Chapter 2

Best-Fit Intervals and Consensus Sequences

COMPARISON OF THE RESOLVING POWER OF TRADITIONAL BIOSTRATIGRAPHY AND COMPUTER-ASSISTED CORRELATION

PETER M. SADLER AND ROGER A. COOPER

PETER M. SADLER • Department of Earth Sciences, University of California, Riverside, Riverside, CA 92521 ROGER A COOPER • Institute of Geological and Nuclear Sciences Limited, Lower Hutt, New Zealand

Approaches in High-Resolution Stratigraphic Paleontology, edited by Peter J. Harries,

1. THE BIOSTRATIGRAPHIC SEQUENCING PROBLEM

Biostratigraphers subdivide geologic time using the sequence of first and last appearances of only a small fraction of the available fossil taxa. Thus, there appears to be considerable scope for improving biostratigraphic resolution simply by incorporating more first and last appearance events. The problem, of course, is that the proper sequence becomes more difficult to determine as the number of events rises. Figure 1 illustrates how a mild example of the problem can be visualized in the form of a tangled fence diagram. It is constructed from the observed ranges of trilobites collected by Palmer (1954) from the Cambrian Riley Formation of Texas. The seven fence posts represent seven different stratigraphic sections. Each wire line connects the horizons of the lowest or highest find of a taxon that has been recovered from adjacent sections. The wire lines cross wherever a pair of events have been preserved in contradictory order in adjacent sections. These Riley data are a classic example of a relatively internally consistent set of range charts. They have become a benchmark for testing and illustrating new quantitative methods: graphic correlation (Shaw, 1964); unitary associations (Guex, 1991); and constrained optimization (Kemple *et al.*, 1995). Nevertheless, 253 of all the possible pairs of first and last appearance events in this data set have been observed in contradictory order. For another 543 pairs, at least one section is equivocal, in the sense that the events are preserved at the same stratigraphic level -- a condition that is consistent with either order. Traditionally, the Riley ranges are used to separate only six or seven zones.

In effect, traditional biostratigraphy restores the fence to an untangled appearance by selective removal of lines until those that remain do not cross one-another. But it is likely that many of the lines meet many of the posts at the wrong levels; the highest and lowest finds in these sections are not at the horizons which correspond in age with the true first and last appearances of the taxon. This chapter presents the case for *adjusting all* the fence lines rather than *eliminating most* of them. The resulting fence will have much higher resolution because it retains many more lines. The “best-fit intervals”

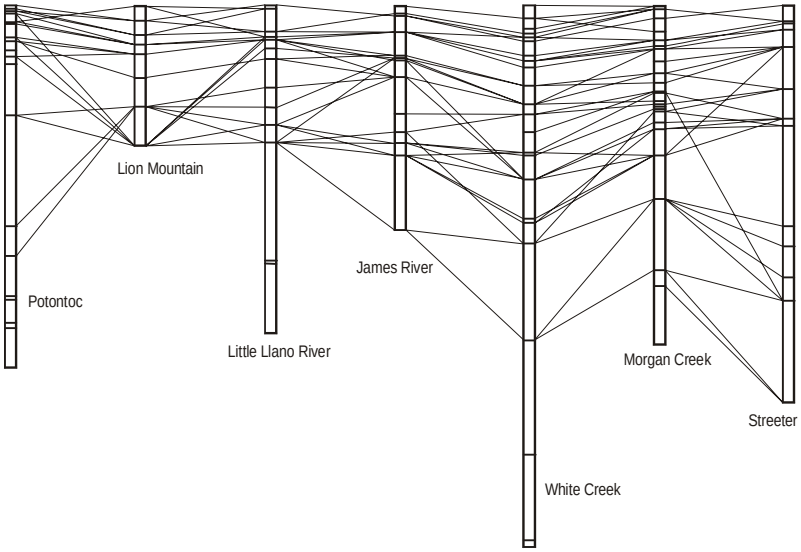

Figure 1. Tangled-fence diagram of the observed contradictions between 7 sections that preserve 62 taxa of the Cambrian Riley Formation of Texas (data from Palmer, 1954; Shaw, 1964). Note that fence lines are drawn only for taxa found in both of the adjacent sections; in order to illustrate the full extent of the tangle we would need to plot every section next to every other (for 7 sections that requires 3 more fences with the posts in a different order). Fossiliferous horizons are marked by horizontal bars in the columnar sections.

mentioned in the title will be shown to express the relative resolving power of the adjusted lines – the range of positions they may occupy in different but equally good solutions to the problem of untangling the fence. The "consensus sequence" will be the largest set of lines that are internally robust in the sense that none of their best-fit intervals overlap.

A few straightforward rules govern the adjustment of the fence lines. They are simpler than the rules for Shaw's (1964) graphic correlation. Nevertheless, the approach requires computer assistance to select the minimum sufficient adjustments from an overwhelming number of options. The results are reproducible and uncertainties in the placement of events can be fully quantified in terms of the fit between the model sequences and the observed local ranges.

Tangled-fence diagrams expose very good reasons why traditional methods use only a fraction of the available first and last appearance events. The task of untangling the fence rapidly becomes totally bewildering as the number of fence posts and the number of wire lines increases. First, in each section there is a possibility that random factors in preservation and collection will cause some pairs of events to be preserved in the wrong order. Thus, the number of pairs of events that are found in contradictory order must be expected to increase with the number of stratigraphic sections

in which both events have been found. This expectation will be confirmed below, using empirical data. Second, the number of possible sequences that can be built from the observed disorder of events grows with the number of taxa. This result will be derived from a consideration of permutations.

The increase in the number of possible solutions with the number of taxa is extremely pernicious. The growth rate is worse than exponential. Familiar problems that encounter the same difficulty include the search for the most parsimonious cladogram and the inversion of geophysical data. Fortunately, this difficulty arises in a large class of very practical commercial problems (termed "non-polynomial complete," Dell *et al.*, 1992), like the routing of delivery vehicles. As a result, the discipline of operations research has focused considerable effort on developing efficient computer algorithms that can solve these tough problems. It is not our purpose to criticize traditional biostratigraphy. Knowing that non-polynomial complete problems have been dubbed "the Mt. Everest of computer science" (Anderson, 2001), it is evident that the traditional culling of events was an intelligent and entirely justified strategy for solving the stratigraphic correlation problem, prior to the availability of computer methods.

Before examining the adaptation of these computer methods to biostratigraphy, we must explore several aspects of the biostratigraphic sequencing problem: 1) the contradictions within the field data that lie at the core of the problem; 2) the rules for resolving them; and the size of the set of feasible sequences that could resolve the contradictions. At the end of the chapter, we discuss selected aspects of four case histories to illustrate the types of results that can be expected: 1) a classic data set for which the results have been compared with a manual solution; 2) a controversial data set for which the differences between two published interpretations can be quantified; 3) a time-scale and biodiversity project that avoids the limitations which traditional zones and stages impose upon resolution; and 4) a set of exploratory wells in which borehole caving compromises the first-appearance data and for which the results have been compared with seismic correlation.

1.1 Contradictions

Inconsistencies in the order of first and last appearances, as recovered from different stratigraphic sections, can be expressed as a contradiction ratio: that proportion of the evidently contradictable pairs of events (A and B) for which the two possible sequences (A-below-B and B-below-A) have each been found in at least one section. Some pairs of events cannot be preserved in contradictory order because the ranges of the corresponding

taxa do not overlap. The "evidently contradictable" pairs are those whose ranges have been observed to overlap. For them, every additional section that is examined means another chance to discover a contradiction.

The causes of contradiction are well known. The distribution of living organisms is patchy. Few individual organisms become fossils and few of these fossils are recovered in identifiable condition. As a result, the stratigraphic ranges of fossil taxa tend to be too short; they underestimate the range of the living organisms. Lowest finds tend to be too late and highest finds too early at any one stratigraphic section when compared with the regional or global range ends that would be the basis of a time scale. The discrepancies are inevitable; they vary from place to place and from taxon to taxon. Thus, contradictions must be expected between the sequences of range-end events preserved in different sections.

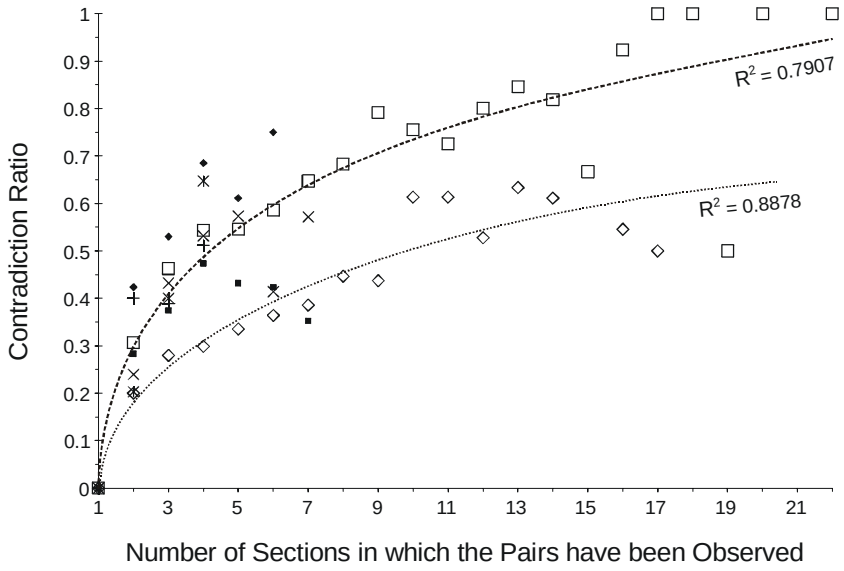

Figure 2. Increase in the pairwise contradiction ratio with increase in the number of sections in which the pairs of taxa are observed. Large open squares and coarsely dashed logarithmic regression: Ordovician through early Devonian graptolites (Sadler and Cooper, unpubl.). Large open diamonds and finely dashed logarithmic regression: Late Miocene to Recent foraminifera and radiolaria from North Atlantic ODP sites. Other data, for which the number of sections is too limited to justify regressions: trilobites of the Cambrian Riley Formation (Palmer, 1954); graptolites of the Mohawk Valley (Goldman *et al.*, 1994); Maastrichtian ammonites of Seymour Island, Antarctica (Macellari, 1984, 1986) and the Bay of Biscay (Ward and Kennedy, 1993).

Empirical evidence (Fig. 2) confirms that contradiction ratios generally increase with the number of sections from which a pair of events has been recovered. The rate of discovery of new contradictions slows as more

sections are added because the odds of contradiction operate on the dwindling population of pairs that still appeared to be consistent after the previous section was examined. But the rate slows faster than would be predicted if the odds of contradiction were constant for every pair of events. Better models for the shape of the regressions in Figure 2 allow the odds of contradiction to vary from pair to pair. Then, the pairs of events with long odds of contradiction are likely to appear consistent until very large numbers of sections have been examined.

It is reasonable to suppose that the properties of pairs of events render some pairs harder than others to preserve in contradictory order. The section of this chapter that deals with traditional zonation explains some of the variation in terms of range lengths and the time separation between events. Other factors probably relate to ecology and preservation. It will require much more data, however, to tease such effects apart on diagrams such as Figure 2.

Of course, failures of preservation and collection alone do not account for all the contradictions. Some reflect genuine differences in the timing of local appearances and disappearances, such as result from slow migrations. Nevertheless, a time scale that is based upon first and last appearances needs the proper order of the regional extreme events - the oldest of the local first appearances and the youngest of the local last appearances. Shaw (1964) established this principle in graphic correlation. The size of the region will be determined by the geographic scope of the correlation problem.

As in graphic correlation (Shaw, 1964), resolving the contradictions requires adjustment of the locally observed range ends until all localities fit one regional sequence of events. An acceptably parsimonious procedure must minimize the sum of all these adjustments. We will demonstrate the natural logic of such a procedure by first considering a trivial correlation problem and then enlarging it to a realistic size. The rules for untangling fence diagrams emerge in the process.

1.2 Rules for Resolving Contradictions

Figure 3 extracts from Shaw's (1964) Riley data an elementary contradiction that involves only two trilobite taxa and two local stratigraphic sections (White Creek and Potontoc). Assuming that the observed ranges are either correct or underestimate the regional range, we can resolve this contradiction without difficulty. (This assumption would be less straightforward for microfossils that are susceptible to benthic mixing; the thickness of the mixing layer then leads to a natural limit for meaningful precision in the measurement and adjustment of ranges.)

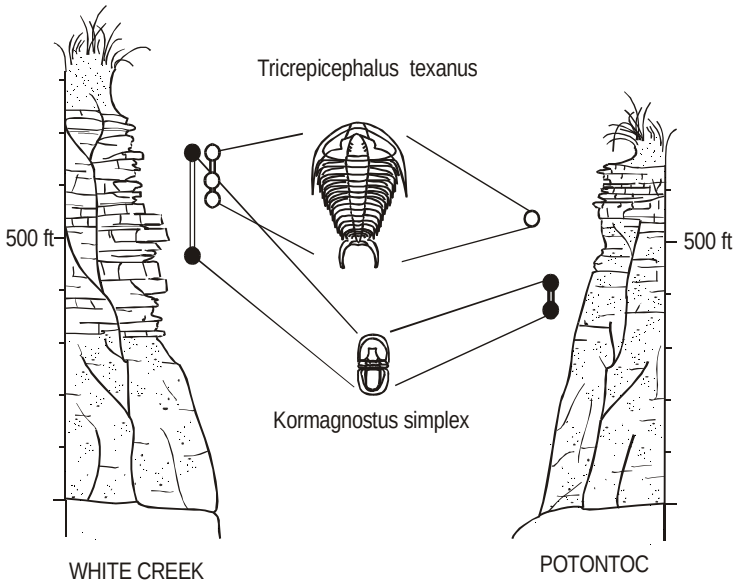

Figure 3. Contradictory ranges of two trilobite taxa from two measured stratigraphic sections in the Upper Cambrian Riley Formation of Texas (Palmer, 1954; Shaw, 1964). The Riley data are presented in feet rather than meters because they are a classic comparative benchmark that has always been analyzed in the original thickness scale (Shaw, 1964; Guex, 1991; Kemple *et al.*, 1995).

The White Creek section proves that the two ranges did overlap in time; the two trilobite taxa must have coexisted. The Potontoc section cannot be read literally as the correct sequence of events, because it fails to record the proven coexistence. So, for a sequence of events that can form the basis of a time scale we choose the sequence of first and last appearance events as preserved at White Creek. Any other sequence for these four events would be less parsimonious, in the sense that it implies either a greater failure of preservation and collection, or a more complex pattern of migration. Thus, very straightforward logic has resolved this elementary contradiction. It is only the huge number of pairs of events in a real data set that overwhelms such mental reasoning. In order to develop a rigorous method that will untangle large data sets, let us pursue this trivial example farther by considering all the possible histories for two taxa.

Two taxa generate four events: two first appearances and two last appearances. The number of biologically feasible sequences is constrained to be less than the total number of permutations of four events because each first appearance must precede the last appearance of the same taxon. In general, two taxa present six possible sequences of events (Fig. 4). The number six arises as follows: there are two possible choices for the oldest

first appearance and then three options each for the subsequent history. The second taxon might not appear until the first has already disappeared (sequences 1 and 6 in Figure 4); the second taxon might appear within the time span of the first and then outlive it (sequences 2 and 5), or the range of the second taxon might be relatively short and fall entirely with the range of the first (sequences 3 and 4).

Before trying to choose which of these six options best fits the field observations, notice that they exclude the possibility that two events are exactly simultaneous. That possibility raises questions about meaningful and practical limits to the measurement of time that do not yield simple

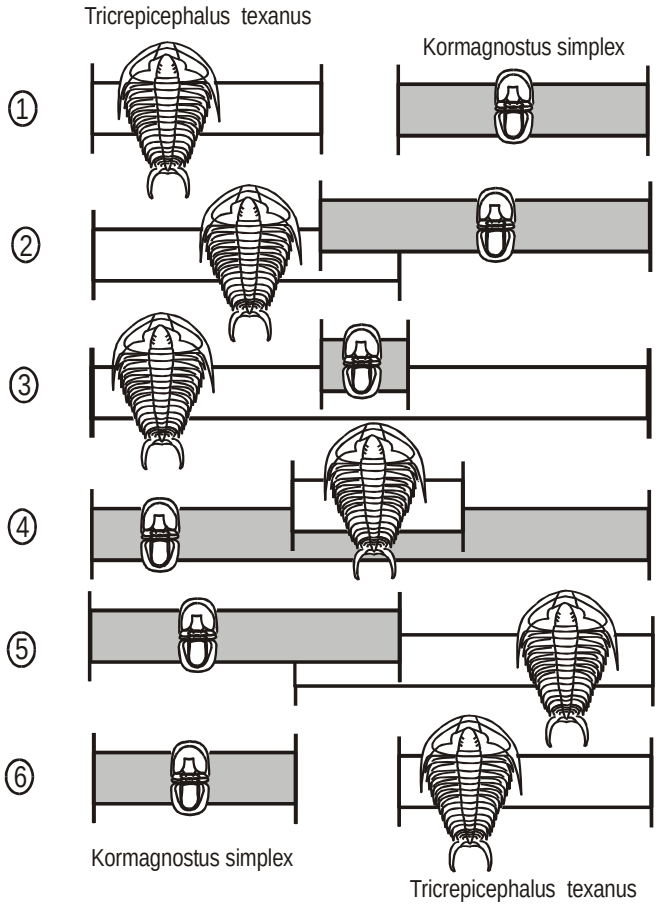

Figure 4. Six feasible sequences for the first and last appearance events of two taxa. Time progresses from left to right. For each case, the gray and the white rectangles represent the two taxon ranges, first appearance on the left, last appearance on the right.

answers. Accordingly, the possibility of simultaneous events is deliberately reserved for the task of scaling the time intervals between events in the best sequence. Separating questions of sequence and spacing brings several advantages (Kemple *et al.*, 1995). It means that, although the choice of best sequence does not allow ties, the resolvable interval between events may later be assigned a practical value of zero. Zero spacing would be permitted where no section can resolve a stratigraphic separation between adjacent events, either because they occur at the same level or because the separation between them is smaller than the thickness of the benthic mixing layer.

Of course, if pairs of events can have zero spacing, they can generate sequences that are practically indistinguishable. The sequences would differ only in the order of pairs of adjacent events. They would be one case of different sequences that fit the field data equally well. The concept of equally well-fit sequences is central to the development of best-fit intervals and consensus sequences.

Another critical concept is constraint. The stratigraphic observation that two taxa must coexist provides a constraint on the number of *geologically* feasible sequences. In our elementary example, this constraint eliminates two of the six sequences (1 and 6) because the two taxon ranges overlap at White Creek. In order to choose the best of the remaining four sequences (2, 3, 4, and 5), determine the minimum net adjustment that would be necessary to make both stratigraphic sections fit each sequence. The best, or most parsimonious, sequence would then be the one that requires the least adjustment of the observed ranges. Thus, the problem naturally takes the form of a constrained optimization (Kemple *et al.*, 1989) with two fundamental rules. First, eliminate sequences that fail to honor observed coexistences; then search among the remainder for the one that best fits the field observations in the sense that it minimizes the implied failings of the preservation and collection process. The magnitude of the misfit is the sum of all the necessary *ad hoc* range adjustments (Table 1). It can be measured in a variety of ways (Kemple *et al.*, 1995). The raw total length of all the range adjustments, expressed as stratigraphic thickness, will suffice to apply the rules of optimization to Figure 3. Table 1 summarizes the results.

Table 1. Misfit between field observations and the six model sequences in Figure 4. ("INFINITE!" misfit indicates that the sequence fails to meet the coexistence constraints)

Model Sequence:	1	2	3	4	5	6
Misfit with observations:	INFINITE!	390 ft	270 ft	120 ft	120 ft	INFINITE!

Remember that the adjustments may only lengthen the observed ranges, not shorten them. Also, where two events are preserved at the same stratigraphic horizon either of them may be interpreted to have occurred

before the other, because we can assume that sediment accumulated too slowly (or with too much mixing) to resolve the true separation. Operationally, this leads to two more rules: a single configuration of observed ranges may be compatible with more than one sequence of events; and range extensions do not alter the observed sequence of events unless they extend an observed range to an horizon where another event has been observed.

Sequences 4 and 5 in Figure 4 both resemble the field observations, to the extent that taxon *Kormagnostus simplex* appears before *Tricrepicephalus texanus*. These two sequences share the same minimal misfit (120 ft) because they require only that the top of the range of *K simplex* at Potontoc be extended up to the horizon of the last appearance of *T. texanus*. Fitting the observed ranges to sequences 2 or 3 requires substantially more adjustment because the range of *T. texanus* must be extended downward in both sections. The field observations do not fit sequence 2 unless all three of the extensions are adopted from Figure 5, giving a total misfit of 390 ft. Sequence 2 has the worst fit.

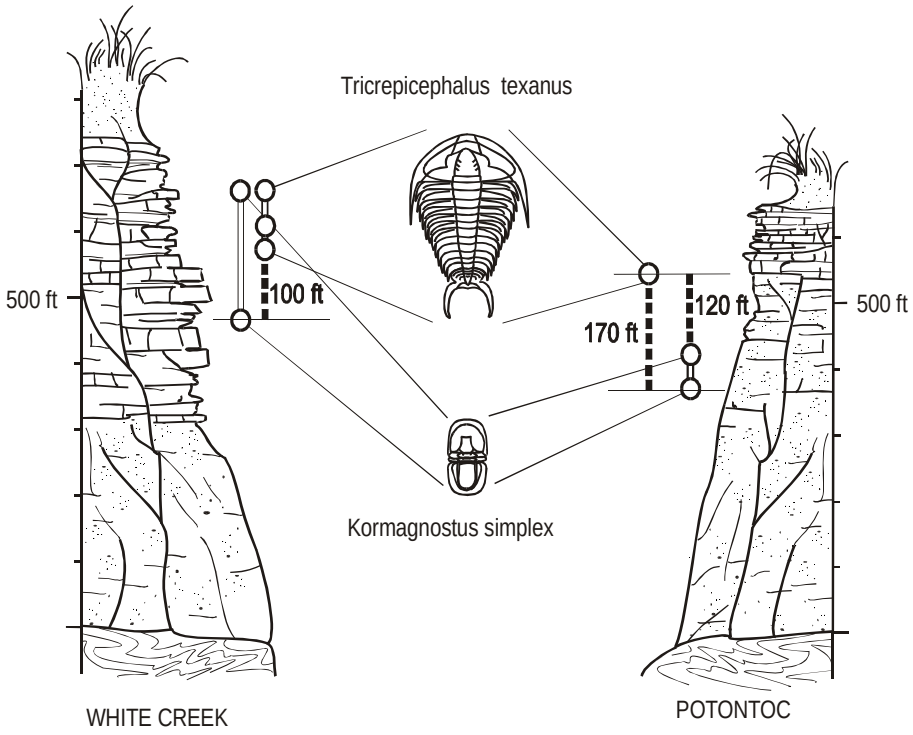

Figure 5. The set of minimum range extensions which, in different combinations, can adapt the observed ranges in each section to fit any feasible sequence of events (Fig. 4).

Together, sequences 4 and 5 are the set of most parsimonious solutions for this elementary problem in the sense that they imply the minimum failing of the fossil record *in these two sections*. After more sections and taxa are added, this element of the solution might need to change in order to minimize extensions to other observed ranges. Therefore, the optimization of larger data sets cannot readily be constructed piecemeal from a set of two-by-two elements, with each element remaining as simple as our example. Rather, the necessary adjustments must be considered for whole sequences. The addition of more taxa improves the potential resolution. In general, it is to be hoped that the inclusion of more sections improves the quality of the solution. Additional sections improve the chances of finding the earliest and latest occurrences and of observing all pairs of events in the correct order.

1.3 The Number of Possible Sequences

The rules for resolving contradictions do not alter when the number of taxa and sections increases. An increase in the number of sections simply lengthens the computation time required to find and sum all the appropriate range extensions for any one sequence. But increasing the number of taxa has an alarming effect on the number of feasible sequences that must be evaluated. Consider adding just one more taxon to the previous example.

The third taxon (Fig. 6) adds a fifth and sixth event to the model sequences. The first two taxa have filled four of the six positions (large open circles in Fig. 6) and they can do so in six different ways. For the two new events, imagine a pair of possible place-holders (small filled circles in Fig. 6) at the beginning of the line, at the end of the line, and between each pair of the four filled positions. The place-holders are paired to admit the possibility that the new first appearance and last appearance events will occupy adjacent positions.

Now consider *all* the ways to insert the two new events, proceeding systematically, as follows. Place the first appearance of the new taxon in the earliest placeholder and count the number of possible younger positions for the last appearance (only one in each pair of small filled circles except that shared with the first appearance). Move the first appearance to the next oldest pair and repeat the process. Total the options for the two new events (5+4+3+2+1 = 15) and multiply by the permutations of the previous four events (15 x 6 = 90) to discover that there are ninety sequences of events for three taxa. Some of these may fail to satisfy coexistence criteria; the remainder will all be compatible with the observed ranges plus various combinations of range extensions.

Figure 7 adds a fourth taxon and repeats the previous calculations, but with more placeholders for the new events. The previous three taxa filled

six positions in ninety possible ways. The two new events have twenty eight (7+6+5+4+3+2+1) different possible options for their insertion into the sequence and the number of possible sequences has jumped to 2,520 with only four taxa. The number of possible ways to consider untangling a four-taxon fence diagram would be daunting. The whole Riley data set (Fig. 1) has 62 taxa!

Figures 6 and 7 have revealed the rules for biostratigraphic permutations in a form that allows the total to be calculated for any larger number of taxa. Each additional taxon causes the previous size of the problem to be multiplied by the factor:

$$(2n\text{-}1) + (2n\text{-}2) + (2n\text{-}3) + \ . \ . \ . \ + 2 + 1$$

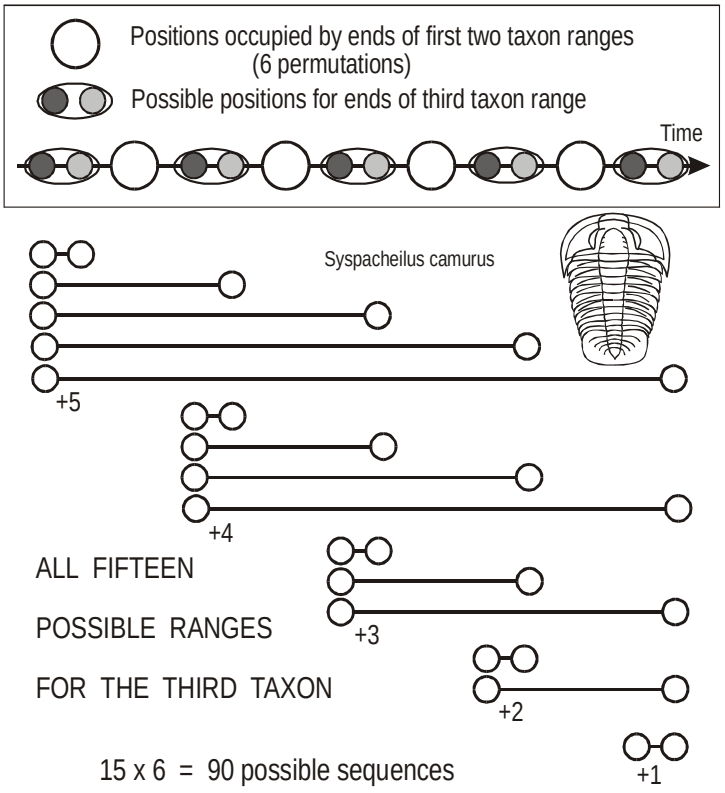

Figure 6. Calculation of the number of biologically feasible sequences of first and last events after the addition of a third taxon to the problem in Figure 3. New positions in the sequence (gray circles) disappear unless occupied. The new positions are shown in pairs, the darker gray for first appearance (FAD) the lighter gray for the last appearance (LAD). The distinction between the two positions in the paired gray circles disappears unless the FAD and LAD of *S. camurus* occupy adjacent positions. Further explanation in text.

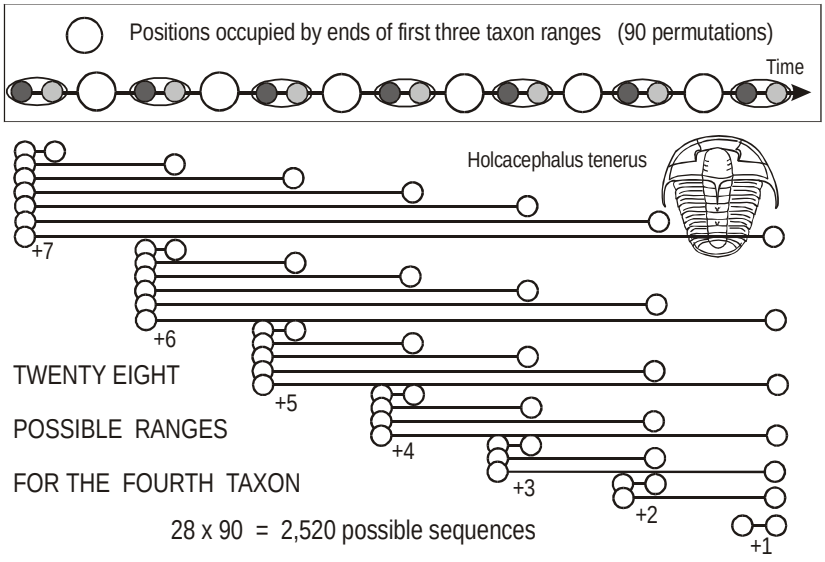

Figure 7. Calculation of the number of biologically feasible sequences of first and last events after the addition of a fourth taxon to the sequencing problem posed by Figure 3. Symbols as in Figure 3.

where *n* is the new total number of taxa and 2n the total number of events. Table 2 depicts how the number of feasible sequences grows to much more than half a billion for only seven taxa.

Although observed coexistences reduce the number of feasible sequences, even the most extreme constraint - requiring that every taxon must coexist with all others - still leaves over twenty-five million possible histories for just seven taxa. Calculating the number of possible sequences with this maximum coexistence constraint for *n* taxa is relatively straightforward. The *n* first occurrences must all lie in the first half of the sequence because no taxon disappears until all have appeared. So, there are *n* factorial ways to arrange all the first occurrences. This number must be multiplied by another *n* factorial, which is the number of ways to arrange the *n* last occurrences in the second half of the sequence.

Figure 8 plots the further increase in the number of permutations up to 100 taxa. Notice that the number of feasible sequences grows to exceed 10^{300}, a colossal number. For the 62 taxa in the Riley data set, the number of feasible sequences surely exceeds 10^{150} regardless of the number of coexistence constraints. To appreciate just how daunting the problem has become, imagine that it takes only one nanosecond to evaluate each sequence in terms of its fit with the field observations in the seven sections in the Riley problem. All of geological time amounts to far less than

Table 2. The growth of the number of biologically feasible sequences with the number of taxa

Number of taxa:	1	2	3	4	5	6	7
Feasible sequences - with no coexistence constraints	1	6	90	2,520	113,400	7,484,400	681,080,400
- with the maximum set of coexistence constraints	1	4	36	576	14,400	518,400	25,401,600

10^{30} nanoseconds. It is clearly beyond the scope of simple mental arithmetic to untangle (optimally) all the contradictions among the 62 taxa and seven sections in Figure 1. Furthermore, even a computer will need an intelligent algorithm that can find the best-fit sequence without conducting an exhaustive search of all possible sequences.

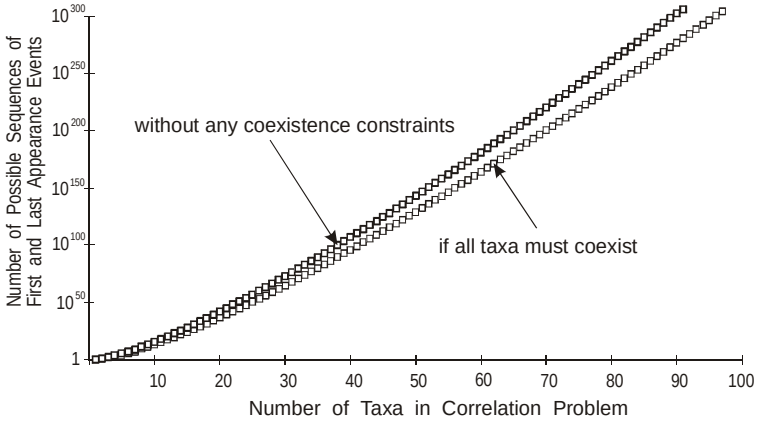

Figure 8. The increase in the number of feasible sequences of first and last appearance events as the number of taxa increases. Realistic levels of coexistence constraint lie between the two curves. The upper curve includes no coexistence constraints; it requires only that each taxon must appear before it can go extinct. The lower curve represents the most constrained case - one in which all taxa are known to coexist with all others.

Correlation problems may include non-biostratigraphic events. These are unpaired events, such as ash falls, for which there is no *a priori* constraint equivalent to the sequencing of first and last appearances of the same taxon. Assuming such event beds are distinctive and positively identified, however, their position in a stratigraphic section cannot be moved up or down like a taxon range-end to fit the section to a particular sequence of events. Thus, these events are constrained to remain in the order of their appearance in stratigraphic sections. First appearances of taxa that lie stratigraphically below an ash bed must be earlier in sequence, because first

appearances may not be adjusted upward. Similarly, a last occurrence is certainly younger than any ash bed that lies below it in a single stratigraphic section. In this way, ash beds that can be matched across several sections serve to constrain the number of feasible sequences to be lower than the totals shown in Figure 8 for biostratigraphic events alone. Ash beds whose relative age has been established by radiometric dating can serve a similar purpose even if they are not known from the same stratigraphic sections.

2. TRADITIONAL BIOSTRATIGRAPHIC INTERVAL ZONES

Traditional zonal biostratigraphy justifiably avoids the nightmares inherent in Figures 1 and 8 by selecting a relatively small number of first and last appearances to define the boundaries of interval zones. Of course, the price for simplifying the problem by culling events is likely to be a loss of resolution. The data set for the Riley Formation includes 124 range-end events; yet only six or seven biozones were traditionally used to subdivide this stratigraphic interval. Although it may be unrealistic to expect all 124 events to be practically applicable to time correlation, Shaw's (1964) graphic correlation showed that reducing them to six was far too conservative. If the exercise is to determine the history of standing diversity, however, then the proper permutation is needed for all 124 events - including the first and last appearances of taxa that are known from only one section.

For very practical reasons, the traditional selection of range ends favors abundant, distinctive, and cosmopolitan taxa. But the most critical requirement is to limit the selection to those events that are almost always found in the same order. Consider how this requirement inevitably leads to the culling-out of long-ranged taxa and closely-spaced events. The longer the taxon range, the greater will be the maximum possible discrepancy between the real and observed range ends - the observed ends might fall anywhere within the real range. The smaller the interval between a pair of events, the smaller will be the discrepancy (in one range or both) that is sufficient to reverse their preserved order.

If few taxa are chosen and they are relatively short ranged and their range-end events are sufficiently widely separated from those of other chosen taxa, then their order of preservation will inevitably appear relatively stable. There may be equally good events that must be excluded from the index set because they are too close to a selected event or two. Nevertheless, exceptions to the most frequently observed order of selected events do occasionally emerge. These anomalies are corrected on an ad hoc basis and comfortably explained by appeal to one or other of the processes that cause

ranges to be under-represented. Although the same biases affect all observed range ends, no attempt is made to correct the others.

Traditional correlation uses the observed range ends, not adjusted range ends. When comparing the sparse tie lines of a traditional fence diagram with a full set of adjusted tie-lines drawn by a computer, it is natural to ask whether the impression of increased temporal resolution might be partly illusory. Remember, however, that the traditional lines cannot be assumed to be precisely located, simply because they do not cross. The precision may be partly illusory in both cases. The computer-assisted correlation addresses this question internally by analyzing the full set of equally best-fit solutions. Traditional biostratigraphy uses independent tests, such as chemo-stratigraphic markers and paleomagnetic reversals, to test for diachronism in zone boundaries. In this way, many traditional zones have been shown to be remarkably reliable. Computer-assisted correlation can easily and more effectively incorporate this same strategy. As explained above, marker beds and age constraints should be added to the list of events for which the computer algorithms seek an optimal sequence. Such fixed markers serve to direct the range adjustments to the more diachronous range ends.

2.1 Assemblage Zones

A significantly different method of building a time scale from biostratigraphic information also deserves attention. Rather than observed range ends, it exploits only the more trustworthy observations of overlapping ranges. These overlaps record the coexistence of taxa. The method groups coexisting taxa into assemblages and then orders these assemblages, from oldest to youngest, as criteria for determining relative age in a succession of coarse assemblage zones. The Wood Committee (Wood *et al.*, 1941) built the sequence of North American Land Mammal Ages this way. The approach has been quantified and automated by Guex (Guex and Davaud, 1984; Guex, 1991) under the term “unitary associations.” Alroy (1992) used essentially the same quantitative approach, developed independently, to improve upon the resolution achieved by the Wood Committee.

This seriation of faunal assemblages does not always produce a dramatic increase in resolution relative to the traditional event-based zones. Guex (1991) extracted eight assemblages for the Riley data set, for example. Nevertheless, his unitary association method is a very powerful approach to correlation problems in which direct evidence of superposition is rather sparse; this problem may arise because the faunas are isolated, or the fossiliferous sections span very short stratigraphic intervals, or the fossil-bearing strata are highly deformed. Radiolarian faunas from cherts at convergent plate boundaries frequently fall into the last category. For an

example of isolated faunas, see the chapter by Webster *et al.* in this volume.

The method of unitary associations is more conservative than our treatment of the Riley example, but it suggests a means to augment our measure of misfit. Unitary associations do not consider range extensions at all; instead they exploit coexistence as the means of both constraint and optimization. Hypothetical sequences of first and last occurrences tend to imply more coexistences than are actually observed. An optimal sequence would minimize the number of coexistences ("conjunctions" in Alroy's terms) that are implied but not observed in any of the faunas. This "excess coexistence" factor can easily be added onto the measure of misfit derived from range extensions. The additional penalty proves to be essential for correlation problems that span a time interval which is much longer than the individual stratigraphic sections (e.g., the graptolite case study, below).

Readers who are wary of relative stratigraphic thickness as a proxy for time may have been uncomfortable with its use as a measure of misfit in our trivial example from the Riley Formation. Unitary associations then seem very attractive because they avoid connections between rock thickness and time. In the next section we explain how constrained optimization can measure misfit without reference to stratigraphic thickness.

3. COMPUTER-ASSISTED CORRELATION

Although the biostratigraphic sequencing problem is substantial, our analysis of a small part of the Riley data set has shown that it should be tractable. For all 124 range-end events, the number of possible sequences becomes dauntingly large, but the rules for selecting the best sequences were found to be few and simple. Iterative application of simple rules is a task well-suited to computers. Kemple *et al.* (1989, 1995) have formalized the rules as a constrained optimization. The CONOP9 software for Windows 95/98/NT platforms (see Chapter 13 and accompanying CD; Sadler, 2000) implements this method and has been used for the remainder of this chapter. There are alternative automated approaches and, to some extent, CONOP9 can mimic them. The method of unitary association (Guex 1991; Alroy, 1992) has already been mentioned. The RASC software (Agterberg and Nel, 1982a,b; Agterberg and Gradstein, 1996, 1999) builds sequences from the most commonly observed pairwise ordering of preserved events. It must exclude events that are observed in very few sections, but is significantly faster than CONOP9 as a result. A comparable approach appears in CONOP9 as an alternative measure of misfit (see below).

3.1 Constrained Optimization

As in our trivial example of two trilobites in two stratigraphic sections, constrained optimization eliminates sequences that fail to reproduce all observed coexistences (constraint) and then searches among the remainder for those that best-fit the field data (optimization). The choice of constraints is obvious. Two critical procedural choices must be made in order to implement an optimization by computer: 1) select a measure of the misfit between possible sequences and the field observations -- the best-fit sequence is likely to vary with this choice; and 2) select a search procedure that chooses the sequence with the best fit -- this decision should determine the efficiency of the search, but not the outcome.

3.2 Measures of Misfit

Several options are available for measuring the misfit between the observed ranges and a possible sequence of events. Our example used an interval measure, formulated from the length of the adjustments needed to make observed ranges fit the proposed sequence. Some other measures are ordinal and based upon the number of locally observed pairwise sequences of events that are contradicted by the proposed sequence (comparable with the ranking and scaling method, Agterberg and Gradstein, 1996). A third class uses the implied but unobserved coexistences (comparable with the unitary association method; Guex, 1991). Compound measures may be produced by summing two or more of these three measures. The more complex the measure, however, the more difficult it becomes to interpret the outcome of the optimization.

The best-fit solution is likely to vary with the measure of misfit. Therefore, the choice should be guided by purpose. For interval measures a critical decision concerns whether the size of the range adjustments will be measured in raw stratigraphic thickness or in numbers of event horizons. The former (penalty = ‘interval’ in CONOP9 syntax) favors the sequences preserved in relatively thick sections; its use is justified where the sedimentary facies are uniform or thicker sections are thought to provide better resolution. The latter (penalty = ‘level’ or penalty = ‘eventual’ in CONOP9 syntax) favors sequences preserved in the most fossiliferous sections; it is generally preferable and certainly advisable when the sedimentary facies and the rate of accumulation are known to vary significantly from section to section. By measuring range adjustments as the number of local events or event horizons that must be “leapfrogged,” it avoids using relative thickness as a proxy for duration and shares some of the conservatism inherent in the unitary associations method.

Ordinal measures of misfit (penalty = 'ordinal' or penalty = 'rascal' in CONOP9 syntax) favor the most frequently observed order of events. They may be minimized faster than interval measures because they make local range adjustments only once, after the best-fit sequence has been found. A solution that minimizes ordinal misfit predicts the most likely sequence of events at a single locality, such as the next well drilled in a program of petroleum exploration (Cooper *et al.*, 2001). It might handle the problems of reworking better than other measures. A solution that minimizes interval misfit estimates the regional sequence of origination and extinction events; it is therefore likely to be better suited to building time scales and solving paleobiological sequencing problems.

3.3 Search Heuristics

The size of a correlation problem is given by the number of events to be placed in sequence and the number of stratigraphic sections in which they have been observed. Figure 8 shows that the size of the stratigraphic correlation problem, as measured by the number of solutions to consider, grows at a faster than exponential rate as the number of taxa increases. The number of events and the number of sections together determine the number of operations required to evaluate the misfit for any one of the model sequences. In practice, the best estimator of total computation time is the number of locally observed range ends. Every case of a taxon that is not found in a particular local section reduces the computation time because it reduces by two the number of locally observed range ends that are candidates for adjustment.

All the efficient options for this optimization search by trial and error, iteratively improving their trial solutions until finding the best fit without the hopeless task of examining all possible sequences. We have yet to find any satisfactory means of implementing the most popular forms of genetic algorithm that involve both mutation and hybridization. Because each event in a biostratigraphic sequence may appear only once, the task of making a viable hybrid from two or more model sequences appears to be less efficient than simply continuing the trial and error process for one sequence. So far, search algorithms that iteratively mutate a single sequence reach the solution fastest. Kemple *et al.* (1995) discuss several options and give a detailed account of "simulated annealing" (Kirkpatrick *et al.*, 1983), the search heuristic which CONOP9 uses most extensively.

For this chapter, we briefly outline the progress of a typical search by simulated annealing. It has three basic requirements: 1) an initial guess, which may be any feasible trial sequence of events; 2) a means of mutating one trial to generate new trial sequences; and 3) rules for deciding whether to

accept the new sequence or return to the previous one and make a different modification. Simulated annealing always uses the same standard rules for accepting or rejecting each trial. But the means of mutating one trial sequence to generate the next must be suited to biostratigraphic sequences if simulated annealing is to converge efficiently on the best fit.

The most efficient means of mutating a sequence selects one event at random from the previous sequence and moves it to a random new position. Sadler (2000) explains other options and their uses. For the early trials, the simulated annealing algorithm accepts quite "bad" mutations that significantly increase the misfit. Bad mutations may lead to a better fit in subsequent trials. More importantly, they prevent the search from becoming trapped at a sub-optimal solution. Throughout the search, however, the rate of acceptance of bad mutations diminishes as the algorithm progressively lowers the maximum size of the temporarily acceptable deteriorations in fit. In effect, this allows an initially rather coarse search strategy to become progressively finer and "home in" effectively on the best-fit solution. The strategy of accepting some bad mutations allows simulated annealing to proceed without hybridization; it is one end-member in the range of evolutionary algorithms.

While searching, CONOP9 displays each trial sequence as one frame in an animated and rather stylized range chart (Figs. 9-inset, 10-inset; if the number of taxa exceeds the number of lines on the screen, a standing diversity curve can be animated instead). At first, the range chart shows that most ranges are long, and all taxa coexist in the middle of the sequence. For its initial guess, CONOP9 has placed all the first occurrences in random order in the first half of the sequence and all last occurrences, independently randomized, into the last half. This simple ploy satisfies all possible coexistence constraints. Of course, such an initial guess fits the empirical data very poorly; and early mutations of the sequence rapidly reduce the misfit (Fig. 9).

The progressive reduction of both the acceptance threshold for bad mutations and the misfit between the trials and the field observations are plotted on top of the animated range chart. Ideally, both curves fall smoothly and exponentially from a steep initial descent in the upper left of the graph to a horizontal finish in the lower right (Fig. 10). The initial value of the acceptance threshold and its exponential reduction must be programmed in advance; they correspond to the initial temperature and the slow cooling schedule of the annealing analogue (Kemple *et al.* 1995). In contrast, the smoothly diminishing returns in the improvement of fit emerge at run-time, but only if the annealing schedule was set appropriately for the size and complexity of the problem. The number of sequences examined in a successful simulated annealing may reach many tens of millions for very

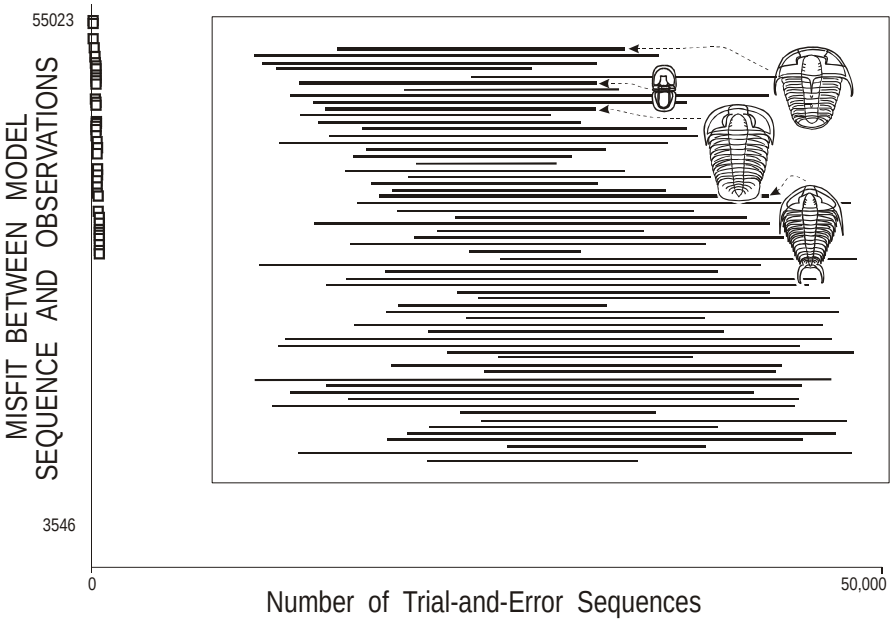

Figure 9. Rapidly improved fit after about one thousand trials in a CONOP9 search for the best-fit solution to the Riley correlation problem, from a random initial guess. This is modified from a snapshot of the computer screen during a CONOP9 search. Two diagrams that are superimposed on the screen have been separated here for clarity. 1) Main graph: open squares in the upper left plot the falling misfit, from top to bottom, as the number of trials increases, from left to right. 2) Inset box: the horizontal lines are relative ranges for the 62 trilobite taxa, first appearance to the left, last appearance to the right. This snapshot shows the ranges corresponding to the latest trial sequence; during a search, they become an "animated range chart" in which range ends continually adjust to depict the current trial. Illustrations of trilobites have been added to locate the four taxa used in Figures 5-7; they would not appear on the screen during an actual CONOP9 run.

large data sets, but is far fewer than the totals charted in Figure 8. The heuristic rapidly finds optimal solutions for moderately large instances (up to 100 taxa) and very good solutions for extremely large instances (in excess of 500 taxa). For the latter, it is possible to improve the solution by re-running the search, but starting from the previous best sequence and with a low rate of acceptance for bad mutations.

4. BEST-FIT INTERVALS

The best-fit sequence found by constrained optimization is seldom a unique solution for a biostratigraphic sequencing problem; several slightly

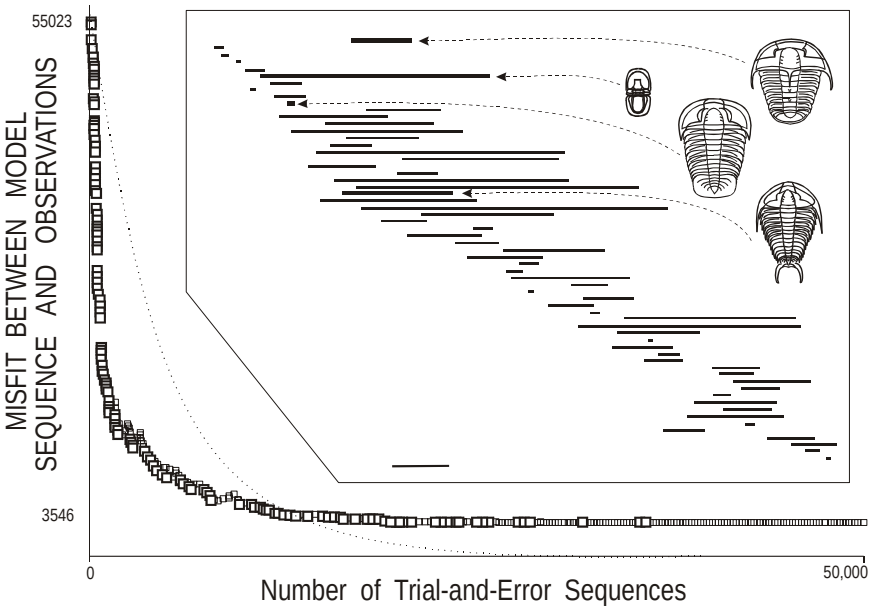

Figure 10. Snapshot of the final trial in a CONOP9 search for the best-fit solution to the Riley data set. Large open squares plot the reduction of misfit with the field observations through a search of 50,000 trial sequences. Small squares indicate that a trial mutation was accepted even though it did not immediately reduce the misfit. At the end of a well-tuned search by simulated annealing, the rate of discovery of better solutions has slowed exponentially leading to a run of unsuccessful trials. The dotted line charts the falling probability of accepting "bad" mutations (increases in misfit; see text). Inset box: taxon ranges in the best-fit solution; compare Figure 9; the diagonal arrangement of the best-fit ranges is an artifact of Shaw's (1964) numbering of taxa that generally proceeded up-section.

different sequences of events typically fit the field observations equally well. Some events may have a constant position throughout this set of equally best-fit sequences; others will be less tightly constrained. The best-fit interval for an event is the range of positions that it takes in the whole set of equally best-fit sequences. The best-fit interval is the flat base of a "relaxed-fit" curve that plots the smallest known misfit as a function of the position of one event across all positions in the overall sequence (Fig. 11).

The operational challenge in mapping out the best-fit intervals, is to find the full set of equally best-fit sequences. Given any one of the best-fit sequences it is straightforward to begin to explore the extent of each best-fit interval; while keeping the remainder of the events in the same relative position, move one event up and down the sequence, recalculating the misfit each step of the way. This strategy can be rapidly completed after a single search, but may not establish the full limits of the best-fit interval. For that, all other events should be allowed to reoptimize around the moving event when it appears to have stepped beyond the best-fit interval.

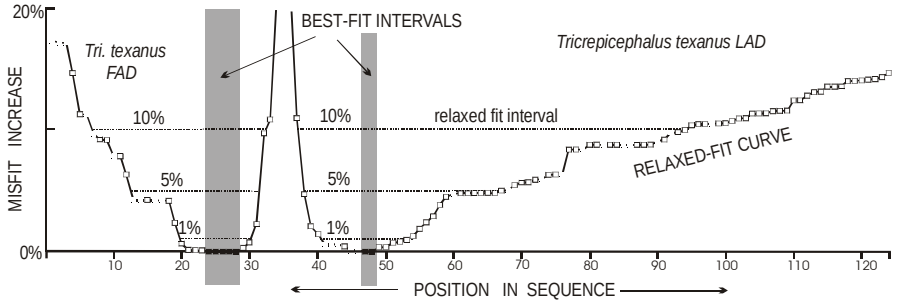

Figure 11. Best-fit intervals and relaxed fit curves for the range ends of *Tricrepicephalus texanus* in the Riley data set. The curves plot the best-known fit as a function of the position of one event in the sequence. Relaxed fit is expressed as a percentage increase in misfit relative to the fit of the optimal sequences. The best fit for the LAD is tighter than for the FAD; but the LAD interval expands faster when the fit is relaxed. Because the FAD constrains the lowest possible position of the corresponding LAD, and vice versa, both curves are asymmetric; the steeper limbs mark the younger end of the FAD curve and the older end of the LAD curve.

For every event and each of its possible relative positions in the sequence of all events, there will be one or several arrangements of all other events that minimize the misfit. It would not be feasible to complete all these minimizations except for small data sets. An exhaustive mapping of the best fit intervals faces the unmanageably huge numbers of possible sequences predicted by Figure 8. Fortunately, a large majority of the overwhelming number of possible sequences lie far from the best-fit intervals (Fig. 12). It is feasible to reduce the efficiency of the simulated annealing heuristic so that the search deliberately dawdles among the near-optimal and optimal sequences close to the region of best fit. Figure 11 was prepared in this fashion: sub-optimal searches were allowed to loop for days, visiting many millions of near-optimal sequences in order to reduce the risk of underestimating the width of the best-fit intervals.

The average size of the best-fit intervals for all the events measures the resolving power of the whole set of N events. The best-constrained events occupy the same position in all best-fit sequences; their best-fit intervals have size 1 because they span only one position in sequence. Totally unconstrained events fit equally well in all N positions; their best-fit interval would have size N. Therefore the mean best-fit interval for all events ranges from 1.0, for the best possible resolution, to N for a completely unconstrained set of events. For the Riley data set, the mean best fit interval is 3.2 positions wide. The reciprocal of the average best fit interval might be a more intuitive measure; it decreases below 1 as resolution deteriorates. But the minimum value would vary from problem to problem, unless the total number of events is rescaled to a constant value based upon 1/N.

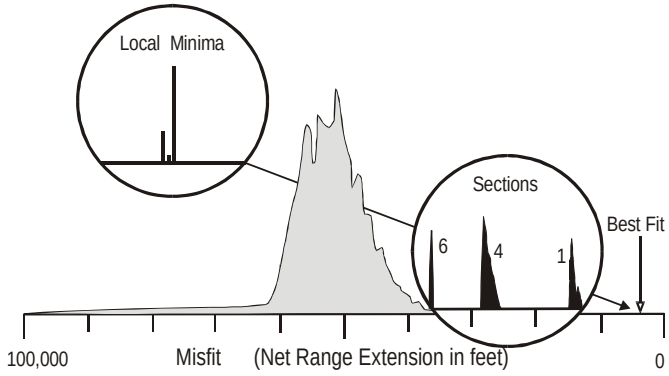

Figure 12. Grey shaded curve: misfit frequency distribution for a million random sequences fit to the Riley data set of 62 taxa and 7 sections. The best-fit sequence, local minima, and sequences that would be generated from local sections are not visible at the scale of the gray curve – they are typically not discovered in one million random mutations. Inset circles greatly exaggerate vertical scale, showing what might become visible in the tail of the distribution after trillions of trials; they are accurate representations, constructed by long sequences of non-random trials that targeted the tail. Black histograms: misfit frequencies for sequences developed from a single section, with missing taxa inserted at random. Section 1 (Morgan Creek) is superior to sections 4 (Little Llano River) and 6 (Potontoc). Local Minima: frequency distribution of outcomes of another set of searches designed to stop at any local minimum encountered in the misfit landscape. In this instance, the tallest of the local minima is the best fit; it captures more of the random searches than the sub-optimal minima.

4.1 Consensus Sequences

The best-fit intervals for a set of events typically overlap with one another to some degree. A sub-set of the events can be selected on the condition that their best-fit intervals do not overlap form a consensus sequence in the sense that they have the same positions, relative to one another, in all the equally best-fit sequences. Although the maximum consensus sequences contain less than the full set of available events, they often exceed the traditional biostratigraphic resolution by an order of magnitude (compare “Traditional resolution” to “Consensus sequence” in Table 3). Sets of events with the same best-fit intervals are alternates forone-another in the consensus sequence. Subsets of events with substantial overlap in their best-fit intervals, but no overlap with other subsets, might be used as the basis for defining the limits of assemblage zones. Thus, best-fit intervals and consensus sequences can place traditional zonation methods on a rigorous footing. Algorithms in CONOP9 rapidly extract consensus sequences once the best-fit intervals have been determined.

4.2 Relaxed-Fit Intervals

Some of the precision of the best-fit measures is likely illusory. This becomes most evident when different measures of misfit generate small differences in the set of best-fit sequences. The best-fit intervals can often be reduced in length simply by combining two measures of misfit into a single optimization criterion. Thus, for any one measure, solutions that almost match the best fit might, for practical purposes, be regarded as equally good. Similarly, knowledge of the thickness of the benthic mixing zone could be used to place limits on the meaningful differences in fit, when these are measured in stratigraphic thicknesses.

Relaxing the best-fit criteria to include all sequences within 0.01% or 0.05% of the best fit would be a conservative approach to the sequencing problem. This allows the intervals to include sequences that are best fit by other measures of misfit. The curves of Figure 11 define all possible relaxed-fit intervals for a pair of events. If these relaxed-fit curves are regarded as cross-sections of a bathtub, then progressive relaxation corresponds to filling the bath and monitoring relaxed-fit intervals by the increasing width of the water surface.

As the relaxation progresses, and larger differences of fit are ignored, the number of non-overlapping best-fit intervals decreases, and the resolving power of the consensus sequence deteriorates. In order to lower the resolution to the level of traditional biostratigraphic zonation, the best fit criterion must be relaxed as much as 20% in some case studies and as little as 0.2% in others (Table 3). This percentage relaxation varies with the number of events and sections in the problem. The size of the total misfit, and any percentage of it, tend to grow with the number of locally observed range ends. Thus, relaxation by a fixed percentage of the best-fit permits the greatest relaxation in instances of the problem with the most sections and taxa. When the purpose is to compare data sets of different size, relaxation should be expressed as a multiple of the average misfit for a single event in the best-fit sequence of each data set. Unlike the more familiar percentage, this *ad hoc* average measure corrects for raw differences in the total number of locally observed events and in the degree of internal contradiction.

4.3 Rating Stratigraphic Sections

Best-fit intervals can measure the resolving power of individual stratigraphic events and, if averaged, the resolving power of whole sets of events. They cannot measure the quality of single sections because best-fit intervals only emerge after the sections have been composited into a single sequence. Raw measures of misfit offer two possibilities to rate stratigraphic

sections. First, individual range extensions may be summed across all ranges observed in a section. This effectively measures the departure of the section from the composite sequence, but must be treated carefully as a measure of the relative quality of different sections. As mentioned above, the total misfit tends to increase with the number of taxa that a section preserves. Every observed taxon adds two range-ends that might need adjustment. Calculating the average extension per taxon can compensate for the correlation between net adjustment and the number of observed taxa.

A better measure would not simply correct for taxonomic richness, however; it would include taxon richness as a positive attribute of a section. The second possible use of raw misfit values achieves this goal by generating a set of sequences of events from the range chart of a single section. Taxa missing from that section are added in random positions, constrained only by the coexistences observed elsewhere. Each sequence in the set can then be assigned a misfit value which is the total range adjustment needed to fit all other sections to this sequence. The whole set produces a frequency distribution of misfit values that is characteristic for the section from which the set was generated (black histograms in Fig. 12). High values of misfit result naturally from two attributes of poor sections: missing taxa and poorly recovered taxa. Sequences generated from poor sections fail to place many of either kind of taxa within their best-fit intervals, the missing taxa, because they are inserted at random into the sequence, and the poorly recovered taxa because their observed ranges are too short, whether as a result of poor preservation or sparse sampling.

5. CASE STUDIES

We illustrate the benefits of computer-assisted correlation by reference to selected features of the four case studies listed in Table 3. The data sets vary considerably in size and complexity, presenting different challenges for the search heuristic. The best-fit solutions for the Riley and Mohawk problems have both been reproduced by numerous relatively short searches. Because the graptolite clade and Taranaki basin problems include much larger numbers of taxa, very long searches have identified several near-optimal solutions that vary in detail but might not include very best solution.

Table 3 reports the length of the searches as a number of trials, because running time varies with the hardware used. Representative running times with 400-600 MHz Pentium III processors and 64-128Mb of RAM are as follows (hardware and solution times have improved considerably since this table was written). The best-known solution for the Riley data set emerges in less than 30 seconds. A fairly good solution for the graptolite clade

Table 3. Statistical Summary of the Selected Case Studies

Case Study	Riley Formation	Mohawk Valley	Graptolite Clade	Taranaki Basin (3 studies)
Location	Texas	New York	World-wide	New Zealand
Problem	correlation	correlation	seriation	correlation with range tops*
Age	Late Cambrian	Late Middle Ordovician	Cambrian - Devonian	Paleocene - Miocene
Duration	~ 10 Ma	< 10 Ma	~ 100 Ma	~ 45 Ma
Traditional resolution	6-7 zones	3-4 zones	52-78 zones	16 stages
Number of sections	7 sections	6 sections	169 sections	8 wells
Clade	trilobites	graptolites	graptolites	microfossils
Number of taxa	62	21	570	87 – 178
Observed local ranges	201	58	2997	508 – 1054
Number of taxa observed to coexist	348 pairs	74 pairs	6881 pairs	1,035 – 5,905 pairs
Rate of pairwise contradiction	18%	23%	24%	58 – 61% (tops and bases)
Number of tuff beds	0	5	22	0
Number of trials in search	40 - 50,000	50 - 100,000	>25 million	1 - 2 million
Minimum misfit	3546 feet	461 meters	7632 levels	23,383 – 68,541 meters
Average misfit per range end	8.8 feet	3.97 meters	1.27 levels (7.6 meters)	48.2 – 34.8 meters
Additional implied coexistences	64 pairs	10 pairs	5295 pairs	1,593 – 6,302 pairs
Conjunction index (Alroy, 1992)	0.84	0.97	0.57	0.39 – 0.48
Consensus sequence	50-60 events	31-40 events	>1000 events	>50 events*
Resolution increase relative to traditional	7 - 10 fold	7 – 10 fold	> 10 fold	> 9 fold
Relaxed consensus that matches traditional resolution	6 - 10% relaxation	25 - 40% relaxation	0.2 - 0.5% relaxation	n.a.*
References and sources	Palmer, 1954; Shaw, 1964; Kemple *et al.*, 1995.	Cisne *et al.*, 1982; Goldman *et al.*, 1994; Sadler and Kemple, 1995.	Sadler and Cooper, in prep.	Cooper *et al.* 2000; Cooper et al, 2001; Sadler, unpubl.

*borehole caving compromises first appearances and the consensus sequence algorithms

emerges in less than two days, but the search may continue to find small improvements for another 10-15 days. To map out the best-fit intervals for the Riley problem, such hardware should be left to loop through suboptimal searches for 2-3

days. For a data set as large as the graptolite clade, the mapping of best-fit intervals should be run continuously as a background task for a month or more.

5.1 Riley Formation, Texas: A Classic Correlation Problem

Trilobite faunas collected by Palmer (1954) from the Upper Cambrian Riley Formation of Texas became a classic data set when Shaw (1964) used the range charts to demonstrate his graphic correlation method. Kemple *et al.* (1995) used Shaw's data in order to test constrained optimization against Shaw's best graphical solution. Shaw's solution has the appearance of considerable subjectivity because of his *ad hoc* justification of more and less reliable range ends. Kemple *et al.* achieved considerable agreement with Shaw's result, without recourse to CONOP9's ability to weight individually every observed range end in every section. They weighted all events equally.

Shaw's graphic method built the solution piecemeal, one section at a time. Several reiterations were needed to converge on a stable solution and Shaw wrote many pages of *ad hoc* justification concerning the local observations that he considered most reliable. Because Kemple *et al.* (1995) reached essentially the same solution without using differential weighting, they argue that Shaw's evaluation of the relative reliability of different events was not really subjective. Rather, it was Shaw's only means to incorporate objective information from the whole data set while evaluating one or two sections at a time. Of course, the computer-assisted optimization uses all the information throughout. Its optimization criteria are explicit in advance and require less than one minute to generate a best-fit solution that unscrambles the fence diagram (compare Figs. 1 and 13).

Shaw used 62 taxa from seven sections, which all span approximately the same interval. Thus, one measure of the size of the problem is 868, the number of local range ends that must ultimately be located. In fact, no section yielded all taxa. Only 402 range ends were actually observed. This smaller number, is a better guide to the time required to compute the misfit for any one geologically feasible sequence of events. Of the observed range ends, only 325 can be usefully adjusted in the course of constrained optimization. Some of the remaining range ends cannot be meaningfully extended because they already lie at the top or bottom of the investigated stratigraphic section. For others, no contradictions arise concerning their placement in sequence because the taxon was found in only one section.

The best-fit sequences for the Riley Formation require the adjustment of

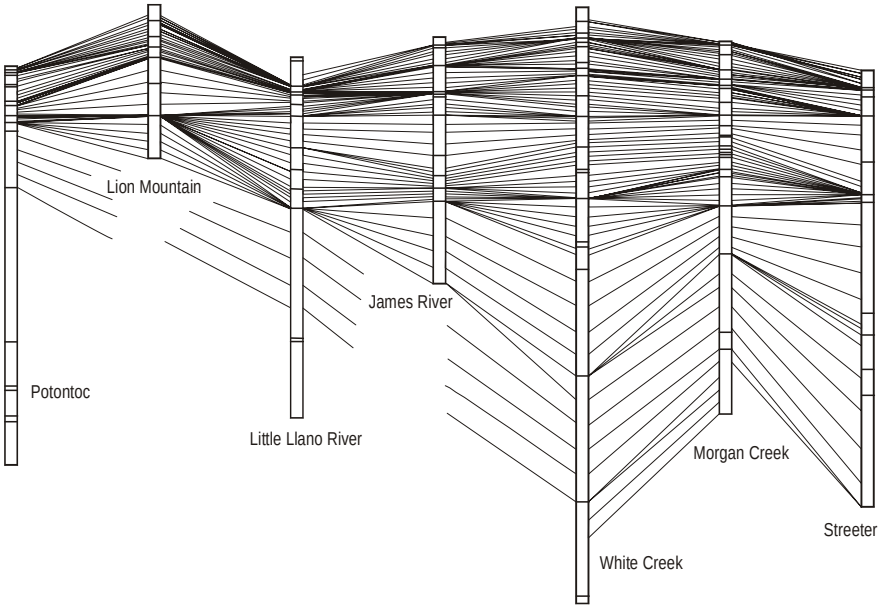

Figure 13. Best-fit fence diagram for the Riley data set. The best-fit solutions place tie lines only at observed event horizons, shown by horizontal bars that divide the section columns. For this figure, the placed levels were subjected to a 5-level moving average that reduces the number of spurious convergences and highlights more robust evidence for condensed levels.

117 of the observed range ends for a total misfit of 3546 feet. Because two thirds of the improvable range ends required no adjustment, the average extension is less than ten feet per event. This compares favorably with the average 21-foot spacing between observed event horizons. The extended ranges imply only 71 more coexisting pairs of taxa than the 348 actually observed coexistences that served to constrain the solution.

How real is the fine resolution implied by all the adjusted tie-lines in Figure 13? Figure 14 summarizes the best-fit intervals for all 124 events. Clearly, a few events may occupy ten or more positions without impacting the best-known fit. For many others, the best fit is achieved in a unique position. In the best resolved portions of the sequence, the best-fit intervals of adjacent events do not overlap. Five such intervals, spanning two to eleven positions in sequence, can be subdivided by using all the first and last appearance events within them. They separate six portions of the sequence in which many larger best-fit intervals overlap with one another. In some ways, these six portions resemble assemblage zones that might be defined on the basis of a suite of events. The number of these gross subdivisions of the sequence resembles the number of traditional zones, suggesting that the traditional resolution is entirely reasonable, given the lack of computer assistance to resolve contradictions.

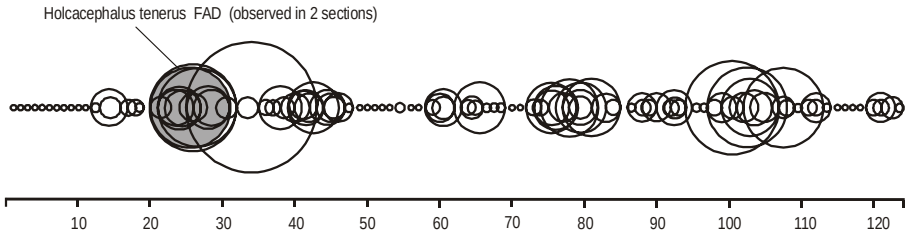

Figure 14. Best-fit intervals for all 124 events in the Riley data set, plotted against an ordinal time-scale with 124 divisions that young from left to right. Each circle represents the range of positions that an event occupies in the set of all equally best-fit sequences. One circle is labeled to identify the event. There is a circle for every event, but some coincide exactly.

Figure 15 illustrates a maximum-consensus sequence -- the largest number of non-overlapping circles that can be extracted from Figure 14. For the Riley Formation, these maximum consensus sequences include 50-60 events, depending on the measure of misfit. This amounts to a seven- to nine-fold increase in resolution relative to traditional zonation. To reduce the consensus sequence to only 6 or 7 events, a resolution comparable with traditional zonations, we must relax the fit criteria to accept all sequences within 6 – 10% of the best fit.

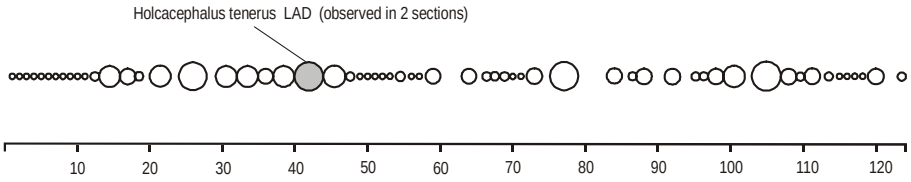

Figure 15. A maximum-consensus sequence for the Riley data set. This figure is developed automatically from Figure 14 by selecting the largest number of non-overlapping circles. Each circle represents one event; one is labeled as an example.

5.2 The Ordovician of the Mohawk Valley: Quantifying Alternate Correlation Models

The late Ordovician graptolite-, conodont-, and K-bentonite-bearing sequences of the Mohawk Valley in New York State alternate between limestone facies (Trenton, Dolgeville, etc.) and shale facies (Utica). The facies boundaries are diachronous. The preferential recovery of graptolites from the shales and conodonts from the limestones leads to severe contradictions between the sequences of first and last appearance events preserved in the local sections.

Not surprisingly, two quite different model correlations have been

proposed for these rocks. One relies primarily on matching K-bentonite beds from locality to locality (Cisne and Rabe, 1978; Cisne *et al.*, 1982a,b); the other places more reliance on the observed ranges of graptolites (Goldman *et al.*, 1994). The differences are significant because they substantially alter the interpretation of the off-shelf Dolgeville limestone facies: in one it is a long-lived shelf-fringing facies and in the other it is a short-lived pulse, perhaps indicative of a sea-level excursion. In the former interpretation, all sections overlap considerably in time; in the latter, two sections, Dolgeville Dam and Nowadaga Creek, are significantly younger than the others (Fig. 16; Sadler and Kemple, 1995).

Sadler and Kemple (1995) showed how constrained optimization can be used to quantify and thus compare the fit of the rival solutions to the field data. They examined 15 graptolite taxa from 6 sections. Here, we present the results for a slightly larger data set with the same six sections but using all 21 graptolite taxa from Goldman *et al.* (1994) and five distinctive clusters of K-bentonites. The six additional graptolite taxa occur in only one section and do not alter the solution. The K-bentonite clusters fall outside the controversial intervals.

The model sequence with the best fit to the observed data from the six Mohawk sections has a net misfit of only 461 m. It closely resembles the solution published by Goldman *et al.* (1994). One may visualize this value as the lowest point in a "landscape" of misfit values. Different latitude and longitude coordinates in this analogy refer to different sequences of events. Elevation increases with misfit. Sequences that resemble the solution preferred by Cisne *et al.* (1982a,b) would come from the bottoms of a cluster of closed depressions in such a landscape. The floors of these depressions have misfit values of 595-610 m, about a one third deterioration in fit, relative to the best-fit sequence. Figure 17 indicates that these closed depressions have a large catchment area; they trap a much larger proportion of optimization runs than the true minimum. As a result, the best-fit sequence is extraordinarily difficult to find for this data set.

Compared with this Mohawk example, most correlation problems have less severe internal contradictions. Typical instances of the biostratigraphic sequencing problem correspond to misfit landscapes with fewer and smaller local minima. These minima are closer to the global minimum and they trap fewer runs.

The relaxed fit curves for the graptolite events (e.g., Figs. 18-19) reflect the high rate of internal contradiction and the existence of the cluster of local minima. Rather than the usual progressive deterioration of fit away from the best-fit interval, the curves for several species (e.g., *Dicranograptus nicholsoni* and *Orthograptus pageanus*) show secondary minima or a pronounced shelf at about a 30% relaxation of fit.

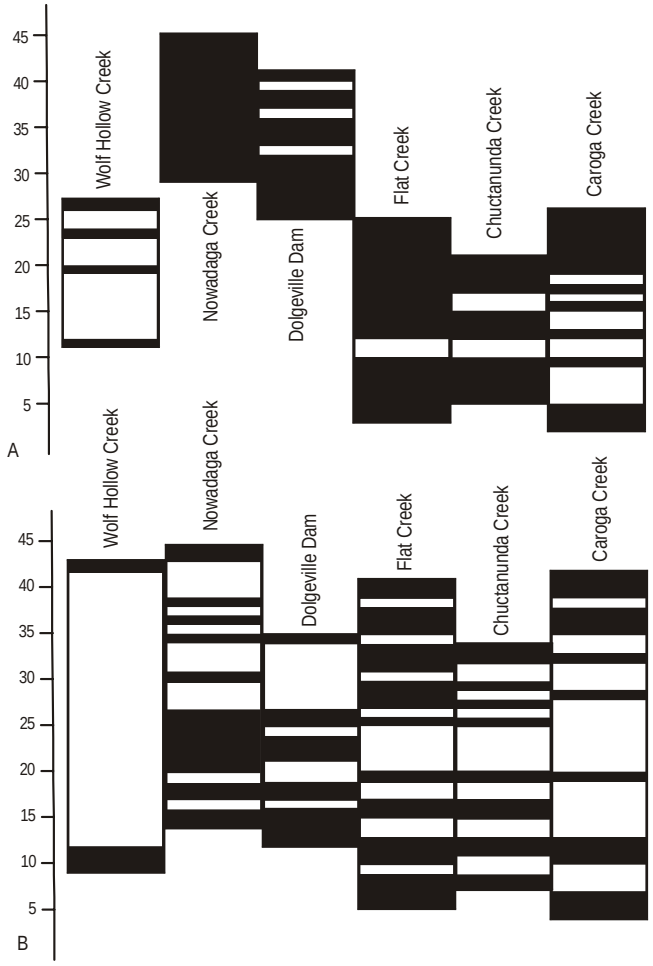

Figure 16. Two correlations of the six Mohawk sections. Vertical "time" axis is the ordinal sequence of events. Black bars indicate range-end events actually observed in the section. Events adjusted to the section top or base are not counted into the range of the section in this plot. Correlation A uses the best-fit sequence with a misfit of 461 meters; it resembles the one proposed by Goldman *et al.* (1994) in its placement of the Dolgeville section. Correlation B is one example of correlations generated by non-optimal misfits in the range 595-610 meters that resemble the placement of the Dolgeville section by Cisne *et al.* (1982).

In addition, there are events with very shallow relaxed-fit curves, indicating that their position in sequence is entirely poorly constrained. The last appearance of *Corynoides calicularis* (Fig. 19) is weakly constrained by field observations. The species was found in only one of the six sections and at only one level; it has no observed coexistences. Although such unique

finds have no local resolving power for correlation, they should not automatically be thrown out of the analysis. First, the best-fit interval

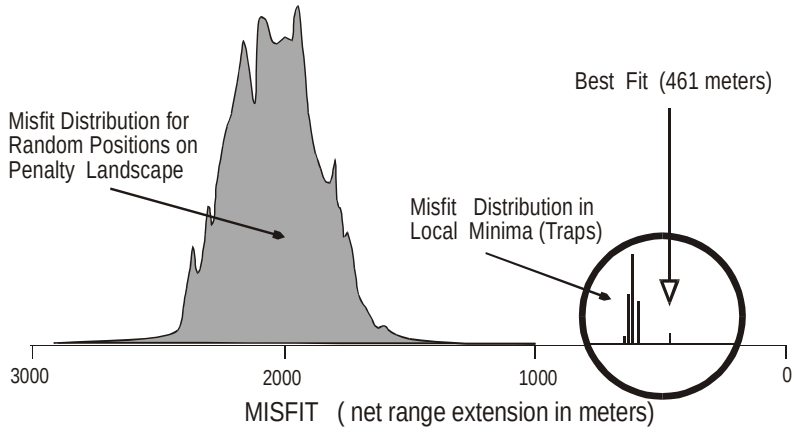

Figure 17. Frequency distribution of misfit for all feasible sequences of the 47 events in the Mohawk data set. Grey curve: misfit distribution for one million random feasible sequences. Five vertical bars in circle with great vertical exaggeration: frequency distribution of local minima found by one thousand runs from random starting positions by greedy algorithms that cannot exit local minima. On the scale used for the gray curve, these misfit values are so rare as to be indistinguishable above the axis of the plot.

provides a measure of the low resolving power that may be preferable to a probabilistic confidence interval. Second, the unique events may deserve inclusion in an accompanying analysis of standing diversity. Third, a rare or unique find may be a key index fossil for zones established in a neighboring faunal province.

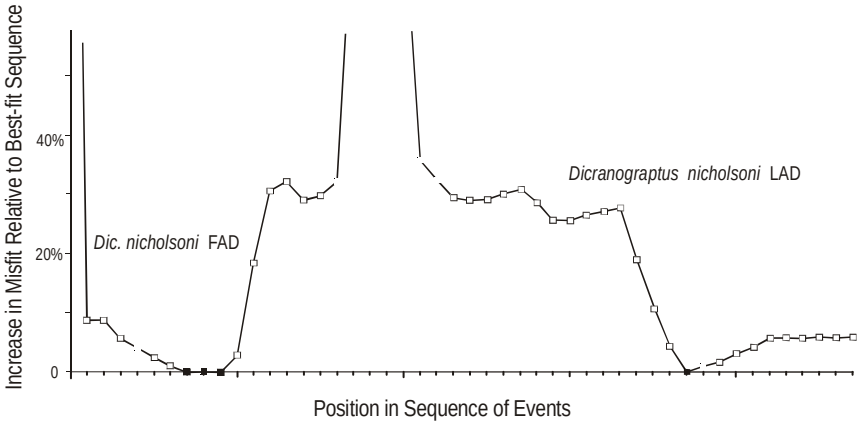

Figure 18. Relaxed-fit curves for the first- and last appearance events of *Dicranograptus*

nicholsoni, found in four of the six sections in the Mohawk Valley data set. Note the secondary minima at about a 30% misfit increase, which compares with the suboptimal sequences that generate solutions of the kind favored by Cisne *et al.* (1982a, b).

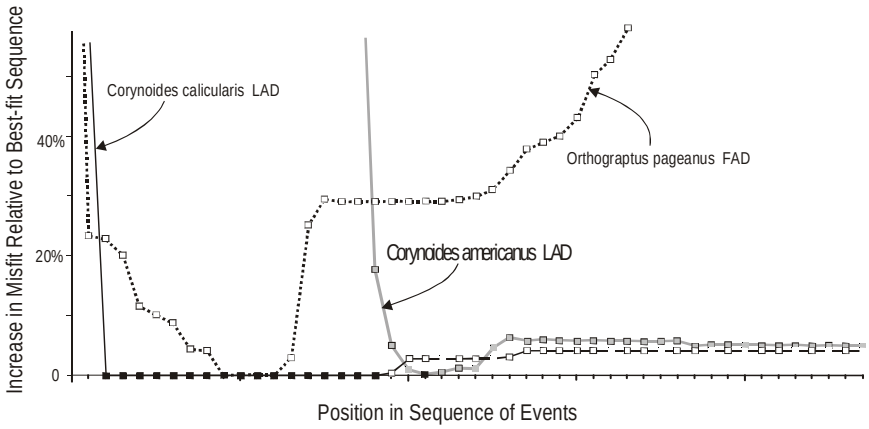

Figure 19. Additional representative examples of relaxed-fit curves for selected graptolite first- and last-appearance events in the Mohawk Valley data set. (Explanation in text.)

Although *Corynoides americanus* is observed in three of the six Mohawk sections, the relaxed-fit curve for its last appearance is still very shallow. This event is preserved in rather inconsistent positions relative to other taxa, but the maximum possible range extension is limited because the highest finds of *C. americanus* are close to the top of the sections in which the species is found. Because observed fossil ranges almost always contain gaps, ranges that end near the limits of measured sections are traditionally considered unsuitable for correlation; another find may lie just beyond the collected interval. The relaxed fit curves confirm that the constrained optimization automatically incorporates this stratigraphic wisdom.

Measures of misfit can be designed which will continue to increase when ranges are extended beyond the limits of a section to higher or lower positions in the regional sequence of events. When added to the primary measure of misfit, these secondary measures tighten the relaxed fit curves. Such secondary measures are justified when the individual sections are so short, relative to the time interval represented across all sections, that some pairs of sections do not overlap in time. Whether or not the Mohawk data fit this criterion depends upon the preferred correlation model. The best-fit sequence (Fig. 16A) indicates that a secondary measure would be justified.

Several different formulations are possible for a misfit measure that discourages the extension of ranges beyond the measured section. The simplest of them minimizes the net length of all the ranges, as measured by

their span of positions in the model sequence; it favors hypothetical sequences with short-lived taxa, but does necessarily improve the fit with the field observations.

A better option borrows from Guex' (1991) and Alroy's (1992) use of coexistences. As a taxon range is progressively extended, it is likely to overlap with an increasing number of other taxon ranges. Thus, the extended ranges tend to imply a greater number of pairs of coexisting taxa than have actually been observed. The number of implied coexistences minus the number of observed coexistences is another measure of misfit between the field observations and a hypothetical sequence of range-end events. Such a measure of misfit manages to limit the growth of poorly constrained ranges beyond the local stratigraphic sections and, at the same time, reinforces the fit with observed coexistences. For the Mohawk problem, addition of such an auxiliary measure of misfit to the primary measure favors solutions of the kind advocated by Goldman *et al.* (1994). It alters the misfit landscape in such a way that the best-fit solution is found more readily. But some searches still tend to get stuck in the local sub-optimal minima in the penalty landscape, which represent solutions like the ones proposed by Cisne *et al.* (1982a, b). These "traps" persist until the auxiliary measure, based on coexistences, is assigned more than twice the weight of the interval measure of misfit.

The full set of best-fit intervals for the Mohawk problem (Fig. 20) reveals that the interval of poorest resolution lies at the beginning of the sequence. The top of the sequence may be only marginally better constrained. As the fit is relaxed, several best-fit intervals near the top of the sequence expand very rapidly. A 5% relaxation of fit would be excessive for well constrained biostratigraphic data sets of this size. For the Mohawk Valley, however, we have seen that the published alternative interpretation corresponds to about a 30% relaxation.

Solutions for many instances of the correlation problem can be improved by adding more fossil clades. Again, the Mohawk problem proves to be more difficult than usual. The addition of the conodont taxa to this correlation problem does not improve the solution (Sadler and Kemple, 1995). The conodont taxa appear to be longer ranging than both the measured sections and the graptolite taxa. Because the conodonts are restricted to the limestone facies, many different limestone intervals appear to record the same suite of ranges. Accordingly, when conodonts are included the CONOP9 algorithms tend to produce a correlation of the limestones that resembles a simplistic lithostratigraphic correlation. Evidently the addition of more taxa does not necessarily improve the resolution unless the taxa are short-lived relative to the measured sections and any facies that preferentially preserves them. Better solutions to the

Mohawk correlation problem are emerging from improved correlation of K-bentonite beds. C. E. Mitchell and his colleagues (pers. comm.) have considerably increased the number of K-bentonites that can be traced in the field and matched from section to section by chemical fingerprinting.

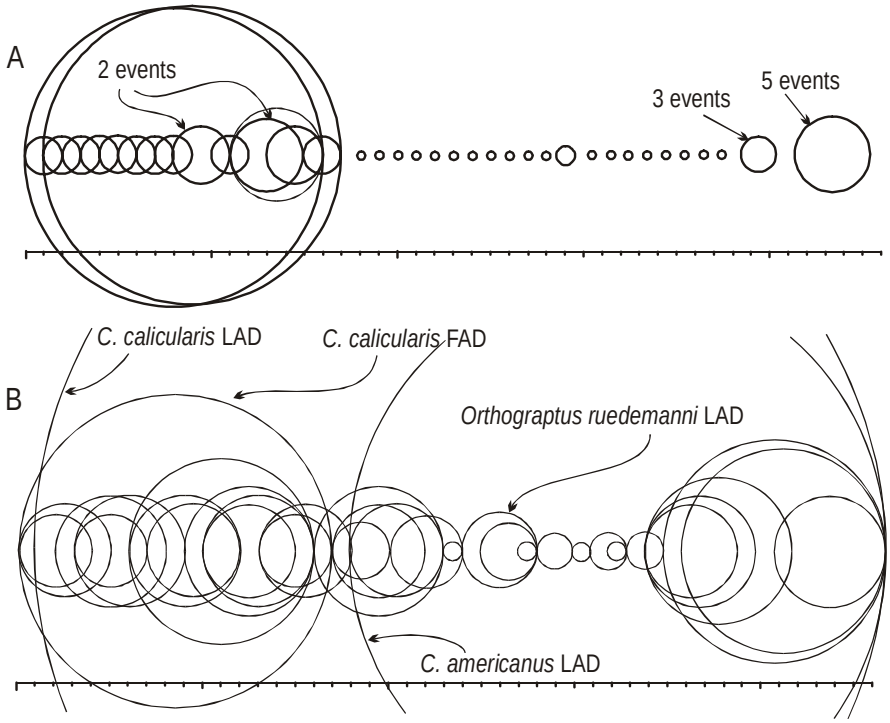

Figure 20. Best-fit intervals (A) and 5% relaxed fit intervals (B) for the Mohawk events plotted as a sequence of overlapping circles. Each event is represented by a circle that extends from the earliest to the latest position that it occupies in the set of best-fit sequences (A) or the larger set with misfit values up to 5% higher than the best-known fit (B). The horizontal scale is the ordinal sequence of events, from the oldest at the left end. A is labeled to show positions in the sequence where two or more circles are exactly coincident.

5.3 The Graptolite Clade: High-Resolution Time Scales and Diversity Curves

This is the largest of the selected case histories. The original purpose was to build an end-Cambrian through early Devonian time scale upon the ranges of 570 graptolite taxa from 169 sections globally. Twenty-two dated K-bentonites provide absolute age control. The search algorithm incorporated these K-bentonites into all the trial sequences according to their

associated faunas. The faunas may be preserved within, above, or below the dated beds; any of these situations allows the dated bed and the faunas to be entered as a short stratigraphic section. It did not matter that the dated beds were very rarely found in long stratigraphic sections with good range charts. Associated conodont faunas could be used as proxies for graptolites if other published sections identified which graptolite taxa could overlap with these conodont ranges. CONOP9 offers several options for building composite sections that scale the intervals between events by reference to the local stratigraphic thicknesses. With radiometric dates included in the best-fit sequence it is possible to identify which of these composites is an acceptable approximation for a time scale.

For this chapter, the case history serves to explain the special problems associated with seriation. Because the individual sections span only a fraction of the total time interval from Cambrian to Devonian, most pairs of sections share no taxa and do not overlap at all in time; they must be stacked in series. A typical correlation problem produces a parallel arrangement of sections with considerable overlap. The term “seriation” better describes the graptolite clade problem. It requires the secondary measure of misfit provided by the number of coexistences that are implied by a trial sequence but represent an excess over those coexistences actually observed in the field.

Good solutions emerge for correlation problems when using only the positive evidence; i.e., misfit is measured solely in terms of the extension of observed local ranges. No additional misfit accrues for inserting a taxon that is completely missing from a given section. This negative evidence becomes necessary in seriation. To understand why, imagine an extreme situation in which this case history includes Ordovician and Devonian sections, but omits the intervening Silurian. Without the Silurian sections, the constrained optimization would still find the best-fit sequences within the Ordovician taxa and within the Devonian taxa. But it would arbitrarily interleave the two solutions into a single sequence that is stratigraphically ridiculous. A stratigrapher would instinctively avoid this error by using the negative evidence that the Ordovician sections and the Devonian sections have no taxa in common. Even with the Silurian included, some arbitrary interleaving is still likely to occur, because the Ordovician taxa are relatively provincial and the available sections are unevenly distributed through the time interval. For example, too many of the measured sections terminate at the Ordovician-Silurian boundary.

Augmenting the interval measure of misfit in proportion to the excess coexistences removes the arbitrary interleaving. Even if the interleaving would eventually be removed by a constrained optimization of positive evidence alone, the augmented measure rids the search of this problem much

sooner, making the optimization far more efficient and less liable to terminate in a local sub-optimal minimum.

There remains only the question of how heavily the auxiliary measure should be weighted. For this seriation problem, one tenth to one-hundredth of the number of implied additional coexistences proved to be a sufficient weight for the negative evidence. If the weight is too high, it tends to force apart sequences of taxa from genuinely coeval geographic provinces. The coeval provinces are held together in the trial sequences by a few shared taxa and by those rare sections that are long enough to extend through changes in provinciality. Therefore, it is essential that the misfit due to coexistences not outweigh the misfit due to range extensions of the cosmopolitan taxa; otherwise, coeval provinces can be teased apart. CONOP9 does not yet have a means to consider geographic proximity and facies similarity; these legitimate aspects of stratigraphers' expertise can be captured only by running subsets of the sections to generate composite sequences for individual provinces and facies. To minimize such problems in this instance, the graptolite seriation was restricted to off-shelf sections in slates and shales.

When augmented by a small fraction of the coexistence excess, the total misfit measure gains in precision and discriminating power. This tends to reduce the width of the best-fit intervals and steepen the relaxed-fit curves (at least the portions close to the best-fit interval). As a result, the consensus sequence for the graptolite clade includes over 1000 events whose best-fit intervals do not overlap. A slightly relaxed fit interval (0.01 - 0.05%) might be more realistic in this case. But graptolites do offer uncommonly high resolving power, even with traditional biostratigraphic interval zones.

For building time scales and for questions of the history of global biodiversity, the computer-assistance brings a significant advantage. Most traditional approaches to these problems are forced to operate at the resolution of zones and stages. For diversity studies, binning the counts by zones may generate spurious peaks and troughs because range ends do not correspond to the zone boundaries and the zones have unequal length. For time scales, radiometric dates are often interpolated at the scale of stages. The computer-assisted process avoids the limitations of zones and stages by operating at the scale of individual events. This is closest to the scale of the raw field observations and removes one level of subjectivity and possible error.

The graptolite case study now includes over 1100 taxa and more than 200 stratigraphic sections. The results are too voluminous to present in detail here. Figure 21 serves as an example. It displays an interpretation of the history of graptolite biodiversity that follows directly from a best-fit sequence. Biodiversity is incremented by one for every first-appearance

event and decremented by one for every last appearance to produce a continuous, high resolution, interval free curve. The time scale for Figure 21 is generated by averaging the stratigraphic separations of the events, across all sections, after ranges have all been adjusted to fit the best sequence

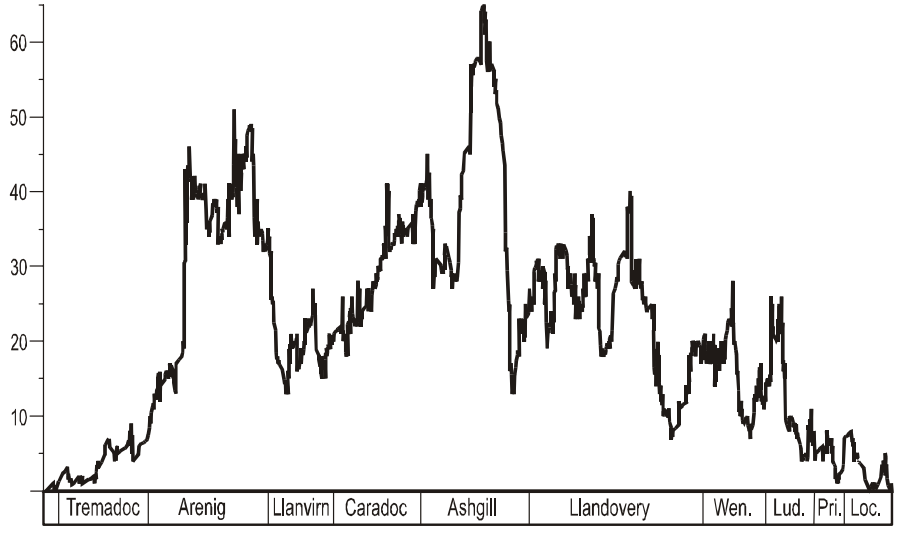

Figure 21. Diversity history of the graptolite clade according to one best-fit sequence for the data set described in Table 3, augmented by all the graptolite taxa found at only one locality. The total taxon count in this diagram is 1136 graptolite species and subspecies. A running total of first appearances minus last appearances determines the standing diversity (vertical axis). The horizontal time axis is a scaled composite section, based on average stratigraphic separations, which gives a good linear regression when the best-fit positions of dated K-bentonites are plotted against a time scale. Stage boundaries were placed according to the position, in this scaled best-fit composite sequence, of the traditional index fossils.

and the section thicknesses have been rescaled to reflect their range of events in the best-fit sequence. This is one of ten scaling options in CONOP9; it produces the best linear correlation with the 22 radiometric dates in this instance.

Traditional binning of diversity data into zones generates discontinuous histograms and introduces severe bias (Foote, 2001). The bias can be eliminated by counting taxa that cross zone boundaries, but the results remain discontinuous and their reliability difficult to quantify. The "continuous" diversity curves described here inherit measures of their quality and reliability directly from the misfit values attached to the sequences from which they were generated. By superimposing diversity curves from all the best-fit sequences, it is possible to generate a reliability interval for diversity curve. By examining the results of progressively

relaxed fit, it is possible to quantify the relative reliability of different peaks and trough in the curve according to the level of relaxation at which they disappear. CONOP9 offers the option to examine the progress of the diversity curve during the search for the best-fit sequence. In this "animated diversity curve" the most robust features emerge at high values of misfit; the least robust features remain volatile to the end of the search for the best-fit sequence. The diversity curve for the graptolite clade is initially unimodal, with a single maximum. The next feature to emerge in most runs is the late Ordovician extinction event.

5.4 The Taranaki Basin: Integration with Seismic Stratigraphy

Eight wells from the Taranaki Basin, off the west coast of North Island New Zealand, were initially correlated to compare three different automated procedures for biostratigraphic correlation – RASC/CASC, CONOP, and GRAPHCOR (Cooper *et al.*, 2001). A line of three of the wells was used to compare the results from CONOP9 with seismic cross-sections (Cooper *et al.*, 2000). The published studies used only 87 of 573 microfossil ranges available for the eight wells. The severe cull was necessary to ensure that all events were found in at least four wells, a precaution to ensure robust results from the RASC program. For this chapter we repeated the optimization with a larger data set that included the 178 most reliable taxa. Table 3 summarizes results for both data sets.

Unlike the stratigraphic sections in the three case histories discussed above, the Taranaki wells do not necessarily record the bases of ranges properly. The taxa are microfossils and some may have been recovered from sediment that fell to anomalously low levels in the well bore during drilling. Accordingly, Cooper *et al.* (2001) omitted the first appearance events from their analysis. Because the CONOP9 algorithms then required that all ranges have two ends, the bases were assigned zero weights so that they would not influence the solution.

Fortunately, some coexistence constraint can be applied without recourse to the first-appearance events. If a pair of last appearances is observed in the opposite order in two different wells, then it follows that the corresponding ranges must overlap. Nevertheless, the unconstrained bases reduce the number of coexistences that can be proven. They also complicate the construction of best-fit curves for the last appearances. The first appearance events do not appear randomly in the best-fit sequences but tend to dominate the lower portions of the sequence, where they bloat the best-fit intervals of other events. Accordingly, this case study has focused on the use of the untangled fence diagrams, in which the tie lines for first appearances are

simply omitted.

The constrained optimization adjusts range ends and inserts missing events. The resulting fence diagrams not only untangle the observed tie lines, they also contain tie-lines that were previously missing. As a result of the increased number of ties generated by CONOP9, as compared with traditional New Zealand stage boundaries, the convergence and divergence of tie lines in the fence diagrams reveal patterns that resemble the condensed sections, lapouts, and sequence boundaries in seismic sections (Fig. 22).

The mathematically smallest misfit with the field observations is never achieved by placing tie lines in the fence diagrams at intermediate positions between observed first- and last-appearance horizons. For more insightful comparison with sequence-stratigraphic correlations and seismic lines, however, Cooper *et al.* (2000) found it useful to adjust the positions of tie-lines in the fence diagrams by smoothing the distances between adjacent lines. Moving average windows have been used that encompass the groups of three, five, or seven adjacent tie lines. The smoothing does not alter the sequence of tie lines and the corresponding events; it adjusts only the stratigraphic distances between them. Smoothing has two useful effects for the visual inspection of possible sequence boundaries; it eliminates convergences that give the appearance of hiatuses where support is poor (few tie lines) and separates coincident tie lines to sharpen the appearance of hiatuses and condensation where support is strong.

The term “support” refers to the number of adjusted range ends that are placed at the same level. In order to fit observed ranges to the model sequences with minimum adjustments, we have seen that computer algorithms extend ranges only to stratigraphic horizons at which other events have been observed. This exaggerates the number of convergences of tie lines in sections where the sampled horizons are sparse. The local stratigraphic distances between adjacent samples may be regarded as uncertainty intervals on the observed range ends in individual sections - additional collecting horizons might have extended the range ends into these intervals. The best-fit sequence can improve substantially upon these uncertainty intervals by compositing information from all sections; but the unsmoothed fence diagram exposes the limitations of individual sections. In effect, smoothing makes a somewhat arbitrary, but more realistic, placement of the tie lines by moving them into the uncertainty intervals, while retaining their order in the best-fit sequence.

Where fossiliferous horizons are wide-spaced and diversity is low, evidence of real convergence must be weak. By spreading tie lines into the uncertainty intervals between collecting horizons, the smoothing effectively eliminates convergences that are supported small numbers of lines and likely to be artificial; e.g., lower parts of the fences in Figure 22. Convergences of

more than three, five, or seven tie lines remain evident after three-, five-, and seven-point smoothing, respectively.

Where the faunas are rich and collecting horizons are closely spaced, the unsmoothed fence diagrams have a different problem: many tie-lines may lie

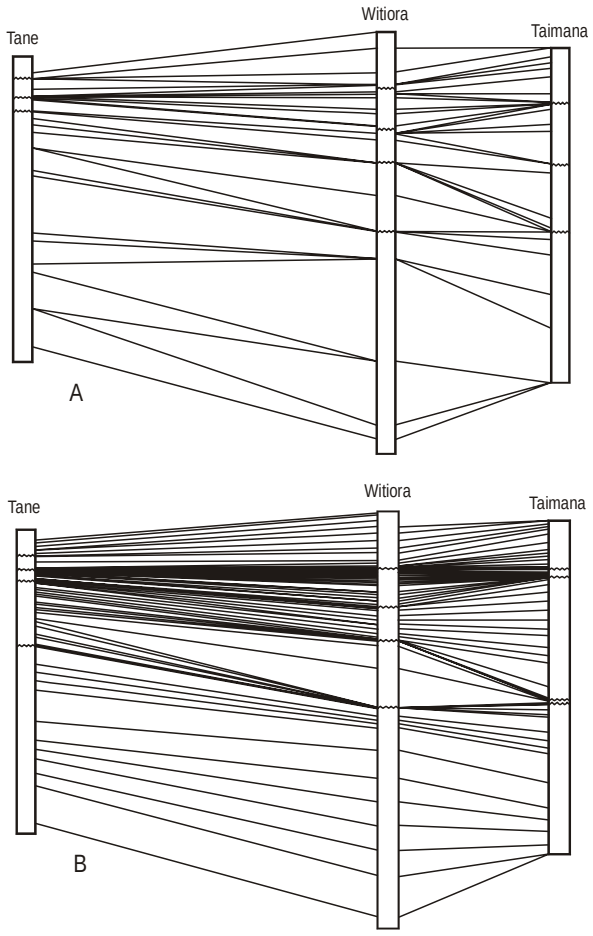

Figure 22. Unscrambled fence diagram of three wells from the Taranaki Basin without (A) and with (B) 7-point smoothing of the elevations of event horizons. The unsmoothed diagram places tie-lines to minimize range extensions; it does not reveal where many lines coincide. The smoothed diagram allows tie-lines to move into intervals between collecting horizons; it eliminates weak convergences and reveals the levels of strong condensation and convergence.

hidden, directly beneath other lines that connect the same horizons. Smoothing causes some of the hidden lines to separate and diverge, but the closely spaced event horizons prevent them from separating very far. Not all the coincident lines are separated; their number is still limited by the number

of points in the moving average. Thus, the smoothed diagrams retain the patterns of condensation and convergence that point to the best-supported and most likely horizons of hiatus or sequence boundaries. Cooper *et al.* (2000) had considerable success matching these convergences with the lapout patterns of seismic reflectors and the lithostratigraphic evidence of hiatus. Indeed, the fit with the high-quality portion of their seismic data was so good that they were emboldened to use the biostratigraphic fence diagram to correct their initial interpretations where the seismic quality was poor and faults complicated the seismic images.

6. CONCLUSIONS

Computer-assisted biostratigraphy may increase resolution by an order of magnitude relative to traditional biostratigraphic interval zones and assemblage zones (Fig. 23) because it solves for the sequence of so many more first- and last-appearance events. The preferred solution is a sequence of events that is characterized by the minimum misfit when compared with all the locally observed taxon ranges. After each measured section has been adjusted to match the solution, the misfit may be summarized by three conceptually different measures. Each totals up a different aspect of the minimum necessary adjustments: the extensions of the local taxon ranges; the locally observed pairs of events whose order is reversed; or the number of additional coexistences that are generated by the adjustments. A weighted sum of the first and third options can solve seriation problems in which the individual measured sections are much shorter than the total time interval under investigation. The coexistence term should be weighted less than the interval term.

Very simple rules govern computer-assisted searches for an optimal sequence of first and last appearance events that minimizes misfit with the observed fossil ranges. The computer algorithms invert the problem and search through possible solutions by trial and error, using the same few rules that guide manual comparison of alternative sequences. Computer memories allow bewilderingly large data sets to be analyzed. Efficient heuristic search procedures find the best of all possible solutions even when these solutions are far too numerous for an exhaustive examination.

The misfit assigned to the best-fit sequence may be apportioned by event and/or section to reveal the sources of discrepancy with the field observations. Misfit measures may also be used to compare the quality of individual sections in a way that balances the number and fidelity of the preserved taxon ranges.

Because they use available taxon range-ends, including those whose

relative age is rather poorly constrained, the computer-assisted solutions are, of course, not unique. The same observed ranges may allow a substantial set of equally well-fit solutions. Best-fit intervals map the range of positions

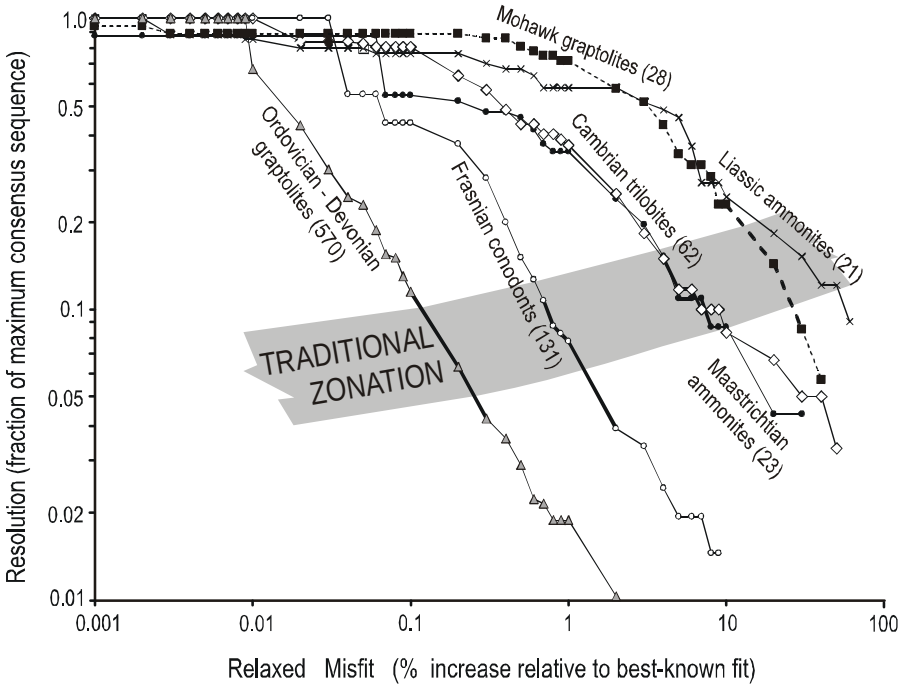

Figure 23. Comparison of the resolution achieved by traditional and computer-assisted biostratigraphy. Computer-generated sequences are charted as the loss of resolution, in terms of the fraction of events that remain in the consensus sequences (vertical axis), as the misfit is progressively relaxed relative to the best-known fit (horizontal axis). Heavy portions of the curves indicate where the numbers of events in the consensus sequence correspond to the numbers of traditional subdivisions. In addition to case histories discussed above, curves are shown for Frasnian conodonts (Klapper, *et al.* 1995 and additional written communications), Upper Liassic ammonite genera (Arkell, 1933, and references therein), and Maastrichtian ammonite species (Macellari, 1984, 1986; Ward and Kennedy, 1993). Numbers in parentheses indicate how many taxa were included in each optimization. The gray band attempts to generalize the resolution achieved by traditional zonation.

that a single event may occupy in the set of equally best fit sequences. After all the best-fit intervals are reduced to a consensus sequence, which retains only the non-overlapping intervals, an order of magnitude more events may remain than in a traditional biostratigraphic zonation. The resulting fence diagrams have far more detail than traditional correlations. Contradictions in the field observations have been resolved by a conservative adjustment of all

the tie lines in the fence, rather than mere culling of lines.

By retaining so many more tie lines, computer-assisted fence diagrams, facilitate comparison with sequence-stratigraphic models and seismic lines. By working at the level of individual observed events, computer-assisted time scales and diversity curves remove the biases otherwise imposed by arbitrary stages and zones. The most significant advantage, however, may be the ability of computer-assisted methods to provide reproducible results and quantify many aspects of the solutions to correlation and seriation problems. The best-fit intervals, for example, provide a natural, but previously unattainable, means to quantify the resolving power of all the individual first and last appearance events.

7. ACKNOWLEDGMENTS

Development of the CONOP software is currently supported by NSF grant EAR9980372 to Sadler. Two of our case studies were facilitated by the generous support for visiting scientists provided by the New Zealand Crown Research Institute for Geological and Nuclear Sciences. Peter Harries, Gilbert Klapper and Charles Marshall provided valuable assistance by drawing our attention to passages in the manuscript that were in need of more lucid explanation.

8. REFERENCES

Agterberg, F. P., and Gradstein, F. M., 1996, *RASC and CASC: Biostratigraphic Zonation and Correlation Software, Version 15*.

Agterberg, F. P., and Gradstein, F. M., 1999, The RASC method for ranking and scaling of biostratigraphic events, *Ear. Sci. Rev*. **46**:1-25.

Agterberg, F. P. and Nel, L. D., 1982a, Algorithms for the ranking of stratigraphic events. *Comp. Geosci*. **8**:69-90.

Agterberg, F. P. and Nel, L. D., 1982b, Algorithms for the scaling of stratigraphic events. *Comp. Geosci*. **8**:163-189.

Alroy, J., 1992, Conjunction among taxonomic distributions and the Miocene mammalian biochronology of the Great Plains, *Paleobio*., **18**:326-343.

Anderson, M. K., 2001, Quantum computing: Souped-up software gets a virtual test, *Science* **292**:419

Arkell, W. J., 1933, *The Jurassic System in Great Britain*, Oxford University Press.

Cisne, J. L., and Rabe, B. D., 1978, Coenocorrelation: gradient analysis of fossil communities and its applications in stratigraphy, *Lethaia* **11**:341-364.

Cisne, J. L., Karig, D. E., Rabe, B. D., and Hay, B. J., 1982a, Topography and tectonics of the Taconic outer trench slope as revealed through gradient analysis of fossil assemblages, *Lethaia* **15**:229-246.

Cisne, J. L., Chandlee, G. O., Rabe, B. D., and Cohen, J. A., 1982b, Clinal variation, episodic

extinction, and possible parapatric speciation: The trilobite *Flexycalymene seneria* along an Ordovician depth gradient, *Lethaia* **15**:325-341.

Cooper, R. A., Crampton, J. S., Raine, J. I., Gradstein, F. M., Morgans, H. E. G., Sadler, P. M., Strong, C. P., Waghorn, D., and Wilson, G. J., 2001, Quantitative biostratigraphy of the Taranaki Basin, New Zealand - a deterministic and probabilistic approach, *AAPG Bull*.

Cooper, R.A., Crampton, J.S., Uruski, C.I., 2000. The time-calibrated composite - a powerful tool in basin exploration, in: *Proceedings of the 2000 New Zealand Petroleum Conference, Christchurch, New Zealand*, Ministry of Commerce, Wellington, pp. 346-354.

Dell, R., Kemple, W. K., and Tovey, P., 1992, Heuristically solving the stratigraphic correlation problem, in: *Proceedings of the Institute of Industrial Engineers First Industrial Engineering Research Conference*, pp. 293-297.

Goldman, D., Mitchell, C. E., Bergstrom, S. G., Delano, J. W., and Tice, S. T., 1994, K-bentonites and graptolite stratigraphy in the Middle Ordovician of New York State and Quebec: A new chronostratigraphic model, *Palaios* **9**:124-143.

Guex, J., 1991, *Biochronological Correlations*, Springer Verlag, Berlin.

Guex, J., and Davaud, E., 1984, Unitary associations method: Use of graph theory and computer algorithms, Comp. Geosci. **10**:69-96.

Kemple, W. G., Sadler, P. M., and Strauss, D. J., 1989, A prototype constrained optimization solution to the time correlation problem, Geol. Surv. Can. Pap. **89-9**:417-425.

Kemple, W. G., Sadler, P. M., and Strauss, D. J., 1995, Extending graphic correlation to N dimensions: The stratigraphic correlation problem as constrained optimization, in: *Graphic Correlation* (K. O. Mann and H. R. Lane, eds.), SEPM Sp. Pap. **53**:65–82.

Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P., 1983, Optimization by simulated annealing, *Science* **220**:671-680.

Klapper, G., Kirchgasser, W, T., and Baeseman, J. F., 1995, Graphic correlation of a Frasnian (Upper Devonian) composite standard, in: *Graphic Correlation* (K. O. Mann and H. R. Lane, eds.), *SEPM Sp. Pap.* **53**:177–184.

Macellari, C. E., 1984, Late Cretaceous stratigraphy, sedimentology, and macropaleontology of Seymour Island, Antarctic Peninsula, Unpubl. Ph. D. dissertation, The Ohio State University.

Macellari, C. E., 1986, Late Campanian-Maastrichtian ammonite fauna from Seymour Island (Antarctic Peninsula), *Paleont. Soc. Mem*. **18**:1-55

Palmer, A. R., 1954, The faunas of the Riley Formation in Central Texas, *J. Paleont*. **28**:709-786.

Sadler, P. M., 2000, *Constrained Optimization Approaches to the Paleobiologic Correlation and Seriation Problems: A Users' Guide and Reference Manual to the CONOP Program Family, Version 6.5*, University of California, Riverside.

Sadler P. M. and Kemple, W. G., 1995, Using rapid, multidimensional, graphic correlation to evaluate chronostratigraphic models for the Mid-Ordovician of the Mohawk Valley, New York, in: *Ordovician Odyssey: Short Papers for the Seventh International Symposium on the Ordovician System.* (J. D. Cooper, M. L. Droser and S. C. Finney, eds.), *SEPM Pac. Sect. Book* **77**:257-260.

Shaw, A. B., 1964, Time in Stratigraphy, McGraw-Hill, New York.

Ward, P. D., and W. J. Kennedy, 1993, Maastrichtian ammonites from the Biscay region (France, Spain), *Paleont. Soc. Mem.* **34**:58 p.

Wood, H. E., Chaney R. W., Clark, J., Colbert E. H., Jepsen, G. L., Reeside, J. B., and Stock, C., 1941, Nomenclature and correlation of the North American continental Tertiary, *Geol. Soc. Am. Bull.* **52**:1-48.