Insights on Why Graphic Correlation (Shaw's Method) Works

Author(s): Lucy E. Edwards

Source: *The Journal of Geology*, Sep., 1984, Vol. 92, No. 5 (Sep., 1984), pp. 583-597

Published by: The University of Chicago Press

Stable URL: https://www.jstor.org/stable/30062441

INSIGHTS ON WHY GRAPHIC CORRELATION (SHAW'S METHOD) WORKS[1]

LUCY E. EDWARDS
U.S. Geological Survey, 970 National Center, Reston, VA 22092

ABSTRACT

In 1964 A. B. Shaw presented a method of correlating fossiliferous sedimentary rocks based on interpretation of graphic plots of first- and last-occurrences of taxa. Because there is no way to determine the true total ranges of fossil taxa, it is instructive to test the accuracy of the method using hypothetical datasets. The dataset used here consists of 16 taxa in six sections with differing known rates of rock accumulation. In all graphs, a single straight-line correlation was a reasonable interpretation. The resulting ranges after the first and third rounds of compositing reproduce the "true" ranges but with small errors. Slight errors in the positioning of individual correlation lines are more likely to lengthen ranges artificially than to shorten them. Shaw's method works well because, whereas actually sampled ranges will be shorter than true ranges, errors in correlation will be likely to extend some ranges. This or any exercise using simulated data is useful only if the hypothetical situation resembles real geologic situations and if insights derived from the hypothetical dataset provide insights into real situations. The method is only as good as the available data.

INTRODUCTION

In his 1964 work, A. B. Shaw introduced the use of a conventional two-axis graph as a means of expressing time-equivalence between stratigraphic sections. By the intrinsic nature of correlatable sections, an infinite number of points of exact equivalency exist. When plotted on a graph, these points form what Shaw terms the Line of Correlation (LOC). The LOC is always there; the geologist's task is to find it.

Shaw's method relies on the observed lowest and highest occurrences of fossil taxa as indicators of geologic time. To illustrate and test his method, Shaw (1964) applied fossil data from eight sections in the Cambrian Riley Formation (Palmer 1955). The results thus produced were in close agreement with Palmer's original determinations. More recently, other workers have applied the method to actual fossil data to produce reasonable and informative results (see, for example, Miller 1977*a*, 1977*b*; Sweet 1979*a*, 1979*b*; Hazel et al. 1980; Murphy and Berry 1983).

One may test the *precision* of Shaw's method by comparing its results with those of previous, more conventional studies (as Shaw did with Palmer's data). Such comparisons, however, do not test the method's *accuracy,* the correctness or degree of conformity with truth.

Because there is no way to determine the true total ranges of fossil taxa, it is instructive to use a hypothetical dataset that includes assorted complications. Here, "true" total ranges are known, and the results produced using Shaw's method can be tested for accuracy. An exercise such as this is useful only if insights derived from the hypothetical dataset provide insights into real geologic situations.

In the following sections, I will summarize briefly the mechanics of Shaw's method (the reader should refer to the latter half of his book for complete details), enumerate and discuss what I feel are the underlying assumptions, and then comment on the accuracy of the method in the context of an example using a hypothetical dataset.

The hypothetical dataset has been designed to be illustrative. It consists of only 16 taxa, the sampling intervals are rather coarse, and stratigraphic complications have been included. Simply put, the conditions are less than ideal.

The design of the dataset is deliberate. If every species showed the same stratigraphic range in every section and if every section were completely sampled, interpretation would be straightforward. Any correlation method would work well. I felt the method should be examined under adverse conditions. In addition, the adverse conditions of the illustrated example should alert the bio-

[1] Manuscript received November 29, 1983; revised April 25, 1984.

[JOURNAL OF GEOLOGY, 1984, vol. 92, p. 583–597]

0022-1376/84/9205-007$1.00

stratigrapher to possible pitfalls in real situations.

This brief summary and short exercise are not intended to represent a complete analysis of Shaw's method. They are only a first step. The purpose of the present paper is twofold: to alert potential users of the method to some of its strengths and weaknesses, and to encourage others to perform similar experiments.

THE METHOD

The logic behind Shaw's method depends on locating time-equivalent levels within pairs of sections (finding a line of correlation), and on using the expression of equivalency (equation of the line of correlation) to express position in either section relative to position in the other section. For more than two sections, information from previously plotted sections is combined into a composite before each additional section is compared. A conventional two-axis graph allows visual evaluation of possible interpretations. Several rounds (cycles of interpretation) are usually used to produce the best model. The method is only semi-objective because alternative interpretations are always possible, and judgment (geologic, paleontologic, taxonomic, biologic, and stratigraphic) must be applied.

Figure 1 illustrates the procedures involved in Shaw's method of graphic correlation. The first step is the compilation of the raw data. Figure 1*a* shows the observed ranges of three taxa (A, B, and C) in each of three sections (striped, plain, and stippled). Next, the order in which the sections will be treated is selected. A well-sampled and unfaulted section containing many species at many levels is chosen as the reference section. Ideally, this section will be complete or nearly complete. Second and later sections are chosen in such a way that the most complete, most fossiliferous, and best-sampled sections are considered before sections that provide less information. In figure 1, the order of compositing is plain (reference section), striped, stippled. In figure 1*b,* the striped section has been plotted on a two-axis graph against the plain section. In Shaw's original work, the reference section (and later, the composite) was plotted on the horizontal axis and the individual sections were plotted on the vertical axis. Some later workers (e.g., Murphy and Edwards 1977; Hazel et al. 1980) plotted the reference/composite on the vertical axis. The choice of axes is immaterial to a compositor familiar with the basics of plane geometry. In the following discussion, I will refer to the composite axis and to the individual-section axis; the reader may translate these phrases

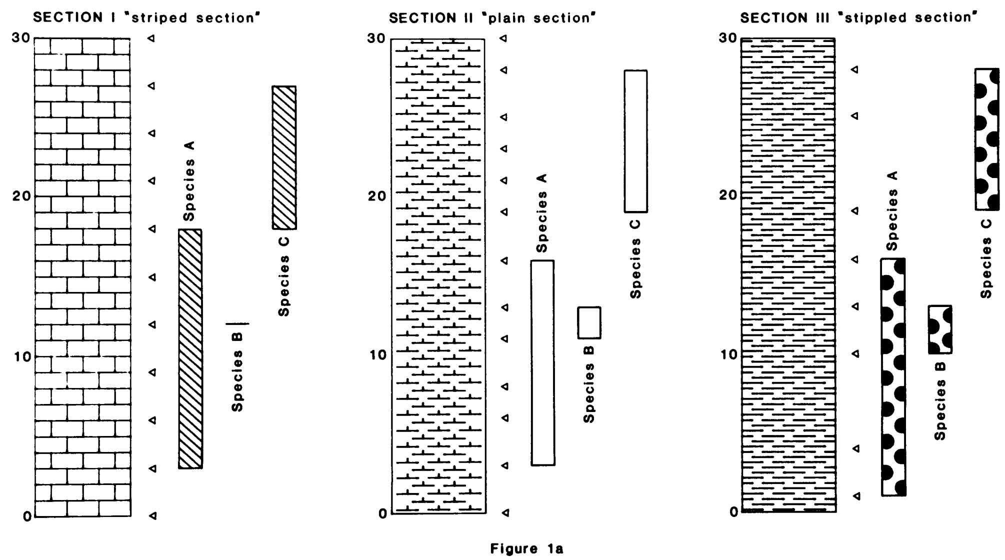

FIG. 1*a*.—The three sections (striped, plain, and stippled) and the observed ranges of Species A, B, and C in each of these sections. Sample locations are denoted by the positions of arrows.

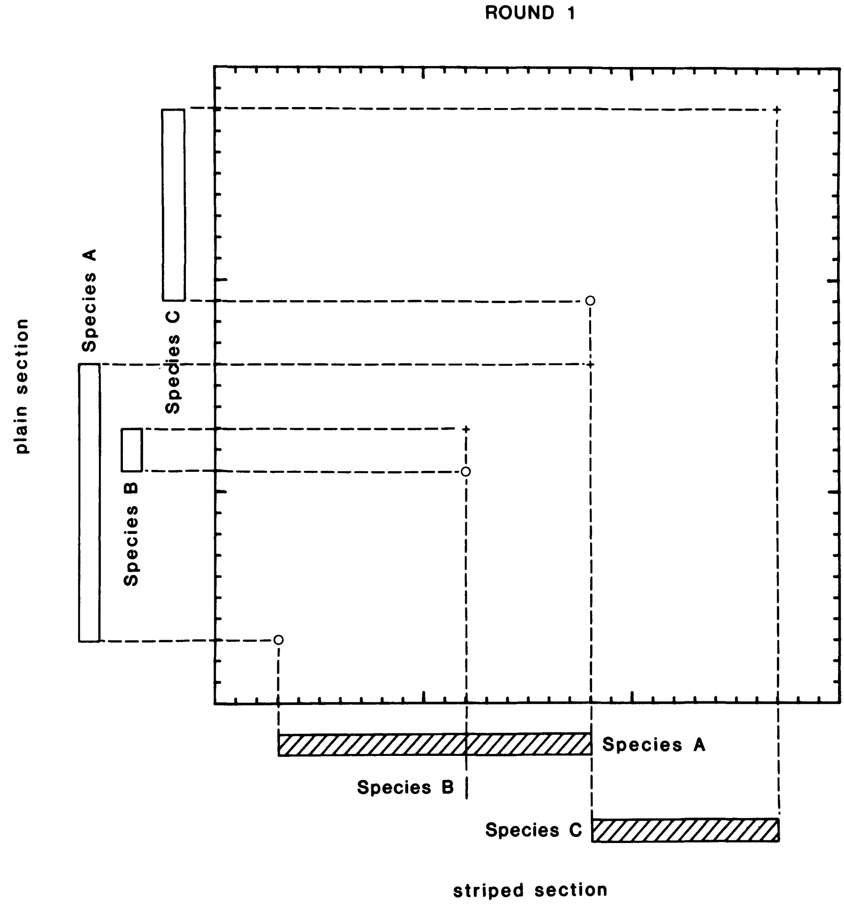

FIG. 1*b*.—The ranges of Species A, B, and C in the striped section (horizontal axis) plotted against the ranges of Species A, B, and C in the plain section (vertical axis). Within the graph, points are plotted according to the species-occurrence levels on the horizontal and vertical axes. "o" denotes first-occurrence events, "+" denotes last-occurrence events.

into the vertical and horizontal axes as desired. For convenience, the oldest material in all sections is plotted nearest to the origin, oldest occurrences (bases) are plotted with "o," and youngest occurrences (tops) are plotted with "+." Coordinates are based on actual measured position (in feet or meters) in the sections being compared.

The next, and most difficult, step is positioning the LOC. Although this is called a "line," it may, in fact, consist of several interconnected line segments. The line segments must either have positive slope or be vertical, because time progresses in the positive direction. The biostratigrapher must determine the best position of the LOC through the array of points. Knowledge of the fossils and their environmental restrictions, sequence of fossil occurrences, and stratigraphic relations MUST play a role in the interpretation. Shaw (1964) originally used a least-squares fit of carefully evaluated points. Miller (1977*b*) advocated the individual consideration of each top and base to produce what the biostratigrapher feels is the most reasonable interpretation.

From a geological perspective, statistical methods are inappropriate for the exact positioning of the LOC for at least three reasons. (1) Some points on the graph more nearly represent time-equivalency than others. A prime example, noted by Shaw (1964) is the base of a correlatable volcanic ash bed. The LOC must pass through this point. The biostratigrapher will also de-emphasize points based on rare or sporadic occurrences in favor of

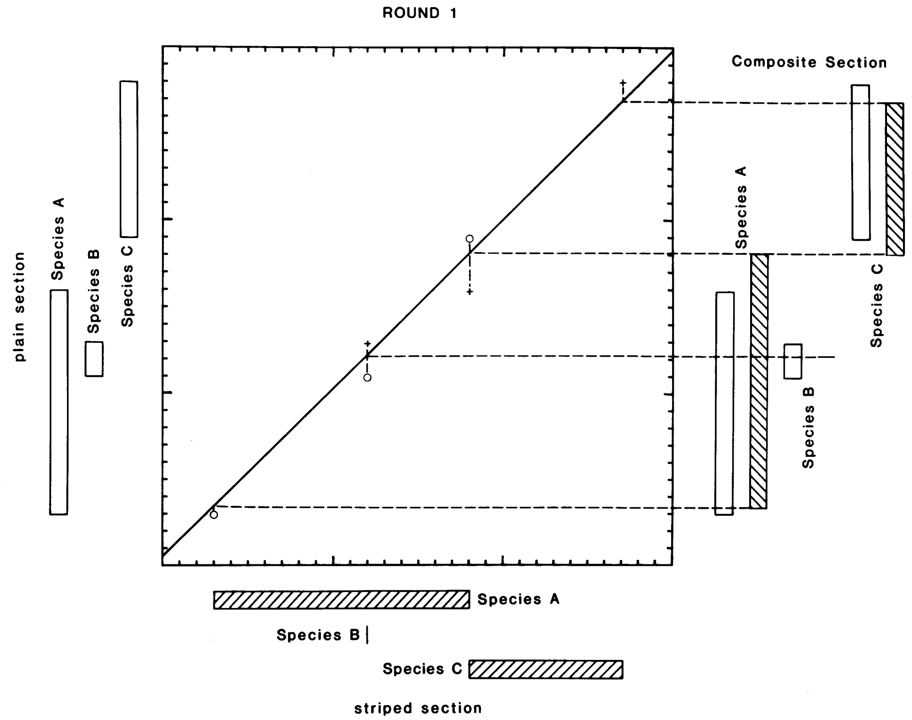

FIG. 1*c*.—Graph of species ranges in the striped and plain sections with a line of correlation (LOC) drawn. See text for comments on the positioning of the LOC. To the right of the graph, the ranges of Species A, B, and C in the plain section (reference section) are displayed again, and the ranges of the three species in the striped section are replotted in the framework on the plain section. The replotting of the striped ranges is accomplished by drawing a line perpendicular to the striped section to the LOC, and then projecting the value parallel to the striped section from the LOC to the plain section.

points based on well-represented taxa. (2) The biostratigraphic implications of first- and last-occurrence events are quite distinct (Edwards and Beaver 1978; Edwards 1982). As we will see later, first- and last-occurrence events are treated differently, and these differences aid in the positioning of the LOC. (3) Sedimentologic details (such as lithologic changes, burrowed surfaces) may offer insight or constrain the position of the LOC. One can not divorce the fossils from the rocks which entomb them.

Once the LOC has been positioned, it is a simple graphic or mathematical exercise to express position in one section in terms of position in the other section (fig. 1*c*). Thus, information from two sections may be combined in a single framework. This combined information becomes the composite section, and a third individual observed section may be plotted against the maximum ranges (lowest bases, highest tops) from the composite (fig. 1*d*). Although combined information is used to form the composite, the separate values must be saved for later rounds.

After all available sections have been plotted and all data projected into the framework of the composite (end of the first round), the procedure must be repeated several times until the LOC's stabilize. In second and later rounds, each section is compared, in turn, with the composite section, exclusive of those values in the composite that are derived from the section being compared (fig. 1*e* shows the beginning of the second round).

Lines of correlation are positioned through what the biostratigrapher considers to be the best points, subject to best judgment. The geometry of the graph allows the interpretations to be evaluated visually. First-occurrence events will fall on the LOC if the lower boundary of the range in the composite is equal to the lower boundary in the section

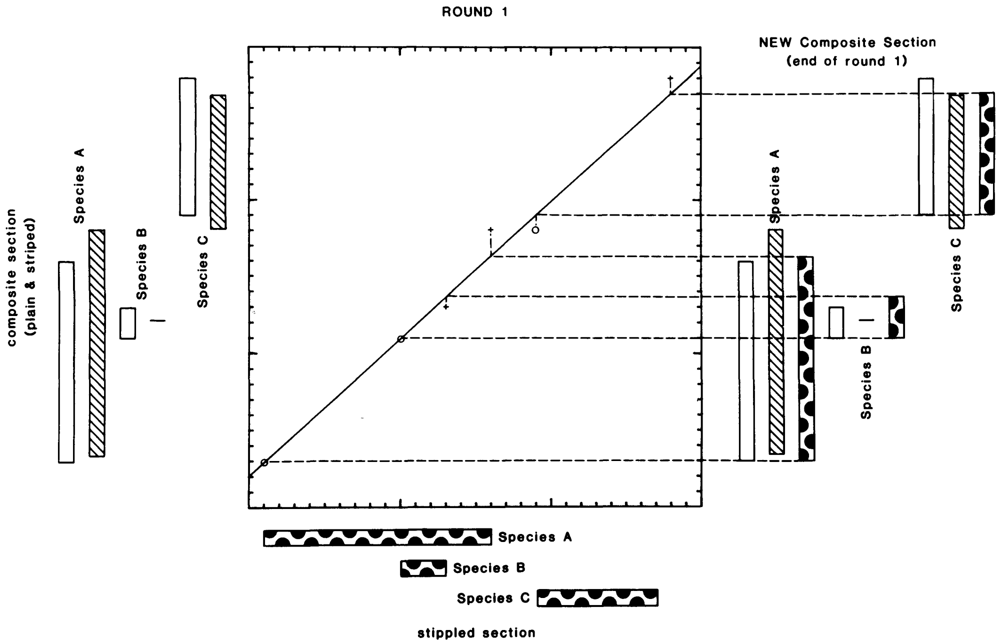

FIG. 1*d*.—Graph of species ranges from the stippled section (horizontal axis) against the combined ranges from the plain and striped sections (vertical axis). Within the graph, points are plotted according to the species-occurrence levels on the horizontal axis and the maximum range values on the vertical axis. These maximum range values are termed "composite section values." The LOC has been drawn, again using the judgment of the compositor. The geometry of the graph is such that points corresponding to ranges that are incomplete in the stippled section lie between the stippled-section axis and the LOC for first-occurrence events and between the composite-section axis and the LOC for last-occurrence events. Ranges that are longer in the stippled section than in the composite fall between the LOC and the composite axis for first-occurrence events and between the LOC and the stippled section axis for last-occurrence events. To the right of the graph, the ranges from the stippled section have been projected and added to the composite. The composite now contains data from all three sections, and this completes round one of the compositing process.

being considered and will fall between the LOC and the individual-section axis if the range in the composite extends lower than the range in the section. If the range in the composite is shorter than that in the section (that is, the first occurrence in the section is lower than that in the composite and a revision of the range based on that section is indicated), the first-occurrence point will fall between the LOC and the composite axis. Similarly, last-occurrence points will fall along the LOC or between it and the composite, unless the range is to be revised. If the last-occurrence point lies between the LOC and the section axis, it indicates a last occurrence at a higher level in the section than in the composite, suggesting that the range should be extended. The graphic chart does not force most interpretations; it merely shows the alternatives. The evaluation of the alternatives (positioning of the LOC) is the duty of the biostratigrapher.

Shaw's method, as originally presented, as updated by Miller (1977*b*), and as practiced here, relies heavily on the best judgment of the paleontologist/biostratigrapher. Anyone seeking a wholly objective means of manipulating data to reach reproducible conclusions may stop reading at this point.

I have experimented with Shaw's method for over 10 years. I find, empirically, that familiarity with the data significantly improves the "geologic reasonableness" of the results. When vast amounts of field and laboratory work, and years of experience, are reduced to a table of numbers representing levels of first and last occurrences, information is lost. Personally, I prefer a method that reintroduces some of this lost information to a completely objective (statistical) approach.

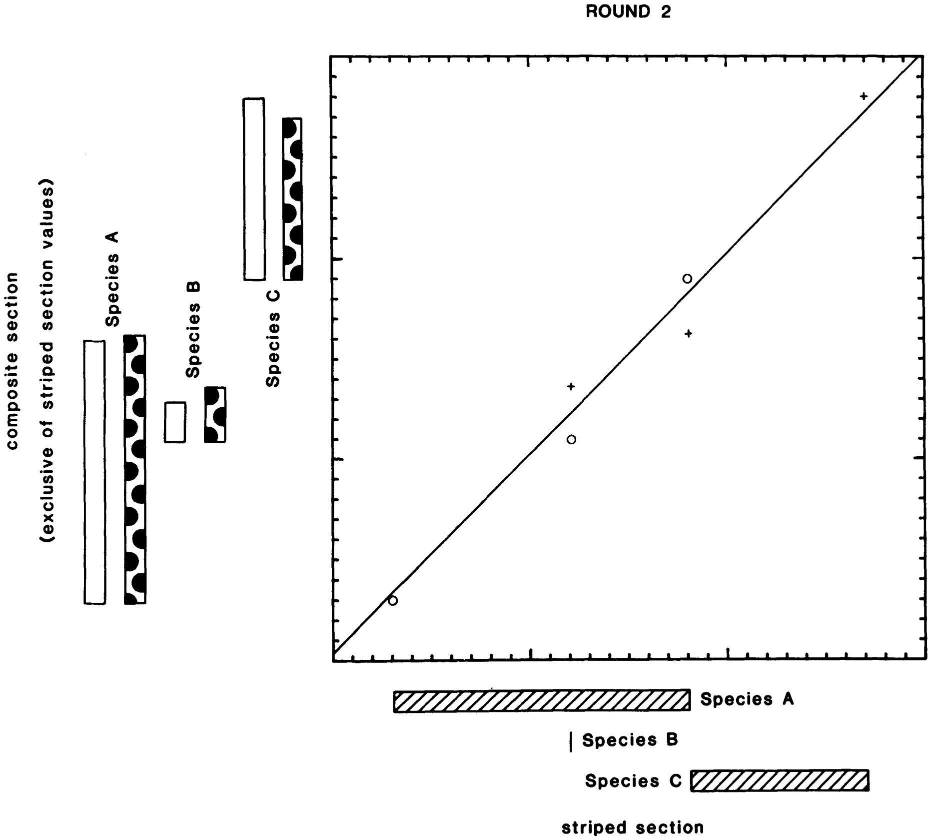

FIG. 1*e*.—The beginning of the second round of the compositing process. The striped section (horizontal axis) is plotted against the composite *exclusive of the values in the composite that are derived from the striped section* (vertical). Note that this graph differs slightly from that of figure 1*b* because the composite now includes information from the stippled section. To complete round two, the stippled section would be plotted against the resulting composite, exclusive of the stippled-section values. Third and later rounds would be completed in a similar manner until ranges stabilize.

ASSUMPTIONS

Any analytical procedure will have certain requirements for its successful application. I term these requirements the method's underlying assumptions. Most methods in biostratigraphy require three basic assumptions: (1) the law of superposition holds; (2) the taxa used are consistently identified; and (3) the evolutionary first appearance and the complete extinction of a taxon are each nonrecurring events (Edwards 1979). The application of Shaw's graphic correlation technique requires the following additional assumptions: (4) the lowest and highest stratigraphic occurrences of fossils, to the extent that they can be inferred to reflect evolution and extinction events, are excellent indicators of geologic time; (5) at the level of precision at which most stratigraphic work is done, the relative rates of rock accumulation in the several sections vary in a manner that can be approximated by a line or series of interconnected line segments—regardless of the actual rates involved; (6) occurrences of fossil taxa used in the method indicate the presence of living taxa at the time of sediment accumulation (reworking and contamination do not exist, or, if they do, the specimens involved can be recognized and disregarded); and (7) the best line of correlation is that which causes the minimum net disruption of the best established ranges, subject to the best judgment of the compositor ("economy of fit," Shaw 1964, p. 254–257).

The assumption of linear relative rates of rock accumulation is the one most open to dispute. Shaw (1964) and Miller (1977*b*) noted that in theory, the LOC should often be a continuous series of short straight lines reflecting minute changes in rates of accumulation. They noted, however, that at the scale and general level of accuracy of most collections, the difference between a straight line and a "dog-legged" line should have at most only a small effect. It is interesting to note that the theory behind graphic correlation does not require any sort of linear relationship among sections. The successful application of the method, however, does. The composite standard should be built up from complete, undisturbed, uncomplicated sections that can be correlated linearly (or with a minimum of linear segments if data are adequate). Only after ranges are well established in such a composite can one hope to recognize nonlinear patterns, which are rare, but not unknown, in the geologic record.

What if the underlying assumptions are not met? Failure to meet these requirements indicates a need for extreme caution on the part of the compositor but need not rule out the method entirely. For example, when the biostratigrapher suspects reworking, the suspect points may still be plotted. If these points fall conspicuously off the LOC, the suspicions are confirmed and the points are then discarded. Similarly, the method may be used to confirm a compositor's suspicions of inconsistent taxonomic identifications. These tests and confirmations will work only if sufficient reliable datapoints also exist.

THE "DATA"

The hypothetical dataset consists of 16 species, all of which have ranges that lie wholly within the spanned interval of time. Geographic considerations are simplified onto a single axis so that time and space are plotted on a two-dimensional graph (fig. 2). The 16 species include four that show a reasonable consistency in distribution in time and space (fig. 2*a*), four that are almost consistent (fig. 2*b*), four that are consistent but geographically restricted (fig. 2*c*), and four that are inconsistent species (fig. 2*d*). Sediment containing these 16 species has been sampled in time and space by six sections. One section (II) consists of three separated samples. Sections III and IV are based on the same line segment in time and space but have differing additional complications.

Differing rates of rock accumulation have been applied to the various sections. Sections I, II, III, and V each have a single linear rate constant. Sections IV and VI have rates that can be visualized as linear segments. Section I has an unconformity; time units 11–19 are missing, and time units 10 and 20 are present together. The rock accumulation rates are as follows (expressed in terms of rock units per time unit):

Section I, rate = 1, unconformity at 10–20 time units (t.u.)
Section II, rate = 1.2
Section III, rate = 0.5
Section IV, rates = 1.25, 1.5, 3, with bends at 8, 20 t.u.
Section V, rate = 0.75
Section VI, rates = 1.5, 2.5, with bend at 17 t.u.

In converting from time-unit values to rock-unit values, all fractions were rounded to integer values prior to sampling. This rounding simulates the effect of sampling from a finite volume of sediment (in contrast to a point-source sample).

The sections were sampled as follows (at rock-unit values):

Section I (11 samples) 0,2,4,6,8,10,12,14,16, 18,20
Section II (3 samples) 18,24,31
Section III (7 samples) 1,3,5,7,9,11,13
Section IV (19 samples) 3,5,7,9,10,13,16,19, 22,24,25,28,30,34,36,38,41,44,48
Section V (11 samples) 4,5,6,7,8,9,10,11,12, 13,14
Section VI (21 samples) 11,12,14,16,17,20,21, 22,25,26,28,30,31,33,35,36,38,41,43,45,47

In real geologic problems, a species is not always represented in every sample that intersects its range in time and space. The probability of actually finding at least one specimen of a species within its temporal, geographic, and ecologic range depends upon the abundance of the species in question and upon the size of the sample (see, for example, Shaw 1964; Hay 1972). In the present simulated example, I have set the probability of finding specimens of each species within its range at 0.8 and the probability of finding

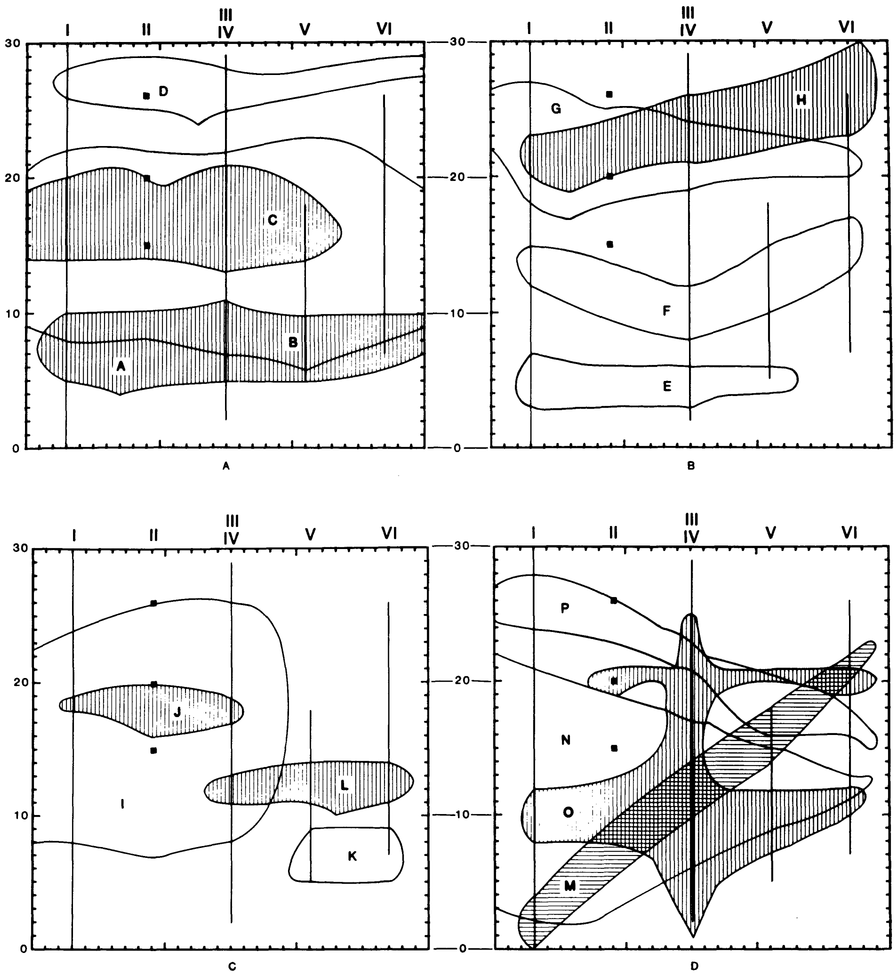

FIG. 2.—The "data." The hypothetical dataset consists of the ranges of 16 species in time (expressed in time units, vertical axis) and space (graphically portrayed on a single axis, horizontal). The six stratigraphic sections are illustrated and labelled. Section II consists of three separated samples. Sections III and IV are the same in time and space but have differing complications, as discussed in the text. The dataset is shown on four graphs to facilitate the visualization of the species ranges; all four represent the same span. All range boundaries should be read to the nearest integer time unit. (A) The ranges of species A, B, C, and D, reasonably consistent species in time and space. (B) The ranges of species E, F, G, and H, almost consistent species in time and space. (C) The ranges of species I, J, K, and L, almost consistent species in time, geographically restricted. (D) The ranges of species M, N, O, and P, inconsistent species in time and space.

specimens of the species outside its range at 0. The setting of this latter probability to 0 is required by assumption (6) above. A probability of 0.8 corresponds to a sample size of 4 for a species that makes up one-third of the population, a sample size of 15 for a species that makes up 1% of the population, or a sample size of 1600 for a species that makes up 0.1% of the population. A random-number function was used to eliminate 1 out of 5 species occurrences to produce table 1 from figure 2.

The hypothetical example simulates actual data with one very important difference—all true relationships in time and space are known. The goal of the following exercise is to apply Shaw's method of graphic correlation to the hypothetical dataset as though these relationships are not known. After the compositing process is complete, we may

TABLE 1

THE DATA

Section I		Section II		Section III		Section IV		Section V		Section VI	
Level	Species present	Level	Species present	Level	Species present	Level	Species present	Level	Species present	Level	Species present
0	M	18	B,C,I,N	1	O	3	O	4	A,E,K	11	A
2	M,N	24	B,C,G,H,J,O	3	E,N,O	5	E,O	5	A,B,E,K,O	12	A,B,K
4	M,N	31	I,P	5	A,B,F,I,M,N,O	7	. . .	6	A,B,O	14	A,B,K
6	A,E,N			7	C,I,L,M,N	9	A,B,N,O	7	A,K,N	16	B
8	B,I,N,O			9	B,C,O	10	B,F,O	8	B,F,L,N,O	17	B,L,N,O
10	A,B,G,H,I,N,O			11	B,G,I,O,P	13	A,B,F,I,M,N,O	9	B,F,L,O	20	B,L,N
12	B,G,I			13	D,H,I,O	16	B,F,I,L,M,N,O	10	F,L	21	B,F,L
14	G,P					19	C,I,M,N	11	L,N	22	B
16	D,G					22	B,C	12	B,C,M,P	25	B,F,P
18	D,P					24	B,C,I,J,O	13	C,M,P	26	B
20	. . .					25	B,C,I,J,O	14	M,P	28	B,P
						28	B,C,G,I,O			30	B
						30	G,I			31	B,O
						34	B,I,O,P			33	B,G,M,O
						36	G,P			35	B,G,M,O
						38	G,I,O			36	B,G,M
						41	H,I			38	M
						44	H,I			41	H
						48	D			43	H
										45	H
										47	H

then compare these results to the known "true" ranges to test the accuracy of the results.

THE COMPOSITING PROCESS—HYPOTHETICAL DATASET

Three rounds of Shaw's method were performed on the dataset given in table 1. The chosen order of compositing was Section III (reference section), followed by Sections V, II, I, VI, and IV. In this simulated case study, Section III was chosen as the reference section because of the presence of many species and the absence of discontinuities or rate changes, despite the relatively coarse sampling interval. In actual geologic settings, the choice of the order of compositing should be based on the completeness of the sections, the sample and coverage recovery, and the geologic judgment of the compositor. The reader may notice the unfair advantage taken here. By choosing Section III as the reference section, I am simulating geologic judgment. In practice the compositor should look for sedimentologic and biostratigraphic evidence and/or experiment with different choices of reference section to find one without apparent discontinuities or rate changes.

To start the first round, Section V has been plotted on the horizontal axis and Section III on the vertical axis (fig. 3). Because in either section the exact positions of first and last occurrences are also a function of sample interval, boxes have been drawn from each symbol to encompass all strata up to the next higher sample for last occurrences and all strata down to the next lower sample for first occurrences. Once the compositing process has proceeded past the first two sections, boxes are more difficult to draw, because the composite axis (here, vertical) has lost a real dimensional scale. However, the compositor should always consider the sample interval when positioning the LOC. Figure 3 shows the LOC selected for first-round correlation of Sections V and III. It is drawn, not by any regression formula, but by the judgment of the compositor. Here, I have attempted to intersect most of the sample-interval boxes. The compositor should expect that the first-round correlation will change in later rounds because more information from additional (later composited) sections will become available.

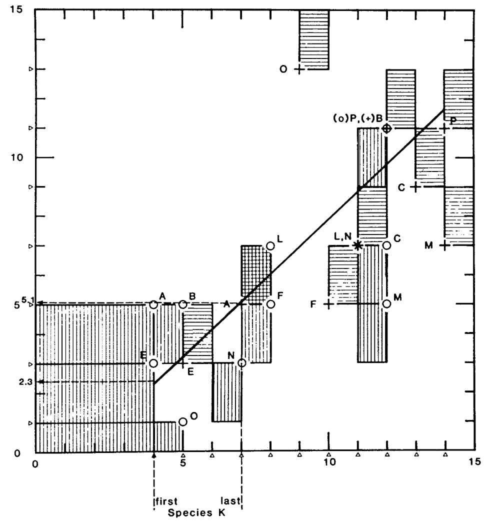

FIG. 3.—Round 1, beginning of the compositing process, first graph. Here, the reference section (Section III) is on the vertical axis, and the first section to be compared (Section V) is on the horizontal axis. Sample locations are noted by arrows along the appropriate axes. First-occurrence events (o) and last-occurrence events (+) are plotted for all species that occur in both sections. Boxes have been drawn to include all strata down to the next lower sample for first ocurrences (vertical hachures) and up to the next higher sample for last occurrences (horizontal hachures) to take the sample interval into account. The chosen LOC has been positioned to intersect many of the sample-interval boxes. Species K is present in Section V but not in Section III. It is added to the composite, as follows: A vertical line is drawn from its lowest occurrence (four rock units) in Section V up to the LOC and projected horizontally, the value 2.3 is read off the vertical axis; similarly, a vertical line is drawn from its highest occurrence (seven rock units) up to the LOC and over to yield a value of 5.1.

Values from Section V were brought to the LOC and translated to values with respect to the reference section. These new converted values were then added to the reference-section axis—which now becomes the composite axis—according to the rule that the lower value for first occurrences and the higher value for last occurrences be plotted against the next section (Section II). Species K, which was found in Section V but not in Section III, was added to the composite, also by translation (see fig. 3). Another LOC was drawn; values in Section II were converted to composite values; the lowest bases and the highest tops became the new composite section. The remaining sections were composited in the same fashion to complete the first round of compositing. Two further rounds of compositing were completed, each section being compared with the composite exclusive of those values in the composite derived from the section being plotted. By the end of three rounds, composite values had stabilized.

Surprisingly, each of the graphs could be interpreted as having a single straight-line LOC. Three of the sections, however, deserve further comment.

Section II includes samples from only three levels. After the first round, it simply did not provide enough information for a line to be drawn. I therefore drew a "least-damage LOC," one that did not extend any ranges at all, in order to continue the compositing process. Using an LOC that does not extend any range is equivalent to omitting Section II in these later rounds. This section is simply inadequate for compositing.

Section I showed a rather wide scatter of points. In the first round, a single straight-line LOC was drawn. In the second and third rounds, I could draw a single straight line; a dog-legged LOC having one bend and two line segments; or a line, an unconformity-caused offset, and a second, poorly constrained line. I chose the simplest hypothesis, the straight-line interpretation. In situations involving actual data, one would hope that knowledge of the fossils and of the physical stratigraphy would aid in choosing among various possible interpretations.

In Section IV, I experimented with a dog-legged LOC in the first round but decided against using a bend, again for reasons of simplicity. By the later rounds, the array of points had become more linear. Section IV merits further discussion because it illustrates several important points. Figure 4*a* shows the third-round plot of points. The linear interpretation shown in figure 4*b* is clearly reasonable, although the "true" correlation is shown in figure 4*c*. Because the example is hypothetical and has been constructed according to known rules, figures 4*d, e,* and *f* can be constructed. Figure 4*d* shows the scatter of points that would have been plotted assuming perfect correlation among all six sections, and figure 4*e* shows the scatter of points, assuming that the "true" total ranges of all taxa are known. Figure 4*f* again shows the true correlation, this time plotted against figure 4*e*. These plots reveal that the sampling interval, the probability of recovery, and the nonsynchroneity of range limits will produce a scatter of points even if section-to-composite correlations are perfect and even if all ranges are completely known. Any compositor should be aware of the above. Second, figure 4 shows that it is nearly impossible to detect a change in slope from $m = 0.8$ to $m = 0.667$ and that with less than perfect recovery and without closely spaced samples, it is difficult to detect a change in slope from $m = 0.667$ to $m = 0.333$. This latter change is equivalent to doubling the accumulation rate in one part of a section relative to another!

RESULTS AND DISCUSSION

Because a hypothetical example having a known linear relationship between the reference section and "truth" was used in the preceding example, the results produced by application of Shaw's method can be directly compared with "truth" by multiplying composite unit values by 2. Table 2 and figure 5 give the results of Shaw's method after the first round of compositing and after the third round, all composite values being doubled. They also give the "true" total ranges, which may not actually have been sampled, and the maximum ranges that actually were sampled.

Two features should quickly be noted from figure 5. First, the results produced by Shaw's method reproduce the "true" ranges with surprisingly high fidelity. They are not, however, perfect. Second, as would be expected, the results after three rounds are better than the results after a single round.

End of Round 1.—Despite the general close agreement at the end of the first round of compositing, some of the ranges are too long relative to the "true" ranges. There is good agreement in the middle of the graph, but the whole section appears "stretched" relative to truth. The span of ranges in figure 5 is 37 units rather than the actual 30 units. The stretching results from the magnification away from the center of slight errors in the placement of the several LOC's. (A small change in slope of any LOC will affect the values at the ends more than the values in the middle. Small errors in placement of LOC's

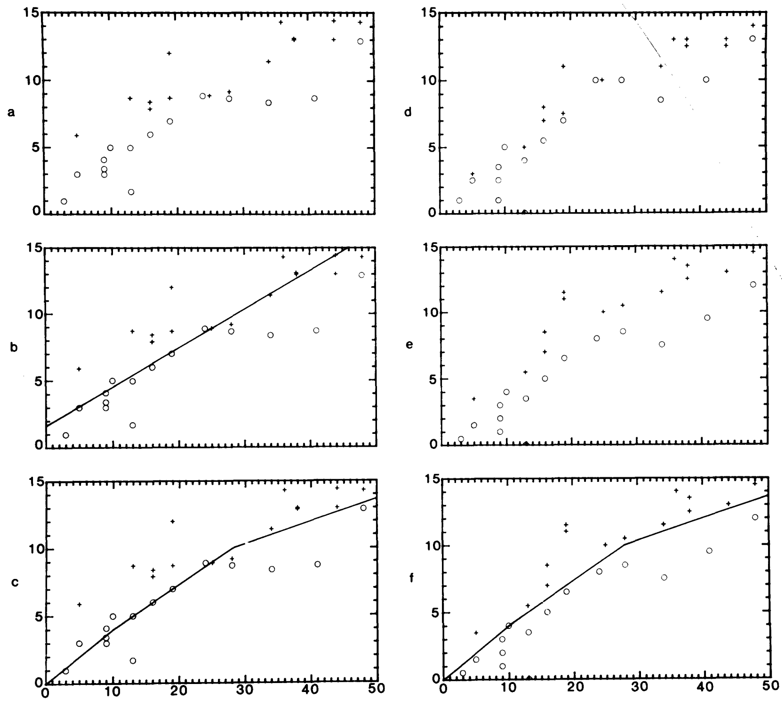

FIG. 4.—Section IV plots. (A) Points for observed position in Section IV (horizontal) plotted against the Round 3 composite exclusive of points derived from Section IV (vertical). (B) Same plot as (A), the chosen LOC drawn. (C) Same plot as (A), a dog-legged line representing the true correlation. This true correlation is known only because this example was constructed artificially. (D) Points for observed position in Section IV (horizontal) plotted against observed position (actually sampled and recovered) in the remaining sections, assuming perfect correlation among the remaining sections (vertical). Compare with (A). (E) Points for observed position in Section IV (horizontal) plotted against the given true total ranges (whether actually sampled and recovered or not) of the species (vertical). (F) Same plot as (E), the dog-legged line representing the true correlation. Note that in (B) and (C), points lie off the line because of one or more of the following factors: sampling interval in Section IV; incomplete recovery in Section IV; incomplete range in Section IV; sampling interval in Sections I, II, III, V, and VI; incomplete recovery in Sections I, II, III, V, and VI; incomplete ranges in Sections I, II, III, V, and VI; and errors in correlation among Sections I, II, III, V, and VI. In (D), this last factor has been eliminated. In (E) and (F), the last four factors have been eliminated.

for different sections in opposite directions will stretch the composite.)

Some of the ranges, even near the middle of the section, are too long. Most notably, these are the top of Species A, the bases of Species G and H, and the top of Species L. These errors result in artificial range overlaps for several pairs of species. The errors in the ranges of Species A, G, and H are caused by failure to detect the unconformity in Section I. The error in the range of Species L is a result of its position in Section V, the first section to be composited. Simply because the compositing process uses lowest bases, highest tops throughout, slight errors in the positioning of the LOC are more likely to lengthen ranges artificially than to shorten them. (Any individual error is as likely to shorten the range in a single section as to lengthen it. However, in order for the range to be shortened in the composite it must be shortened in all individual sections. In order for a range to be lengthened in the composite, it need only be lengthened in a single individual section.)

End of Round 3.—By this time, the section is no longer stretched relative to truth. However, a systematic error remains. As compared with actual sampled values, many more of the events (both bases and tops) are too high in the composite rather than too low. This may result from an unknown systematic

TABLE 2

RESULTS

Event	True value	Results of Rd1 (× 2)	Results of Rd3 (× 2)	Actually sampled w/P = 0.8 (× 2)
16 top H	30	30	29	25
8 top D	29	32	31	28
32 top P	28	30	29	26
14 top G	27	26	26	26
18 top I	26	30	29	26
30 top O	25	26	26	25
7 base D	24	26	26	26
26 top M	23	22	24	22
4 top B	23	24	23	22
28 top N	22	18	17	15
6 top C	21	22	20	20
20 top J	20	22	18	20
15 base H	19	15	17	20
12 top F	17	16	17	16
13 base G	17	15	17	20
19 base J	16	18	17	17
31 base P	15	16	17	17
24 top L	14	18	16	14
5 base C	13	14	14	14
2 top A	11	15	17	10
23 base L	10	12	12	11
22 top K	9	10	11	9
11 base F	8	9	9	8
10 top E	7	7	12	6
17 base I	7	10	10	8
3 base B	6	7	8	7
21 base K	5	5	7	5
1 base A	4	5	7	5
9 base E	3	5	6	4
27 base N	2	− 1	6	2
29 base O	1	2	2	2
25 base M	0	− 5	3	0

bias on the part of the compositor or may be an effect of the design of the dataset (both sections with dog-legs had increasing rock-accumulation rates in the upsection direction). Artificial range extensions still exist but have decreased relative to the first round. The major inaccuracies (such as the overlap of Species A and C and Species E and F) again are caused by failure to detect the unconformity in Section I. Species A and C and Species E and F are never found to overlap in any section. The maximum ranges of the two species in question are set by two different sections, and because of an error in one of these sections, the ranges show false overlap.

Conclusions.—Figures 4 and 5 and the above discussion lead to one of the major discoveries of this paper—why Shaw's method works. It works because (1) actually sampled ranges are always shorter than true ranges, given finite sampling and assumption (6); and (2) if section-to-composite correlation is not perfect, Shaw's method will extend some ranges artificially. Effects (1) and (2) operating together make Shaw's method produce ranges that closely approximate the "true" ranges. The method can be predictive. Species F and G never overlap in any section and, in fact, never overlap in time and space. However, in figures 2 and 5, they do exactly touch in time, and the results of three rounds of Shaw's method predict this!

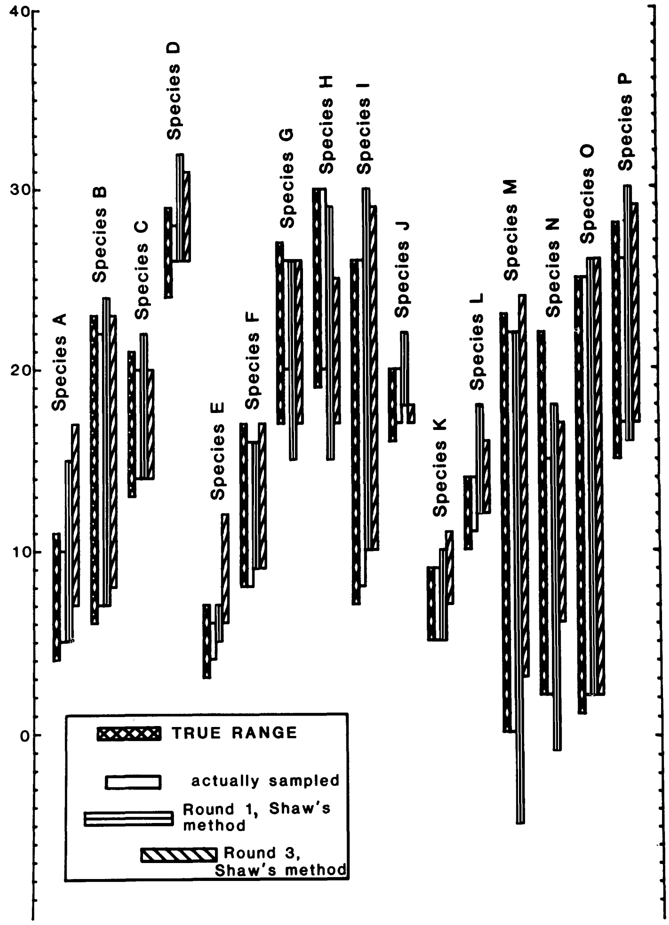

FIG. 5.—Range chart comparisons. The vertical axis represents the time units as shown in figure 2 that are equal to composite units multiplied by 2. Patterns indicate the true total ranges, the ranges as they were actually sampled, the ranges produced after the first round of Shaw's method, and the ranges produced after three rounds of Shaw's method.

With finite sampling, the two effects noted above will always occur. Actually-sampled ranges *will* be shorter than true ranges. One should be able to reduce this effect to an acceptable level by increased sampling. Imperfect correlation *will* lengthen some ranges artificially because a range that is extended artificially by an error in the compositing of a *single* section will be extended in the final results. Anything that can be done to reduce correlation errors will reduce this effect.

Shaw's method is only as good as the available data. As Shaw (1964, p. 154) so succinctly states: "The technique described here does not in any way relieve the paleontologist from thinking about what he is doing. The tool presented by the graph is extremely powerful, but it is also insensate. It cannot overcome inconclusive or inadequate data, nor can it compensate for slipshod thinking in its applications."

CONCLUDING REMARKS

The above example is hypothetical. Many of its features will be unrealistic compared with any given real-world situation. Two of the most notable unrealistic features are that the ranges of all species begin and end wholly within the chosen interval (which makes the bookkeeping simpler but should not greatly affect the logic or results) and that the probabilities of finding all species within their ranges are equal and set at 0.8. The 0.8 may be unrealistically low for microfossil groups such as calcareous nannofossils (large sample size) or unrealistically high as perhaps for vertebrate remains (typically small sample size).

During the compositing process, I may not have always chosen the best LOC, but I have always tried to pick a reasonable line. The best-choice line should always depend on judgment of the fossils and knowledge of the stratigraphy. There was no way to simulate the necessary judgment and knowledge. The reader should be aware of the outcome of using reasonable but incorrect LOC's.

By using hypothetical data, the effects of errors in correlation can be separated from the effects of incomplete recovery (compare figs. 4*a, d,* and *e*). Figure 4, however, clearly shows that incomplete sampling or recovery and nonsynchroneity of range limits will produce a scatter of points even if all correlations are perfect and all ranges are known. The biostratigrapher should be aware of this scatter and should collect adequate material to overcome this part of the problem. The biostratigrapher should also recognize and accept that the geologic record imposes limitations that cannot be overcome completely. We make the best correlations we can with the best information available.

I hope that the above example is useful to any biostratigrapher who is considering applying Shaw's method. All the insights and comments made here should be evaluated in light of the actual problem at hand.

ACKNOWLEDGMENTS.—I thank the many people who have provided insights into graphic correlation, especially M. A. Murphy, P. M. Sadler (both of the University of California, Riverside), J. E. Hazel (Amoco, Tulsa), R. A. Christopher (Arco, Dallas) and N. O. Frederiksen (U.S. Geological Survey, Reston). J. E. Hazel, R. Z. Poore, T. M. Cronin (both of the USGS, Reston), A. B. Shaw, and F. X. Miller (both of Amoco, Tulsa) have provided suggestions and comments on earlier versions of this work. C. Lending and R. Sussman (USGS, Reston) worked with parts of the computer programming. I particularly thank three anonymous reviewers.

REFERENCES CITED

EDWARDS, L. E., 1979, Range charts and no-space graphs: Computers and Geosciences, v. 4, p. 247–255.

——— 1982, Quantitative biostratigraphy: the methods should suit the data, *in* CUBITT, J., and REYMENT, R., eds., Quantitative Stratigraphic Correlation: Chichester, Wiley, p. 45–60.

———, and BEAVER, R. J., 1978, The use of a paired comparison model in ordering stratigraphic events: Jour. Math. Geology, v. 10, p. 261–272.

HAY, W. W., 1972, Probabilistic stratigraphy: Eclogae geol. Helv., v. 65, p. 255–266.

HAZEL, J. E.; MUMMA, M. D.; and HUFF, W. J., 1980, Ostracode biostratigraphy of the Lower Oligocene (Vicksburgian) of Mississippi and Alabama: Gulf Coast Assoc. of Geol. Soc. Trans., v. 30, p. 361–401.

MILLER, F. X., 1977*a*, Biostratigraphic correlation of the Mesaverde Group in southwestern Wyoming and northwestern Colorado: Rocky Mtn. Assoc. of Geol. 1977 Symposium, p. 117–137.

——— 1977*b*, The graphic correlation method in biostratigraphy, *in* KAUFFMAN, E. G., and HAZEL, J. E., eds., Concepts and Methods of Biostratigraphy: Stroudsburg, Pa., Dowden, Hutchinson and Ross, p. 165–186.

MURPHY, M. A., and BERRY, W. B. N., 1983, Early Devonian conodont-graptolite collation and correlations with brachiopod and coral zones, central Nevada: Am. Assoc. of Petrol. Geol. Bull., v. 67, p. 371–379.

———, and EDWARDS, L. E., 1977, The Silurian-Devonian boundary in central Nevada, *in* MURPHY, M. A.; BERRY, W. B. N.; and SANDBERG, C. A., eds., Western North America: Devonian: Univ. Calif., Riverside, Campus Mus. Contrib. 4, p. 183–189.

PALMER, A. R., 1955, The faunas of the Riley Formation in central Texas: Jour. Paleontol., v. 28, p. 709–786.

SHAW, A. B., 1964, Time in Stratigraphy: New York, McGraw-Hill, 365 p.

SWEET, W. C., 1979*a*, Late Ordovician conodonts and biostratigraphy of the western Midcontinent province: Brigham Young Univ. Geol. Studies, v. 26, p. 45–85.

——— 1979*b*, Graphic correlation of Permo-Triassic rocks in Kashmir, Pakistan and Iran: Geologica et Palaeontologica, v. 13, p. 239–248.