CONTOUR MAPPING AND DIFFERENTIAL SYSTEMATICS OF GEOGRAPHIC VARIATION

ROBERT P. ADAMS

Reprinted from SYSTEMATIC ZOOLOGY
Vol. 19, No. 4, December 1970
pp. 385–390
Made in United States of America

CONTOUR MAPPING AND DIFFERENTIAL SYSTEMATICS OF GEOGRAPHIC VARIATION

ROBERT P. ADAMS

Abstract

Adams, R. P. (Colo. St. Univ., Ft. Collins, Colo. 80521) 1970. Contour mapping and differential systematics of geographic variation. Syst. Zool., 19: 385–390.—Algorithms are presented for automatic contour mapping of characters and displaying of the composite differential of several characters by computer. An example is given using morphological and chemical characters from *Juniperus pinchotii* Sudw. (Cupressaceae). Several attributes of contour mapping of geographic variation are discussed. Differential systematics is shown to be of considerable interest in delimiting areas of speciation and/or ecotypic differentiation.

Computerized contour mapping of characters, when used with an appropriate statistical text, can be used to show, graphically, regional trends within a taxon. Correlation of these trends with ecological factors may be readily visualized such that further fieldwork may be concentrated on the investigation of these areas. Differential systematics can be a powerful tool for the populational worker in revealing areas of rapid differentiation using not one, but several characters simultaneously. These characters can be of several kinds, such as morphological, chemical, cytological, etc., in order to obtain a more representative sample of the organism's genome.

This paper is concerned with the presentation of algorithms for determining contour maps and the composite differential of several characters automatically by computer. An example is given using both morphological and chemical characters from trees of *Juniperus pinchotii* Sudw. (Cupressaceae) from a study by Adams (1969).

CONTOUR MAPPING

Recently Fisher (1968) has studied animal distributions by plotting the sixth degree trend surfaces of five centroid factors (obtained from correlations between taxa distributions). The value of these contour surfaces lies in the fact that they may be easily correlated with environmental and/or geological variables, and intercorrelations between these surfaces (whether they be morphological characters, chemical characters, abstract factors, etc.) are readily visualized by the reader.

Although Sokal (1965) warns that the drawing of isophenes (contours) ". . . is still largely a subjective procedure . . ." and "given the same basic data, different observers may arrive at quite differently drawn isophenes," these objections can be overcome by the use of standard contouring algorithms on automatic plotting devices controlled by computers. A second objection of Sokal (1965) is that "maps generally give the impression of being more reliable than they are in fact." This aspect can be countered by placing the results of a statistical test of the means (such as the SNK test) next to the contour map as shown by Sokal and Rinkel (1963).

The basis of contour mapping is a set of mesh points over a geographical area for which the three-dimensional surface is desired. Unfortunately, biological samples can seldom be taken at the regular mesh points of a grid since most taxa do not have a continuous distribution in both time and space. Although many computation centers have programs for drawing contour maps, they usually require a large matrix of points which are equally spaced. Thus, one is confronted with a collection of "semi-random" points from which all of the mesh points (matrix values) must be estimated. The approximation presented in this paper is only one of several which could be used (notably sixth degree or higher polynominals). The mesh point approximation is

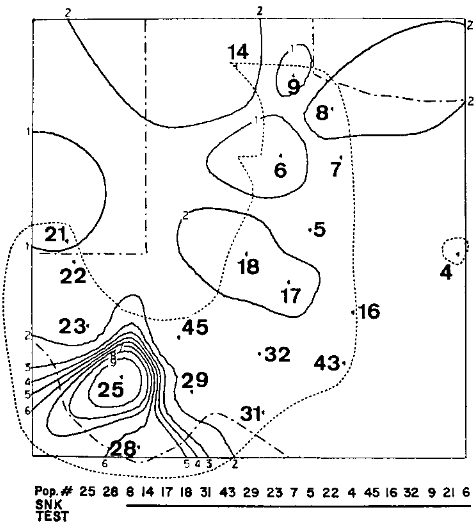

FIG. 1.—Contour map of the percentage of citronellal found in natural populations of *J. pinchotii*. Contour values are the percent of citronellal in the total oil extracted from populational samples of *J. pinchotii*. Contour symbols and values are: 1 = .47%; 2 = .97%; 3 = 1.46%; 4 = 1.96%; 5 = 2.46%; 6 = 2.95%; 7 = 3.45%; 8 = 3.94%; 9 = 4.44%.

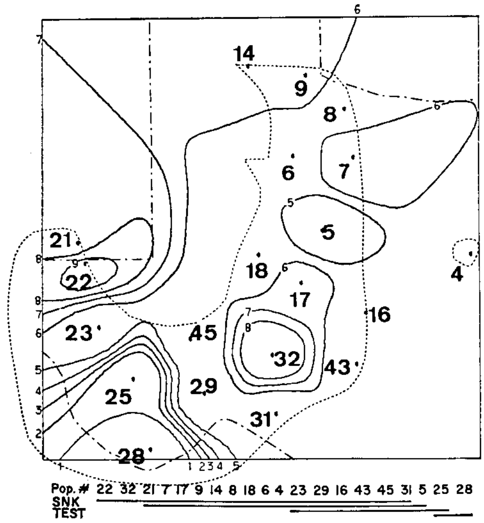

FIG. 2.—Contour map of the female cone color in populations of *J. pinchotii*. Colors are: 2.0 = rose; 3.0 = red/brown; and 4.0 = yellow/brown. Contour symbols and values are: 1 = 2.26; 2 = 2.42; 3 = 2.59; 4 = 2.76; 5 = 2.91; 6 = 3.08; 7 = 3.24; 8 = 3.40; 9 = 3.56.

that of the weighted average of all of the data points, with the weights (W) being a function of the inverse of the distance of each data point from the mesh point to be estimated. Empirical data indicate that a weighting of distance (D) raised to three times the logarithm of the distance plus one (i.e., $W = D^{3 \cdot \log(D+1)}$) yields a very close approximation to the true surface.

That is: If X_{ij} is a mesh point whose value is to be determined, R_k is the kth of n data points, and D_{ijk} is the distance from R_k to X_{ij};

Then: $$W_{ijk} = (D_{ijk})^{3 \cdot \log(D_{ijk}+1)}$$

And: $$X_{ij} = \frac{\sum_{k=1}^{n} [R_k \cdot (1/W_{ijk})]}{\sum_{k=1}^{n} (1/W_{ijk})}.$$

Previous experiments (Adams, unpublished), using the distributions of myrcene and carvone in populations of *Juniperus ashei* Buch., have indicated that the contour maps obtained by use of these weights are in good general agreement with 3-dimensional models of the population means of these compounds. In a test of geological data from O'Leary, Lippert, and Spitz (1966), this approximation fit their data much more closely than did sixth degree polynomials. This approximation tends to deviate from the "true" surface in planar areas which are adjacent to areas with large gradients. These deviations are small and in general seem to have little effect on the overall trend of a contour map.

After the mesh points are calculated, a standard contouring routine can be used to draw the contour map (Some installations utilize a 35 mm film system in which the maps are displayed on a cathode ray tube and photographed. Prints are then made from the 35 mm film).

Figures 1 and 2 show the contour maps of the percent of citronellal (a volatile terpene) and the relative female cone color (2.0 = rose, 3.0 = red/brown, 4.0 = yellow/brown) in populations (5 or more observations per population) of *Juniperus pinchotii*.

The contour levels are the midpoint values of the n equal intervals of the range of the function to be contoured. The dot-dash line and the dotted line show the boundaries of the state of Texas and the distribution of *J. pinchotii*, respectively. These contour maps are based on 3721 (61×61) mesh points approximated from the original 20 data points (it should be noted that extrapolation to a large number of mesh points was necessary to obtain smooth contour lines and does not imply more accuracy from the original 20 data points). A summary of the SNK tests is given at the bottom of each map (any means underscored by the same line are not highly significantly different: any two not underscored by a common line are highly significantly different).

From Figure 1, it appears that most of the differentiation is in the area of populations 25 and 28, with the concentration of citronellal increasing in these populations. Although some slight variation is evident in the other populations, the SNK test reveals no highly significant differences except: 1) between populations 25 and 28; 2) between the rest of the populations and populations 25 and 28. Figure 2 indicates much the same pattern on the basis of female cone color, except populations 22 and 32 are tending to be divergent from the rest of *J. pinchotii* and in the opposite direction from populations 25 and 28 (i.e., toward more yellow/brown cones as opposed to the tendency toward rose-colored cones in populations 25 and 28). In general, cone color appears to be more variable than the percentage of citronellal. A more detailed discourse of the extent, possible causative factors, and the significance of this and other differentiation (involving 19 morphological characters and over 100 chemical (terpenoid) characters) in natural populations of *J. pinchotii* will be presented at a later time.

DIFFERENTIAL SYSTEMATICS

Differential systematics was proposed by Womble (1951) and defined as "a methodology for synthesizing multiple measurements, indices, and frequencies into a composite variable, the *systematic function*, which, for all loci, evaluates the average change with distance of a total reality." Even when contour maps are constructed as previously mentioned, one is faced with trying to summarize the total trend of all the characters taken together in order to assess both the amount and direction of differentiation within a taxon. Differential systematics will allow one to sum up the rates of change with distance (differential) of several characters to show zones of differentiation within a taxon.

Apparently only a few biologists (e.g., Hagmeier, 1958) have actually used this approach, presumably due to the laborious procedures outlined by Womble to obtain the differential of a set of characters. With the advent of modern computing methods this drudgery can now be eliminated. Since Womble's proposal has been dormant for so long, I will restate the basic principles.

Figure 3 shows the synthesis of two characters (traits) into a composite derivative. Traits g_1 and g_2 of a taxon show clines of opposite slope in part (*a.*). Averaging them only confuses the possible taxa *X*, *Y*, and Z. Trait g_1 and the absolute value of its derivative (dg_1/ds = rate of change of trait g_1 with respect to distance) are shown in part (*b.*). Likewise trait g_2 and the absolute value of its derivative are shown in part (*c.*). Part (*d.*) illustrates the absolute derivatives averaged as a composite derivative whose peak values separate the three taxa *X*, *Y*, and Z. When several characters (traits) are treated in this manner the areas in which there are only random changes will tend to accumulate a more or less uniform value of the composite derivative. If there is a region with rapid changes in several characters, the sum of these changes (absolute value of the derivatives) will be additive and accumulate much more rapidly. If these differentials are then contoured, the areas of high differentials will appear as "mountain ranges" (in analogy to topographic maps) separating divergent populations.

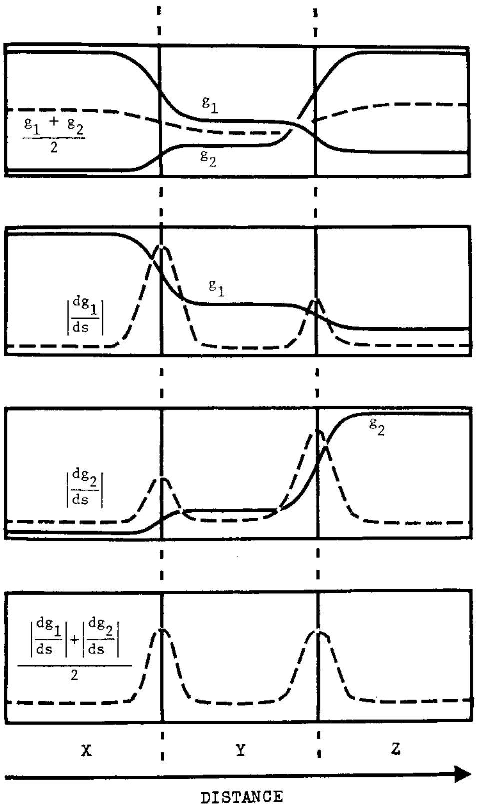

FIG. 3.—Profiles demonstrating, somewhat diagrammatically, the synthesis of multiple variable traits into a composite derivative which separates the three taxa *X*, *Y*, and *Z* (adapted from Womble, 1951).

These areas may then be compared to standard rainfall maps, topographic maps, etc., in assessing the causative factors involved.

The algorithm for computation of the composite differential involves the use of the matrix of grid mesh points generated by the contour program. If one considers the matrix generated by the contour program as a table of Z coordinates (in the *X*, *Y*, *Z* coordinate sense) and with n^2 (i.e., $n \times n$) elements, then there are $n^2 - 2n + 1$ $(n-1 \times n-1)$ squares determined by 4 mesh points (such as *a*, *b*, *c*, and *d* as illustrated in Figure 4). The mesh points *a*, *b*, *c*, *d* are the interpolated values of some trait g_i (i.e., $a-c=$ the value of g_i at point *a*,

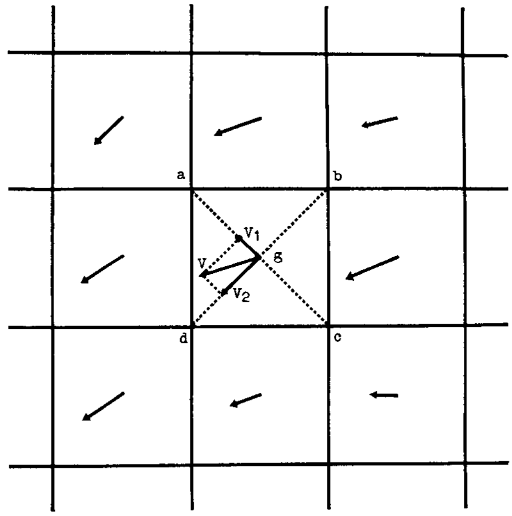

FIG. 4.—A hypothetical grid section with the slope vectors at the midpoints of each square. Slope vector *V* at point *g* is resolved into its two components, V_1 along the *ac* axis and V_2 along the *bd* axis. Points *a*, *b*, *c*, and *d* are four of the mesh points generated by the contouring program.

minus the value of g_i at point *c*). When *n* is large, these squares approximate a plane surface, such that the slope of the plane at its center point, *g*, is some vector, *V*. A hypothetical grid section is shown in Figure 4 with the slope vectors indicated at the center of each square. The slope vector can be resolved into 2 components: one component, V_1, along the *ac* axis and the other component, V_2, along the *bd* axis.

Let:

$$G = \frac{\sum_{i}^{n} |dg_i/ds|}{n} = \text{average differential of all } n \text{ characters}$$

$D =$ relative distance from *a* to $b = ab = bc = cd = da$.

$g =$ midpoint of the square determined by points, *a*, *b*, *c*, *d*.

$V =$ slope vector at $g \doteq dg/ds$.

$V_1 =$ component of V along the *ac* axis.

$V_2 =$ component of V along the *bd* axis.

Then:

$$V^2 = V_1^2 + V_2^2$$
$$V = (V_1^2 + V_2^2)^{1/2},$$

Where:

$$V_1 = (a-c)/ac;\ \text{since } ac = (ab^2 + bc^2)^{1/2},$$
$$V_2 = (b-d)/bd;\ \text{since } bd = (bc^2 + cd^2)^{1/2}.$$

Then:

$$ac = (D^2 + D^2)^{1/2} = D \cdot 2^{1/2}$$
$$bd = ac = D \cdot 2^{1/2}.$$

Thus:

$$V_1 = (a-c)/(D \cdot 2^{1/2})$$
$$V_2 = (b-d)/(D \cdot 2^{1/2})$$
$$V = \{[(a-c)/(D \cdot 2^{1/2})]^2 + [(b-d)/(D \cdot 2^{1/2})]^2\}^{1/2}$$
$$= [1/(D \cdot 2^{1/2})] \cdot [(a-c)^2 + (b-d)^2]^{1/2}.$$

Since D is only relative, let $D = (1/2)^{1/2}$.

Then: $V = [(a-c)^2 + (b-d)^2]^{1/2} \doteq dg/ds$.

Thus, given the matrix of mesh points (Z coordinates), one can approximate the derivative of that surface by the formulation above.

In addition to the calculation of the new matrix of slope vectors, each character is scaled between 0 and 1 before the contour mesh points are determined so each character will have a basis for comparison. The question of assigning different weights to different characters in the computation of the composite differential has not been resolved but one might wish to weight more heavily those characters that have more significant differences (SNK test). Another possibility is assigning weights proportional to the variability of each character, as suggested for numerical taxonomy by Farris (1966) and Flake and Turner (1968). The contour levels of the differential function are slightly different from those used in contouring the original variables, in that the range of the differential is divided into $n + 1$ intervals but the lowest interval is ignored. The contours are the midpoint values of the largest n intervals. This tends to eliminate some of the background "noise" accumulated in summing the differentials.

The composite derivative of cone color and the percentage of citronellal is shown in Figure 5. Notice that the areas of most rapid changes (highest contour levels) are between populations 25–23, 25–45, 25–29, 28–29, and to a lesser degree between 25–28, 22–23, and 21–22. The moderate changes in cone color are reflected in the small differential between populations 32–29 and 32–31. Areas of rapid changes (high contour levels for the differential function) may then be correlated with geology, climate, and other variables which one might suspect to be influential in ecotypic and/or genotypic variation within a taxon. For instance, the area containing populations 25 and 28 is a mountainous region of volcanic origin approaching the southwestern limit of the range of *J. pinchotii.* A detailed examination of the differentiation within *J. pinchotii* is beyond the scope of this paper and is covered more fully elsewhere (Adams, 1969; Adams, in preparation).

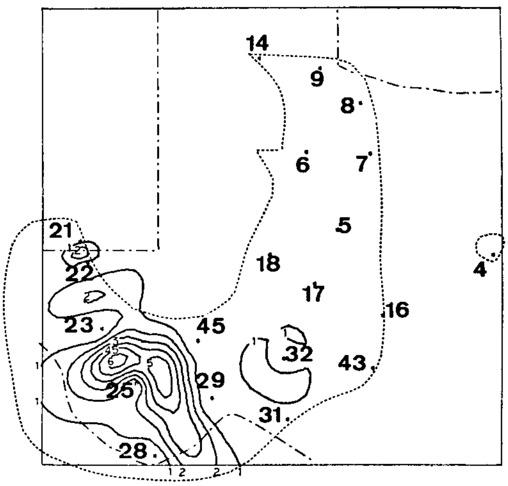

FIG. 5.—The composite differential of the percent of citronellal and female cone color in natural populations of *J. pinchotii.* The contour values are the average differential of the two characters. Contour symbols and values are: 1 = .038; 2 = .064; 3 = .089; 4 = .115; 5 = .140; 6 = .165.

CONCLUSION

Although only a few applications are enumerated here, others will undoubtedly come to mind with other research problems in systematics.

Differential systematics and contour mapping may also prove useful to ecologists, since the variables need not be characters of an organism but might be species' densities within different associations, percentages of the fauna and flora with certain attributes (such as the percentage of succulent plants), etc.

The applicability is perhaps best expressed by Womble (1951) in the closing paragraphs of his theoretical treatise on differential systematics "It is applicable to the analysis of any type, number, and combination of quantifiable traits, provided only they are related as to means of variation, diffusion, and interaction."

REFERENCES

ADAMS, R. P. 1969. Chemosystematic and Numerical Studies in Natural Populations of *Juniperus*. Ph.D. Thesis. The University of Texas, Austin, Texas.

FARRIS, J. S. 1966. Estimation of Conservatism of Characters by Constancy within Biological Populations. Evol., 20(4):587–591.

FLAKE, R. H. AND B. L. TURNER. 1968. Numerical Classification for Taxonomic Problems. J. Theoret. Biol., 20:260–270.

FISHER, D. R. 1968. A study of Faunal Resemblance Using Numerical Taxonomy and Factor Analysis. Syst. Zool., 17(1):48–63.

HAGMEIER, E. M. 1958. Inapplicability of the Subspecies Concept to North American Marten. Syst. Zool., 7(1):1–7.

O'LEARY, M., R. H. LIPPERT, AND O. T. SPITZ. 1966. FORTRAN IV and MAP Program for Computation and Plotting of Trend Surfaces for degrees 1 through 6. Computer Contribution 3. State Geological Survey. The University of Kansas, Lawrence, Kansas.

SOKAL, R. R. 1965. Statistical Methods in Systematics. Biol. Rev., 40:337–391.

SOKAL, R. R. AND R. C. RINKEL. 1963. Geographic Variation of the Alate *Phemphigus populi-transverus* in Eastern North America. Univ. Kansas Sci. Bull., 44:467–507.

WOMBLE, W. H. 1951. Differential Systematics. Science, 114:315–322.

Department of Botany and Plant Pathology, Colorado State University, Ft. Collins, Colorado 80521.