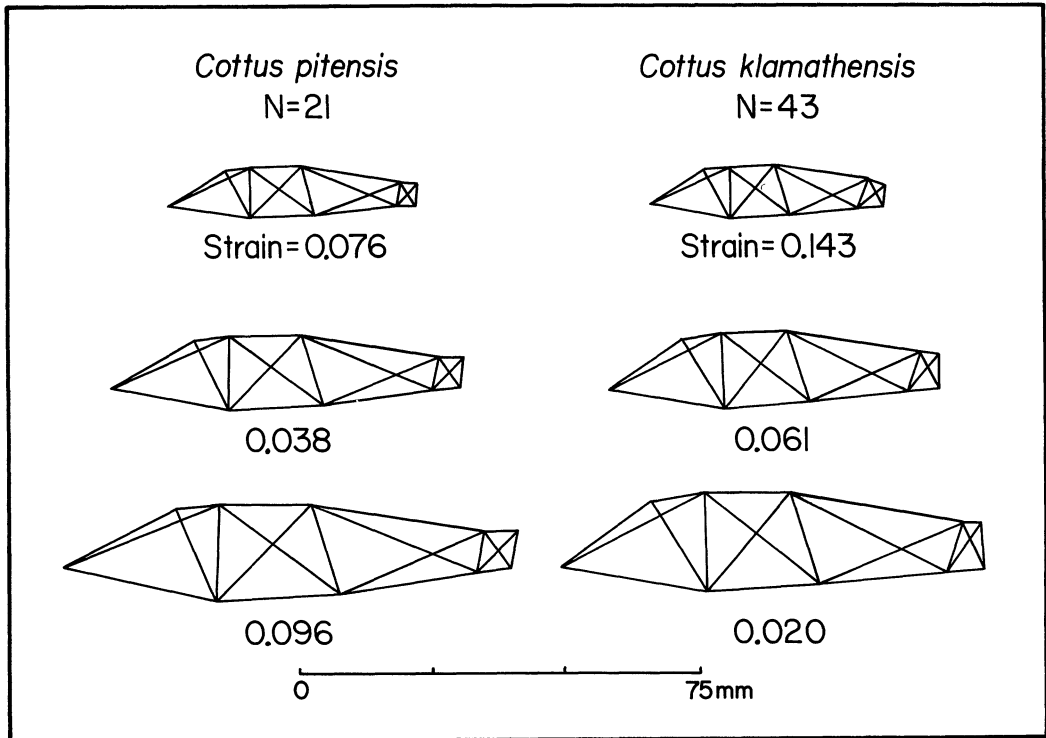
FIG. 11.—Averaged forms for two species of *Cottus* at three different composite body sizes. Reconstructed forms are scaled in millimeters and represent body lengths of approximately 45, 64, and 81 mm.

overall body size (Humphries, in press; Chernoff and Miller, in press). These individual regression coefficients are the same as the loadings of the log-distances on this estimate of body size. We first scale them to a mean-square of 1.0. A variable growing isometrically with "size" (as defined by this set of variables) then bears a loading of 1.0. Loadings greater than 1.0 denote positive allometry while loadings less than 1.0 indicate negative allometry. Compared in this way (Fig. 12), the growth patterns of *Cottus klamathensis* and *C. pitensis* are notably different. (We will draw their comparison in Fig. 15 below.) Because the allometric coefficients are derived from a sample of individuals each measured once, they do not strictly represent patterns of growth (Gould, 1966). However, these patterns of multivariate allometry among individuals seem to account for changes in form

within samples and for differences in shape between taxa.

INTERGROUP COMPARISONS

Because the truss configuration covers the form more completely than do traditional character sets, it can considerably enhance discrimination among groups when the differences are not specific to a very few structures. Figure 13 contrasts the effectiveness of the two styles of measurement. The truss network again consists of the 21 midsagittal distances of Figure 4B. The other character set consists of the measurements shown in Figure 1, with the exception of pelvic-fin length, which is not inside the body outline, and the two measures of body width, which are not measured in the lateral projection—a total of 19 characters. Both sets of distances were measured on the same 64 specimens, 21 of *Cottus pitensis*
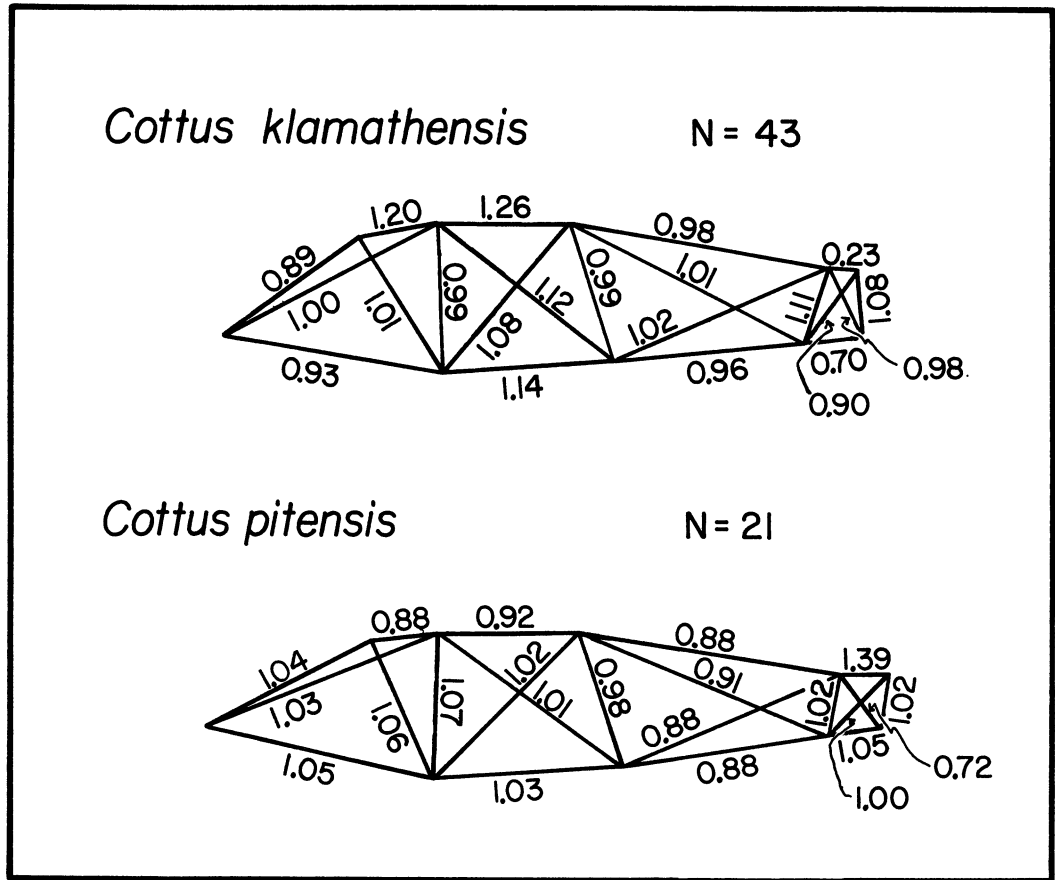
· FIG. 12.—Within-group patterns of allometry for two species of *Cottus*. Allometric coefficients are the loadings (scaled to a mean-square of 1.0) of the truss measures on the first within-group principal component (covariance matrix, log-transformed data). Coefficients greater than 1.0 indicate positive allometry; those less than 1.0 indicate negative allometry.

and 43 of *C. klamathensis*. Five truss distances (3–4, 4–6, 5–7, 6–8, and 9–10) are shared with the traditional set; the distance 1–3 (that from the snout to the base of the pelvic fins) is not equivalent to the length from snouth to pectoral-fin base.

In separate principal component analyses, each pooling the two samples, the sheared second component, independent of size within groups (Humphries et al., 1981), is plotted against the first among-group principal component, which expresses mostly body size. Had we instead used a discriminant function to maximally separate the groups, we would have sacrificed the interpretability of the coef-

ficients (Campbell and Atchley, 1981; Oxnard et al., 1981). The scores on the sheared second component are highly correlated with those on the discriminant axis when the mean body size of the two groups is the same; however, the loadings on the sheared component are directly interpretable with respect to shape differences.

The loadings (Table 2) on the sheared second component of the truss data may be depicted directly on the network (Fig. 13C). The primary contrasts on the form are between positive longitudinal loadings, which indicate that for a given body size *C. pitensis* is relatively longer than
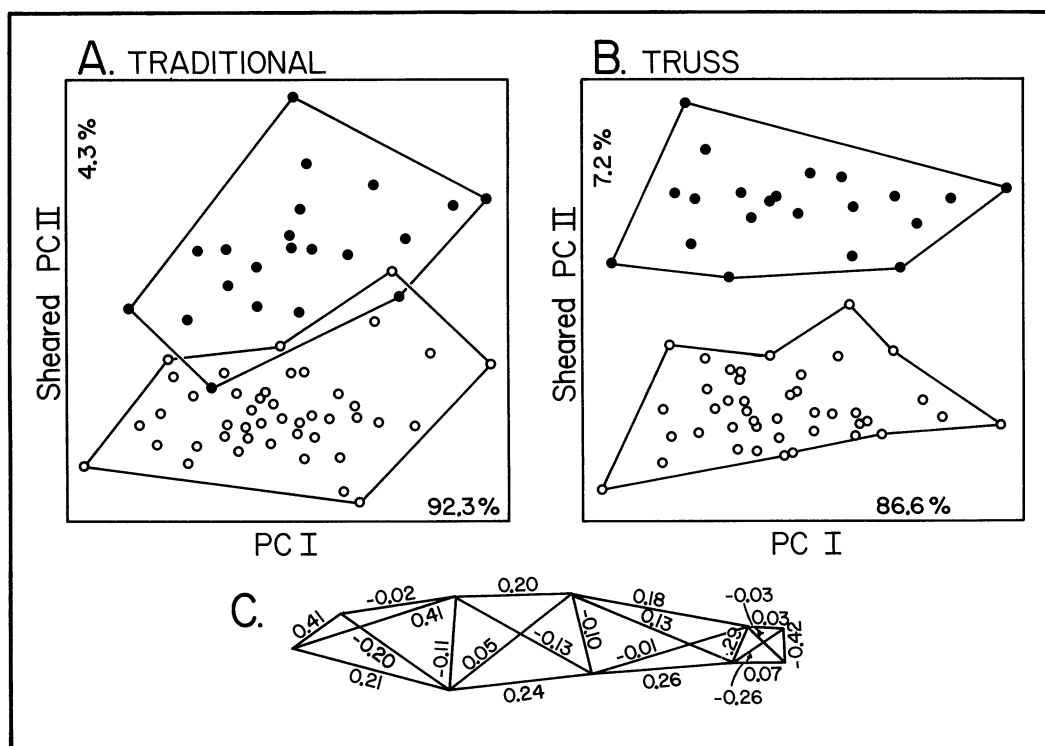
FIG. 13.—Principal component analyses of two character sets on the same 64 specimens, 21 of *Cottus pitensis* (solid circles) and 43 of *Cottus klamathensis* (hollow circles). Characters are listed in Table 2. (A) Scatter plot of the sheared second principal component, calculated with the algorithm of Humphries et al. (1981), against the first among-group principal component of the traditional data set. (B) Scatter plot. of the sheared second principal component against the first among-group principal component of the truss data set. (C) Loadings of the truss measures on the sheared second component, describing the differences in size-free shape between *C. klamathensis* and *C. pitensis*. In (A) and (B), percentages along the axes represent variances explained by the original principal components.

C. klamathensis, and some of the negative diagonal and vertical loadings, which show *C. klamathensis* to be relatively deeper bodied than *C. pitensis*. In particular, the distances with highest negative loadings are those that are posteroventrally oblique in the head and midbody regions and posterodorsally oblique on the peduncle.

The truss network provides better discrimination between the two species than does the traditonal system. This demonstrates the potential fallacy of concluding that two groups are the same simply because an analysis based on a particular data set has not shown them to be different. The principal directions of shape difference are largely unsampled by the tra-

ditional character set except by the measures of maximum body depth and minimum peduncle depth, both of which have high negative loadings on the sheared component (Table 2). Thus, the truss method improves shape discrimination by enforcing the use of cross measurements often not considered.

Composite mapped forms are particularly conducive to biorthogonal analyses, which allow quantitative comparison of forms by revealing the principal directions of shape difference (Bookstein, 1978, 1982). The procedure models the difference in shape as a smooth deformation of a grid in the manner of Thompson (1961), a mapping of one form onto the other in accordance with biological homology. We

TABLE 2. PRINCIPAL COMPONENT LOADINGS OF TWO DIFFERENT CHARACTER SETS FOR SAMPLES OF *COTTUS KLAMATHENSIS* (N = 43) AND *C. PITENSIS* (N = 21).

Traditional character set[1]			Truss character set[2]			
Character	I	Sheared II	Cell	Character	I	Sheared II
			1	1–2	0.21	0.41
				1–3	0.22	0.21
				2–3	0.23	−0.20
				1–4	0.22	0.41
				2–4	0.23	−0.02
Maximum body depth	0.24	−0.18	2	3–4	0.22	−0.11
				3–5	0.24	0.24
				4–5	0.23	−0.13
				3–6	0.23	0.05
Length first dorsal fin base	0.23	0.14		4–6	0.24	0.20
Length anal fin base	0.20	0.14	3	5–6	0.21	−0.10
				5–7	0.20	0.26
				6–7	0.21	0.13
Length second-dorsal fin base	0.22	0.19		5–8	0.21	−0.01
				6–8	0.20	0.18
			4	7–8	0.23	−0.29
				7–9	0.19	0.07
				8–9	0.19	−0.03
				7–10	0.21	−0.26
				8–10	0.18	0.03
Least depth caudal peduncle	0.23	−0.65		9–10	0.23	−0.42
Head length	0.21	0.05				
Postorbital head length	0.22	0.11				
Eye diameter	0.17	−0.19				
Premaxilla length	0.25	0.25				
Maxilla length	0.26	0.21				
Length first-dorsal fin to caudal base	0.24	0.12				
Length second-dorsal fin to caudal base	0.23	0.18				
Length anal fin to caudal base	0.22	0.11				
Length snout to pectoral base	0.23	−0.04				
Length pectoral fin margin to caudal base	0.24	0.08				
Length depressed second-dorsal fin	0.23	0.28				
Length depressed anal fin	0.22	0.20				
Standard length	0.21	0.23				
Total length	0.23	0.28				

[1] The traditional characters include all measurements shown in Figure 1 except pelvic-fin length and the two measures of body width.
[2] The truss characters are those of Figure 4B, identified by the landmarks they connect.

describe the deformation by its effect on a particular coordinate system (the biorthogonal grid) whose elements intersect at 90° in both forms. The method reduces changes in shape to changes in size of homologous grid elements; since all angles with the grid are forced to remain unchanged in the deformation, the observed change in shape is described by the dilatations (relative stretch or shrink)

of homologous grid elements. Because the deformation is assumed to be smooth, the dilatations increase or decrease in gradients along the curves of the grid.

The pattern of shape difference between *C. klamathensis* and *C. pitensis* is clearly revealed by a biorthogonal analysis of the transformation (Fig. 14). The outlines for this analysis are the averaged forms for the two samples, standardized to a common body size. The curves on the forms are an arbitrary sampling from the biorthogonal grid, and the differences in the grid patterns quantify the deformation required to transform one form into the other. The curves of the grid represent the primary directions of shape change at their points of intersection, the directions of greatest local expansion and contraction. The samples of dilatations along the curves are the ratios of homologous grid-element lengths at those points in the two forms. The ratio of lengths to the boundary in these two directions from any point is the mensural character which best discriminates the forms in reference to that point. Dilatations oblique to the axes of the grid may be interpolated according to the formula

$$\delta^2(\theta) = \delta_1{}^2\cos^2\theta + \delta_2{}^2\sin^2\theta$$

where θ is the angle between the oblique length-element and the principal axis of dilatation δ_1.

In our example, the directions of maximum elongation, with dilatations greater than 1.0, are along the more-or-less horizontal grid lines; the pattern indicates that *C. pitensis* is relatively longer than *C. klamathensis* at this standardized body size, but that the direction of elongation is not strictly longitudinal. Dilatations

along the somewhat vertical grid lines are all less than 1.0, showing that the change in shape from *C. klamathensis* to *C. pitensis* is a change from a deep-bodied to a shallow-bodied form. In particular, the greatest differences in body depth are posteroventrally oblique in the head, vertical in the midbody, and posterodorsally oblique in the tail. The pattern of deformation is concordant with the pattern of loadings of truss elements on the sheared second principal component (Fig. 13C). The traditional character set is consistent with this grid as well, but does not sample it so locally nor allow reconstruction of the predicted forms.

Because the species grow differently, the comparison of these forms is to some extent a function of the size standard chosen. The difference between the separate growth trends of the two species (Figs. 11 and 12) may itself be visualized by "growing" a truss according to the difference between the two sets of allometric coefficients. In this way we may deform the averaged 64-mm *C. klamathensis* (Fig. 15A, top) into the form (Fig. 15A, bottom) for which all truss elements are altered according to the differences in predicted allometric growth over a size change of 25% (64 mm to 51 mm). That is,

$$\log d_{iM} = \log d_{iK} + 0.25\Delta_i,$$

where d_{iK} is a truss distance on the averaged 64-mm *C. klamathensis* figure, d_{iM} is the distance of the same truss element on the modified form, and Δ_i is the corresponding difference between the allometric coefficients of the two species (Fig. 12). The resulting deformation, which is the effect of 13 mm of change in body length upon the shape comparison, may

$\rightarrow$

FIG. 14.—Biorthogonal analysis of the transformation from an averaged form of *Cottus klamathensis* to an averaged form of *Cottus pitensis* of the same body size. (A) An arbitrary sampling of curves from the biorthogonal grid, the elements of which intersect at 90° in both forms. The curves represent the primary directions of shape change, one of maximum stretch and one of maximum shrink, at their points of intersection. (B) A sample of dilatations along the grid lines describing the directions of maximum elongation. A dilatation of 1.10 indicates an increase of 10 percent. (C) A sample of dilatations along the grid lines describing the directions of maximum contraction. A dilatation of 0.90 indicates a decrease of 10 percent.
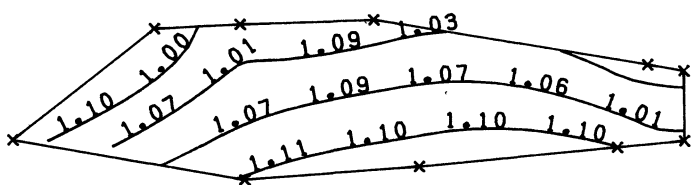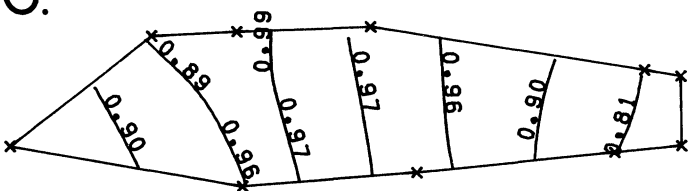
A.
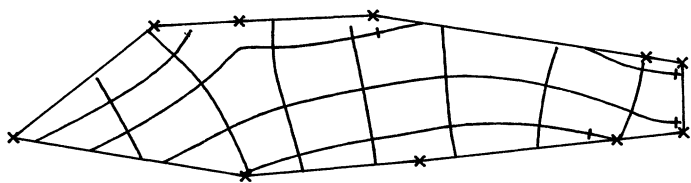
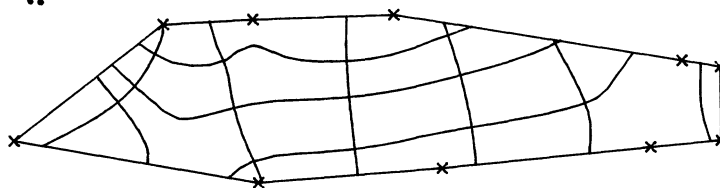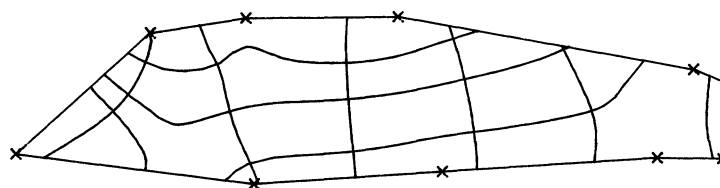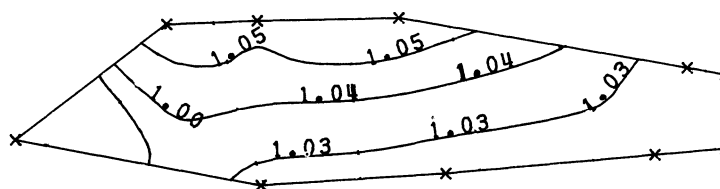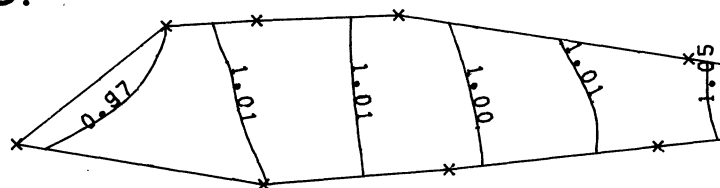Cottus klamathensis
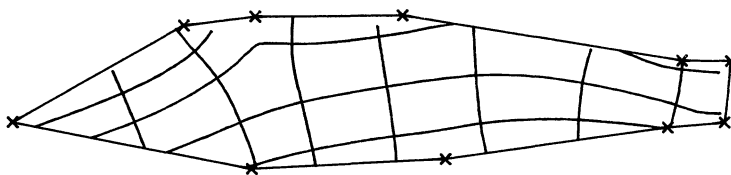
Cottus pitensis

B.

C.

A.

Cottus klamathensis

MODIFIED FORM

B.

C.

be expressed in a biorthogonal grid for the difference between the two trends (Fig. 15; compare with Fig. 12). The grid for comparison of the fishes at any size other than that used in Figure 14 is the product of the grid in Figure 14 by an appropriate multiple of the one in Figure 15. The rules for this manipulation are indicated in Bookstein (1982).

In its original formulation, the biorthogonal method was limited to comparisons of single pairs of forms, usually individuals taken to be representative of populations. Application of the truss network overcomes this limitation: the use of composite averaged forms is preferable to the use of single individuals in describing differences in shape among populations. In addition, because the truss systematically samples variation in many directions (every 45° in a square truss network), the set of ratios of homologous elements in two forms includes one dilatation within some moderate angle (in a square truss, 22.5°) of each principal direction of shape difference. The pattern of loadings on the truss, interpreted as a transformation, therefore provides an initial estimate of the biorthogonal description of the deformation.

CONCLUSIONS

The system of measurement and analysis we have described is based on two premises: first, that collections of anatomical points and the distance measures among them must be homologous from form to form, because only the biological homology of two configurations makes meaningful their scientific description and comparison; and second, that an adequate collection of measured distances should at least permit the reconstruction of the configuration of landmarks it purports to measure, since information is otherwise lost. The truss network is a geometric protocol which fulfills these requirements and offers several additional advantages. Its properties may be summarized as follows.

(1) Use of the truss network as a character set enforces systematic coverage across the form; in contrast, traditional character sets often provide highly uneven coverage.

(2) The truss exhaustively and redundantly archives the form; hence, the original configuration of landmarks may be reconstructed by relaxing the data so as to be coplanar and mapping the Cartesian coordinates of the landmarks.

(3) The least-squares mapping procedure is robust against moderate imprecision of measurement. The degree of measurement error in the data may be assessed graphically and by various indices characterizing the mutual lack of fit of the data.

(4) Forms may be standardized to one or more common reference sizes by regressing measured distances on some composite measure of body size and reconstructing the form using the distance values predicted at some standard body size.

(5) Principal components can be given geometrical interpretations. Component scores are measures of configuration, while loadings are descriptors of shape change.

(6) Composite mapped forms are suitable for biorthogonal analyses of shape differences between forms.

←

FIG. 15.—Biorthogonal grids for the effect upon the interspecific shape comparison, Figure 14, of a change in body length from 64 to 51 mm. (A) An arbitrary sampling of curves from the biorthogonal grid of the transformation from the averaged 64-mm *C. klamathensis* (above) to a modification of that form in which all elements are altered according to the between-species differences in predicted allometric growth over a size change of 64 mm to 51 mm. (B) A sample of dilatations along the grid lines describing the directions of maximum elongation. (C) A sample of dilatations along the grid lines describing the directions of maximum contraction.

Patterns of loadings derived from principal components of truss data are consistent with such analyses. By providing a systematic sampling of variation in many directions, the truss provides an initial estimate of the biorthogonal description of the deformation.

When this system of design and analysis is applied, the results of multivariate analyses may serve a dual role (Bookstein, 1982). (1) As patterns of loadings they may suggest measures which are optimal discriminators between groups, whether of growth stages or taxa. Because we do not know at the outset of a study what measures will ultimately emerge as the optimal characters for analysis and description, there is no point in presuming their identities (e.g., specific angles or proportions) from the beginning. We instead begin with a character set which is sensitive to variation in all directions, and later extract the best simple measures that characterize the observed differences. (2) As factors (Wright, 1954), the results of morphometric analyses give rise to expected values for all distances, expected values that change as "size" or "taxon" changes. A change in the factor may be interpreted explicitly as a transformation, a change of form in accord with biological homology. When the system of linear measures permits a full reconstruction of the original landmark configuration, the predicted values at any factor score will be an archive of a predicted form.

ACKNOWLEDGMENTS

REFERENCES

BOOKSTEIN, F. L. 1978. The measurement of biological shape and shape change. Lecture notes in biomathematics No. 24, Springer-Verlag, New York, 191 pp.

BOOKSTEIN, F. L. 1982. Foundations of morphometrics. Ann. Rev. Ecol. Syst., 13:451–470.

CAMPBELL, N. A., AND W. R. ATCHLEY. 1981. The geometry of canonical variate analysis. Syst. Zool., 30:268–280.

CHERNOFF, B., AND R. R. MILLER. In press. Mexican freshwater silversides (Pisces: Atherinidae) of the genus Archomenidia, with description of a new species. Proc. Biol. Soc. Washington.

GOULD, S. J. 1966. Allometry and size in ontogeny and phylogeny. Biol. Rev., 41:587–640.

HUMPHRIES, J. M. In press. Cyprinodon verecundus, n. sp., a fifth species of pupfish from Laguna Chichancanab, Mexico. Copeia.

HUMPHRIES, J. M., F. L. BOOKSTEIN, B. CHERNOFF, G. R. SMITH, R. L. ELDER, AND S. G. POSS. 1981. Multivariate discrimination by shape in relation to size. Syst. Zool., 30:291–308.

IMSL LIBRARY 1 REFERENCE MANUAL. 1977. 2 vols. International Mathematics and Statistics Library, Houston, Texas.

JARDINE, N. 1969. The observational and theoretical components of homology: a study based on the morphology of the dermal skull-roofs of rhipidistian fishes. Biol. J. Linn. Soc., 1:327–361.

JOLICOEUR, P. 1963. The multivariate generalization of the allometry equation. Biometrics, 19:497–499.

MIKHAIL, E. M. 1976. Observations and least squares. IEP-Dun-Donnelly, New York, 428 pp.

MOYERS, R. E., AND F. L. BOOKSTEIN. 1979. The inappropriateness of conventional cephalometrics. Amer. Jour. Orthodont., 75:599–617.

OLSON, E. C., AND R. L. MILLER. 1958. Morphological integration. Univ. Chicago Press, 317 pp.

OXNARD, C. E., R. GERMAN, AND J. E. MCARDLE. 1981. The functional morphometrics of the hip and thigh in leaping prosimians. Amer. Jour. Phys. Anthropol., 54:481–498.

ROHLF, F. J., AND J. W. ARCHIE. 1978. Least-squares mapping using interpoint distances. Ecology, 59:126–132.

SALMON, G. 1914. A treatise on the analytic geometry of three dimensions, 2 vols. (R. A. P. Rogers, rev. ed.). Longmans and Green, London, 470 + 334 pp.

SCHAEFFER, B. 1976. Practical aspects of homology recognition. Pp. 169–173, in Evolution, brain, and behavior (R. B. Masterton, W. Hodos, and H. Jerison, eds.). Erlbaum Associates, Hillsdale, New Jersey, 312 pp.

SKALAK, R., G. DASGUPTA, M. MOSS, E. OTTEN, P. DULLEMEIJER, AND H. VILMANN. 1982. Analytical description of growth. Jour. Theor. Biol., 94:555–577.

STRAUSS, R. E. 1980. Genetic and morphometric variation and the systematic relationships of eastern North American sculpins (Pisces: Cottidae). Dissertation, Pennsylvania State Univ., Univ. Park.

THOMPSON, D. W. 1961. On growth and form (J.

T. Bonner, ed.). Cambridge Univ. Press, London, 346 pp.

WRIGHT, S. 1954. The interpretation of multivariate systems. Pp. 11–33, *in* Statistics and mathematics in biology (O. Kempthorne, T. A. Bancroft, J. W. Gowen, and J. L. Lush, eds.). Iowa State College Press, Ames, 328 pp.