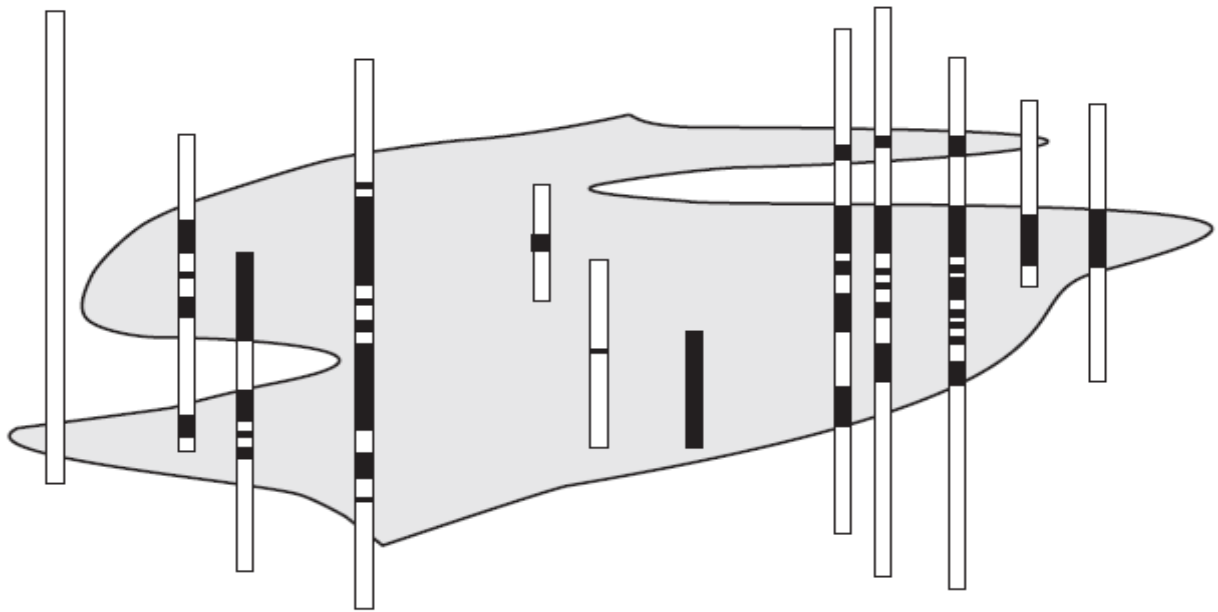


# Biochronology as a Traveling Salesman Problem

## Introduction to the Conop9 Seriation Programs



**PETER M. SADLER**  
Department of Earth Sciences,  
University of California, Riverside CA 92507



**PETER M. SADLER**

Department of Earth Sciences,  
University of California, Riverside CA 92507

# **Biochronology As A Traveling Salesman Problem: Introduction To The Conop9 Seriation Programs**

## **CONTENTS**

0. What is it?
- I. Biostratigraphy and Biozones
- II. Chronostratigraphy and Time Lines
- III. Graphic Correlation
- IV. Constrained Optimization
- V. The Traveling Salesman's Solution
- VI. Inside Composite Time Lines
- VII. Late Cretaceous of Antarctica
- VIII. Silurian of Baltica
- IX. Classic Riley Formation Data
- X. Mohawkian Facies Problems
- XI. Liassic Ammonites of England
- XII. Carboniferous of the Harz Mountains
- XIII. References
- XIV. Acknowledgements
- XV. Worksheets to Print

## ORGANIZATION

Sections I through VI briefly review the history and theory of various aspects of biochronology. The sections enjoy some overlap and may be read separately.

Sections VII through XII introduce case histories to illustrate the use of the seriation program, CONOP9, and the custom data manager CONMAN9.

The CONOP software is compiled for Windows and delivered with three additional PDF documents: a “*Reference Manual*” and “*Users’ Guide*” to CONOP9 and an introduction to CONMAN9 titled “*Starting.*”

## PROGRAM DOWNLOADS:

The CONOP and CONMAN programs, together with manuals, data sets, and classroom exercises, can be downloaded from the Spring 2009 GEO 206B website at UC Riverside [<http://ilearn.ucr.edu>] by clicking the following buttons and tabs:

[Guest Login](#) > [Courses](#) > [Course Catalog](#) > [Geosciences\(GEO\)](#)  
> [Spring 2009\(1\)](#) > [GEO 206B 001 09S](#)

2nd PDF Edition, October 2010  
University of California, Riverside California 92521, USA  
The second edition corrects figures IV4, IV5, and IV6

Prepared for a course sponsored by  
The Chinese Academy of Science  
39 East Beijing Road,  
Nanjing 210008

# 0. What is it?

## What is it?

### What is it like?

In *Save the Cat*, Blake Snyder (2005) asks these questions of every screen-play. Here are answers in his terms:

**Title:** *Brute-Force Biochronology*

**Log-Line:** *Search-algorithms meet field-data and together sequence the history of life.*

**Genre:** *Buddy Movie*

**Species:** *Town Mouse and Country Mouse*

**Story Arc:** *Cautious distrust turns to mutual respect as algorithms and data join forces to perform quality control and interpretation, changing the practice of biochronology.*

## What is biochronology?

Biochronology pieces together time-lines of events in Earth history, mostly the originations and extinctions of ancient species, but also volcanic ash falls, geomagnetic field reversals and excursions in ocean composition. Time lines help fictional detectives solve murders and help paleobiologists discover the scope and circumstances of extinctions, chart the course of evolution, compile the impacts of environmental change, refine the geologic time scale, and more. Maybe it belongs in the “superhero” genre?

## What is Brute Force?

In this context, brute force means imagining millions of working hypotheses and testing them all against the data to see which is fittest – a job for a fast computing machine that slavishly follows simple rules through tedious repetitions.

The human expert might use one rich stratigraphic section as the first approximation of a paleobiologic time-line and modify it progressively

as new information is discovered at other localities. If this considered to be forward progress, brute force can be said to invert the process; it starts with a crude guess at a possible time-line of all events and modifies the sequence by random mutation until the fit with all the field data cannot be improved any more.

The number of mutations far exceeds the number of events in the time-line. There is an obvious analogy with genetic mutation and survival of the fittest..

Recourse to this type of computer-assisted trial-and-error approach becomes *worthwhile* in biochronology after a critical volume of data has been compiled. It becomes *necessary* when that volume overwhelms the human “neck-top” computer.

## How is this a Traveling Salesman Problem?

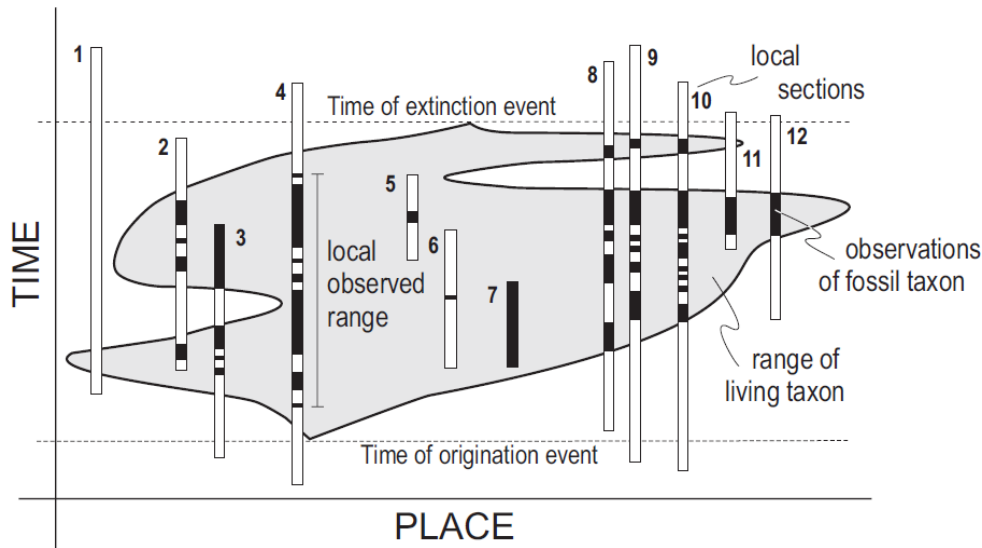
Not because the brute-force method needs to be pitched, door-to-door, like a screenplay! The Traveling Salesman Problem has become one of the favorite challenges for computer scientists. They fondly call it the TSP; it is deceptively easy to state and tedious to solve. Many practical problems are so closely analogous to the TSP that they can be solved by adapting the same methods. The paleobiological time-line task has been shown to be one of the analogous problems.

## Is the Book as Good as the Movie?

No! The colloquial question-and-answer style of this page works well from the podium in a short course setting, especially with software that provides run-time animations. Regrettably, the rest of this accompanying book is much denser reading. It works with still images. It must frequently tell rather than show. And it needs to fill-in rigorous definitional details and provide answers to questions that not all readers would care to ask.

# I. Biostratigraphy and Biozones

Summarizes the origins, evolution, and methods of biostratigraphy.



**Figure I.1** Range of a taxon in life (gray), the observed range of the fossil taxon (black) in stratigraphic sections (rectangles), and the time lines defined by the origination and extinction of the taxon.

## I.1 A Brief History

Biostratigraphers study the distribution of fossils in sedimentary strata. They have two motives – reconstructing the history of life and developing a relative time scale for other geologic studies.

More than two hundred years ago, before formulation of the theory of evolution (Darwin 1859), it became apparent that the same general succession of faunas could be recognized in different rocks at widely separated locations. Trilobites appeared before ammonites, for example, and dinosaurs became abundant before mammals. Such observations led to the major divisions of the Phanerozoic time scale – the Paleozoic, Mesozoic, and Cenozoic eras – and to attempts to resolve much finer subdivisions using fossil species. These subdivisions enable time correlation – the

identification of strata in different places that were deposited during the same time interval.

William Smith (1799) simply traced strata across England using fossil content to augment similarities of rock type and topographic expression (Lauden, 1976). Throughout the Paris Basin, Cuvier and Brongniart (1808) mapped a suite of coarse units on the basis of different faunal assemblages which they believed to be separated by catastrophic events. Phillips (1829) established that some fossils could be used to correlate between different rock types and Murchison (in Sedgwick and Murchison, 1939) ventured to insert newly found faunas from southwest England into a gap in the faunal succession established elsewhere. Adopting bold reasoning from Lonsdale, they used the intermediate nature of the new faunas between those previously known, rather than direct observation of an occurrence in intervening strata.

D’Orbigny (1851) and Opper (1858) began the practice of correlating strata using units of shorter duration (stages and zones) defined by the ranges of fossil assemblages and individual taxa. These and other pioneers established practices of qualitative biochronology that were followed for a hundred years without recourse to numerical methods.

The resolving power of correlation improved significantly when petroleum companies began to apply sequences of microfossil species to the task. A century after Spencer (1864) had embraced natural selection as a means to understand both biological and economic change, petroleum geologist Alan Shaw (1964) formulated “economy-of-fit” as a graphical means to reconcile differences in detail between the successions of species found in separate outcrops of the same age.

At Shell Oil Company, recognition of these differences was fueling concerns about the validity of fossils for correlation (Shaw, 1995); some preferred to proceed as if rock units were of uniform age and only the timing of appearance and disappearance of fossil species varied from place to place (diachronism). Of course, the boundaries of rock units and species ranges can both be diachronous. Shaw’s correlation graphs showed clearly that the problem of separating these two effects and determining the true sequence of all taxa at hand is underdetermined – the unknowns outnumber the constraints. Nonetheless, the practical value of a good approximation had led Shaw to quantify a biostratigraphic implementation of the general scientific practice of seeking simplest solutions (parsimony or Occam’s Razor).

Shaw’s method found favor at Amoco as the petroleum industry exploited the rich microfossil

record to correlate between exploration wells, adding rapidly to the volume of biostratigraphic information, and turning to computers to store and analyze it all (Carney and Pierce, 1995; McGowran, 2005). The Deep Sea Drilling Project led to internationally standardized, high-resolution subdivisions based on microfossils extracted from cores of the ocean floors.

At the same time, computer scientists found in natural selection a trial-and-error process that could be mimicked by algorithms to solve difficult optimization problems. By the 1980’s paleontologists had experimented with a wide range of numerical approaches to biochronology (Tipper, 1988) and evolutionary heuristics were flourishing as a means to compute very good answers to complex scheduling and engineering problems. Personal computers made it practical for individual paleontologists to exploit all these developments and build detailed paleobiological time-lines of faunal assemblages (Guex, 1991) and species first and last appearances (Alroy, 1992).

The basic practices of correlation also adapted to the increasing availability of radioisotopic dates and the calibration of global events such as geomagnetic reversals and excursions in the isotopic composition of oceanic water events. Particularly for Cenozoic periods, this changed the focus from biozone definition to age-calibration of species appearances and disappearances which could then be used as biohorizons for indirect dating.

The fundamental challenge for all these approaches to biostratigraphy as a means of dating remains the discrepancy between biostratigraphic horizons and time horizons (Fig. I.1). To step from biostratigraphic description to time correlation, it is necessary to compensate for the ecological and

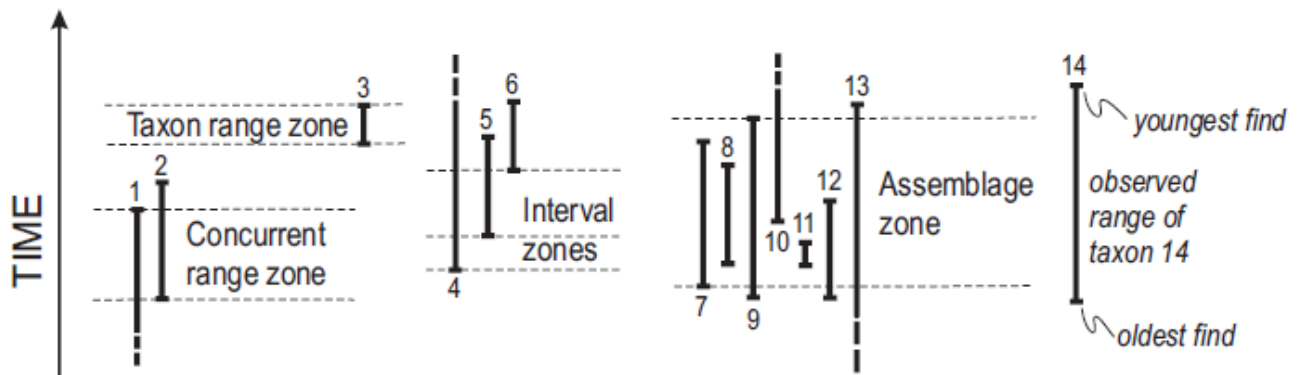
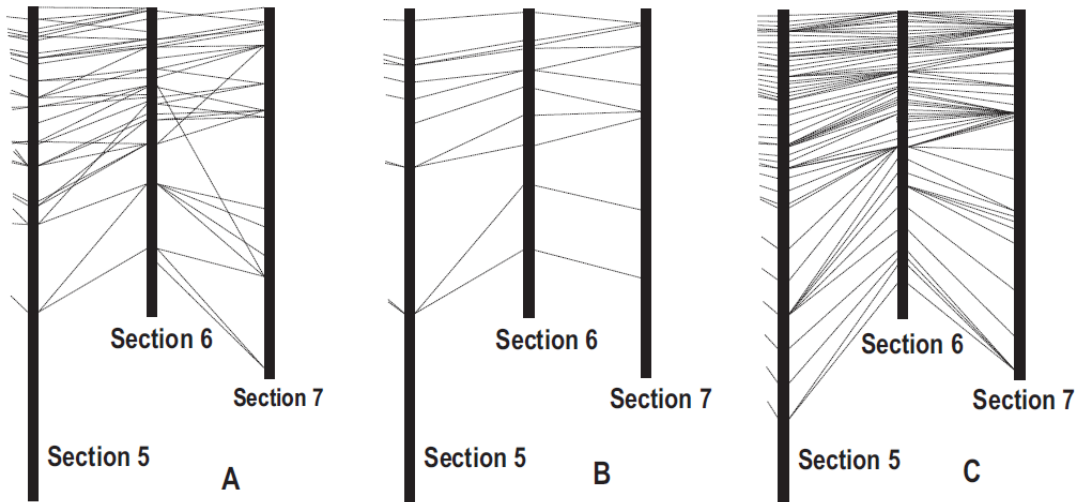


Figure I.2 Biozones defined by subdivision of a range chart.



**Figure I.3** Biostratigraphic correlation as a fence diagram: part of a 7-section project. A) Literal correlation of observed range ends. B) Range ends culled to those found in the same sequence in all 7 sections. C) Adjustment of all range ends to the sequence that best fits all local observations.

stochastic factors which inevitably cause the time of appearance and disappearance of fossil species to vary from place to place; i.e., some local discrepancies which are valid biogeographic information and some which are a nuisance resulting from the vagaries of preservation and collection. In spite of some contention about the best way to handle this problem, it is evident that, in practice, there has been no more cost effective, more generally successful, or more readily available means than fossils to correlate rock strata (Ludvigsen et al., 1986).

## I.2 Terms and Basic Data

A biozone is a body of rock characterized by the fossils it contains (Salvador, 1994). The bounding surfaces of such a unit are almost certainly diachronous; that is, the age of the highest and lowest finds of the definitive species usually vary from place to place. The reasons are simple and unavoidable: new species evolve at unique places and times; their ranges expand to wider but patchy distributions that change over time; the geographic range dwindles more or less rapidly as the species is replaced by others; and incomplete preservation leaves a fossil record of the species that is patchier still.

Furthermore, even deep marine sequences of sedimentary strata do not accumulate continuously

(Aubry, 1995); they include gaps at all scales and some observed ranges will end artificially at such gaps.

Portions of the bounding surface of a biozone are biohorizons – stratigraphic surfaces characterized by a faunal change, such as the lowest or highest local observations of a fossil taxon. The observed biozone will be smaller than the range of the living species because of failures of fossil preservation and collection. The corresponding time interval (chronozone) would extend beyond the living range to encompass all the rocks deposited during the life span of the species, whether or not they contain the fossil. In spite of their essentially diachronous nature, biozones and biohorizons are routinely used to estimate the position of datum surfaces that correspond everywhere to the time of origination or extinction of a taxon.

The raw data of biostratigraphy are faunal lists – inventories of fossil taxa that coexist in the same stratum. Ideally, numerous faunas can be placed in sequence at individual stratigraphic sections, permitting local range charts to be constructed that show the span of strata from the lowest to the highest find of each taxon (Fig. I.2). The range ends, highest and lowest local finds of a taxon, are local estimates of the first and last appearance datums (FADs and LADs).

The diachronous nature of biohorizons becomes evident when lines connecting the horizons of



numerous highest and lowest finds are drawn between stratigraphic sections to form a fence diagram (Fig. 3). Typically many of the correlation lines cross one another and it is clearly not possible to consider them all to be time lines.

### **I.3 FADs and LADs (biohorizons)**

Rather than using a few biohorizons to define zones, modern biostratigraphy seeks to retain as many biohorizons as possible and determine the sequence of the corresponding datums. An essential difference is that this approach must evaluate and adjust the local biohorizons (Fig. 3C). The elimination of cross-cutting biohorizons (Fig. 3A-B) removes the most obvious appearance of diachronism but does not guarantee that the remaining correlation lines are isochronous.

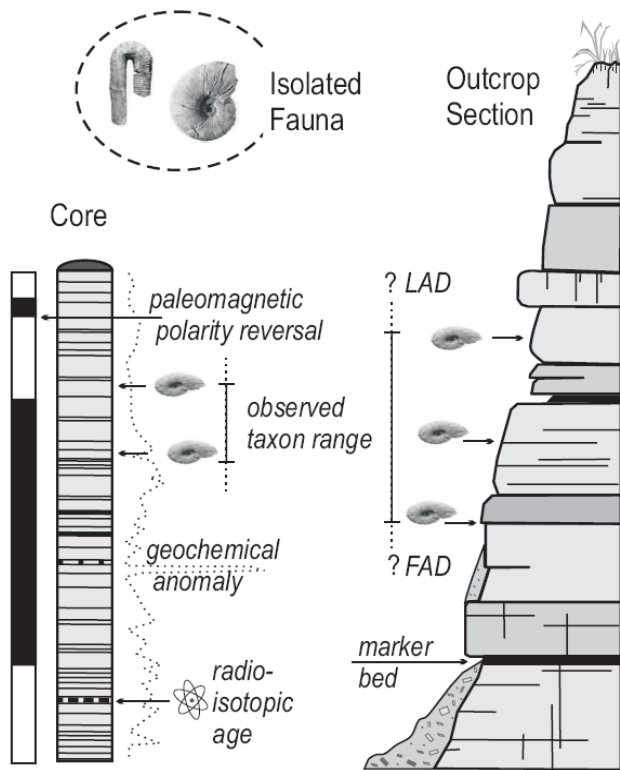
Three means of correction for diachronism are noteworthy. The first uses the gaps within the preserved ranges at one place to add confidence intervals to the observed range ends (Marshall, 1990); it corrects only the stochastic component of the difference between biostratigraphy and time. The second method uses radioisotopic dates and calibrated paleomagnetic reversals to determine the age of range ends at different locations. It is best developed for Neogene FADs and LADs in cores from the ocean floor. The third strategy uses the observed sequences and/or coexistences in many locations and the principle of parsimony: it searches for a composite sequence of events to which all the local range charts can be fit with the minimum of range adjustments. Assuming that most observed ranges are shorter than the true range, observed ranges are extended as necessary to fit them to a single global sequence. Initially, the method was implemented graphically as a form of regression

(Shaw, 1960). Now, computer programs with optimizing algorithms can search for the best-fit sequence in much larger data sets (Sadler, 2004). If available, horizons of known age, such as dated ash falls and calibrated paleomagnetic reversals, may be included to constrain the search for the best sequence.

Sometimes, especially over short distances, it is possible to exploit ecologic factors to achieve finer correlation. Short-lived changes in global temperature or the pattern of winds and ocean currents may become correlatable events when they cause changes in the relative abundances of pollen or foraminiferan taxa, for example. Distinctive fluctuations in stratigraphic trends of the relative abundances of taxa may then indicate correlative strata in different locations.

### **I.4 Resolving power**

In favorable situations, zonation based upon pelagic marine organisms can resolve intervals that are on average less than one million years in duration. Usually such sets of zones are based upon a single biological clade. Conventional ammonite zones and subzones in the Mesozoic resolve 0.4 to 0.75 million year intervals, on average. Mesozoic zones based on foraminifera and calcareous nannofossils resolve 2 to 3 million years. Cenozoic zones and subzones for the same microfossils have average durations of 0.75 to 1.0 million years. Often it is necessary to establish separate sets of zones for different climate belts. For Cenozoic radiolarians there are 32 low latitude zones and 11 mid latitude zones, for example. Ordovician conodonts support “North Atlantic” and “Mid continent” zones with resolving power on the order of 2 to 4 million years. For the Ordovician and Silurian, finer resolution (1 to 3 million years) is afforded by the more cosmopolitan graptolite zones.



**Figure I.4** Summary of the three common sources of information about taxon ranges and coexistences, with examples of other stratigraphic data types suited to correlation and inclusion in time lines.

Time lines that incorporate numerous biohorizons offer the promise of better resolving power. For the Ordovician and Silurian, over 3000 graptolite appearance and extinction events have been sequenced. For the Cenozoic there are on the order of 100 calibrated biohorizons each for foraminifera and calcareous nannofossils. Average resolving power on the order of 0.25 million years or better would appear to be possible by using biohorizons from more than one fossil clade. Whether this is attempted by recourse to radioisotopic calibration or computer optimization of observed sequences, however, there are always clusters of range-end biohorizons whose relative age remains irresolvable. Nevertheless, the resolvable events and clusters of events are often sufficient to increase the resolving power 5- to 10-fold over traditional biozones.

It appears that biostratigraphically based time scales might soon resolve intervals as short as 500,000 years and better through most of the Phanerozoic. This approaches the size of analytical errors quoted for

radioisotopic dates on Paleozoic ash falls. It is likely shorter than the time taken by many taxa to spread from their point of origin to their full geographic extent – an inevitable source of discrepancy between biohorizons and time lines. These figures concern the resolving power of time scales for global correlation. As the distance of correlation is reduced, more biohorizons and strata approximate time lines, and potential resolution improves. The actual resolving power of any given correlation will be limited by the fossil content of the strata.

### I.5 Too Much Information

Too much information can complicate otherwise simple procedures. A growing volume of relevant facts and insight, which should support more precise analysis, can not only overwhelm expert qualitative analysts, but also expose flaws in simplifying assumptions. The fossil record reached this stage as a geologic timekeeper. The use of fossil successions to correlate sedimentary strata began straight-forwardly while the volume of relevant paleontologic information was still meager. Paleontologists now assemble information in large, open-access databases that support the pursuit of long and finely resolved, global time-lines of species first- and last-appearance events.

Unfortunately, the sheer volume of available and relevant information can overwhelm traditional qualitative techniques. Range charts of first- and last-occurrences of fossil species must be combined from many locations to compensate for local incompleteness of the fossil record. Enlarging the geographic scope adds the complications of faunal migration and provinciality – for which the remedy is yet more information. New locations lead to the discovery of new taxa and the time-line task tends to remain underdetermined.

Consider the information volume behind the task of building a time-line for just one “short-lived” extinct clade, the graptolite clade, which supports traditional biostratigraphy of the Ordovician and Silurian. For this exercise, I have compiled a database of 506 sections with 2090 taxa. That provides 19,924 locally observed events and 289,326 pairwise constraints (28,200 proven pairwise coexistence constraints and 261,126 proven instances of a first-appearance that must precede a last-appearance. Alroy (1992, 1994) compiled a larger number of taxa and localities for fossil land mammals of North America!

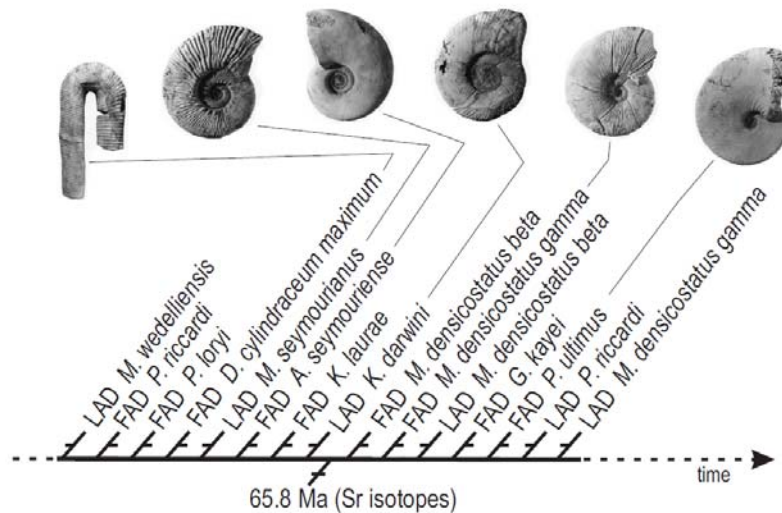
Although such volumes of data surely overwhelm the human “neck-top” computer, there still remain too many unknowns to resolve a unique correct time-line. The remedy lies not only in the application of computers to biochronology but in the acceptance that

uncertainty intervals need to be applied to the position of first- and last-appearance events in the resulting time-lines.



## II. Chronostratigraphy and Time-lines

Reviews paleobiologic time-lines as a foundation for time stratigraphy, macroevolution, and paleobiogeography.



**Figure II.1** Part of an ordinal (unscaled) time line with ammonite range-end events and dated events, built from eight Campanian-Maastrichtian stratigraphic sections on Seymour Island, Antarctica (data from Macellari, 1984, 1986). Time runs from left to right; events are ordered but their spacing is arbitrary. FAD and LAD are abbreviations for first- and last-appearance datum – the global origination and extinction times of fossil taxa – time-stratigraphic horizons, as distinguished from biostratigraphic horizons recognized by local oldest and youngest finds of the fossils.

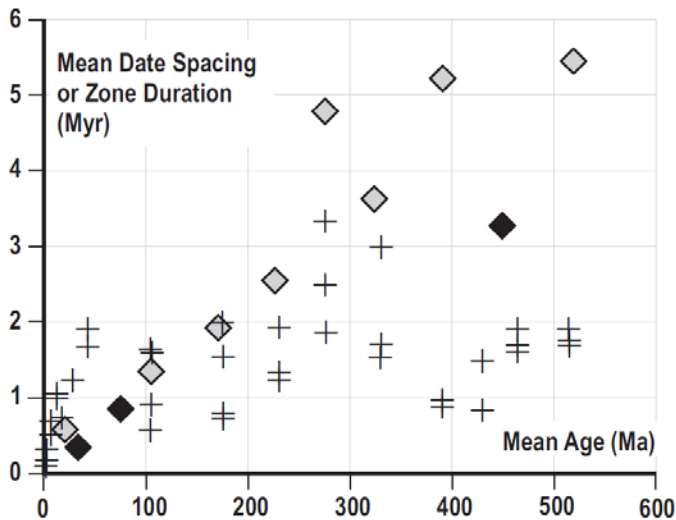
### II.1 Time Lines and Time Scales

Our knowledge of geologic history can be expressed as a time line; i.e., a set of past events placed in the order of their occurrence (Fig. II.1). Every stratigraphic section contributes to the time line, because the local order of events is indicated by the sequence of superposed strata in which they have left a record (Fig. I.4). Isolated fossil faunas and floras contribute too; they provide evidence of taxa whose ranges overlap in time and place.

Typical geologic time lines include large numbers of origination and extinction events of fossil species. These time-stratigraphic first- and last-appearance events (FADs and LADs) usually outnumber other events of time-stratigraphic significance, such as paleomagnetic reversals, volcanic ash falls, and isotopic excursions. They surely outnumber the

radioisotopic dates available to calibrate the ages of events on the time line and to scale the time intervals between them. Fortunately, the sequence of the many biostratigraphic events can be established by stratigraphic superposition, without recourse to radioisotopic dates. Also, the relative spacing of events in time can be estimated, albeit rather crudely, from the average rock thicknesses between the species range ends as observed in local stratigraphic sections. If dated events, such as volcanic ash falls, are included in the sequencing exercise, their spacing serve as a test of the success of the estimates of the spacing of events.

Traditional biostratigraphic time scales are simplified versions of time lines. They group events into zones and show only the events that define zone boundaries. For much of Mesozoic and Paleozoic time, the average intervals between biostratigraphic zone boundaries are shorter than the average intervals



**Figure II.2** – Average time intervals between radioisotopic dates (diamonds) and biostratigraphic zone boundaries (crosses) used in the Gradstein et al. (2004) geologic time scale. Open diamonds: radioisotopic dates used for calibration; filled diamonds: radioisotopic dates for the mammal, ammonite and graptolite examples enumerated in the text; crosses: average durations of biostratigraphic zones in each geologic period for the clades with highest resolving power depicted in Gradstein et al. (2004).

between reliably dated events. The practical resolving power of modern geochronometry is limited by the small number of dated events, not by the analytical precision. Uncertainty intervals on the most precise radioisotopic dates are shorter than the duration of many Paleozoic biozones. Biostratigraphic time-lines offer a better chance than biozones for keeping pace with geochronometric precision.

The resolving power of an ordered time-line and its value for correlation increases with the number of events on the line. The resolving power of the calibrated, scaled time-line increases with the number and analytic precision of the dated events. The difficulty of estimating the true sequence of the first- and last-appearance events increases with their number, but can be mitigated by adding more dated events. Thus, biostratigraphy and geochronometry complement one another.

In practice, there will remain clusters of biostratigraphic events whose relative positions on the time-line cannot be resolved. They must either share one position, or each be drawn with an uncertainty

interval that overlaps the other. The irresolvable clusters are typically much smaller than the number of events traditionally grouped into biozones. The use of uncertainty intervals provides a rigor comparable with the error bars on radioisotopic age determinations

While radioisotopic dating laboratories explore new procedures to reduce the analytical uncertainties in numerical age determinations, biostratigraphers develop new techniques for sequencing large numbers of stratigraphic events. The most recent of these stratigraphic sequencing methods take advantage of computer memories and optimization algorithms. They manage large data sets and increase the speed and reproducibility of the sequencing process.

The sequencing algorithms implement simple logical rules which, nevertheless, need to be expressed in precise terms. Do not let the formality of the logical statements obscure the underlying simplicity. In order to understand the need for computer algorithms, we should review two geologic problems that are more fundamental than the sheer volume of data.

First, is the “unsteady process problem.” It causes the stratigraphic record to contain hiatuses at all scales and means that the long term accumulation rate of the record is a very poor guide to the short term relationship of rock thickness and elapsed time. In practice we must be wary of estimating age by linear interpolation and we need to construct composite sections so that the hiatuses from one locality can be filled with the rock record from others.

The task of correlating many sections into a single composite time-line runs into the “biostratigraphic sequencing problem.” The observed sequence of range ends at any one location is not a reliable indicator of the true sequence and FADs and LADs. The computer algorithms attempt to solve this problem when building composite time lines. Let us consider these two problems in a little more detail.

## II.2 The Unsteady Process Problem

The age of a Holocene or Pleistocene fossil might be estimated directly by a single radiocarbon date. For the older fossil record, neither the fossils nor their host rocks are likely to be suitable for dating. Datable layers, such as volcanic ash-falls, are more sparsely and irregularly distributed than fossils. Consequently, paleobiologists typically need two radioisotopic dates to bracket the age of a fossil. Between two dated ash fall tuffs, it is possible to estimate the age of other layers by

interpolation and to calculate average rates of change across the dated interval. Because stratigraphic sections may contain hiatuses at all scales (Sadler, 1999), however, the interpolation introduces an additional error. Linear interpolation assumes, incorrectly, that sediment accumulates steadily; i.e., that the rate does not change with time and no hiatuses intervene. In fact the time between two dated strata is likely to be non-uniformly apportioned into the intervening rock layers and hiatuses. To minimize the uncertainty introduced by linear interpolation it is desirable that the interval between the dates be as brief as possible.

Unsteadiness characterizes not only the net accumulation rate of the sedimentary stratigraphic record (Reineck, 1960; Sadler, 1981), but also rates of organic evolution (Gingerich, 1983, 2001; Reznick et al., 1997), and rates of change in the physical environment (Gardner et al., 1987). Rapid rates of change characterize short term pulses of activity but are not sustained for long periods. This is evident even at the short time scales of our individual human experiences: consider the changing level of sand and sea at the beach, for example, or the progress of hillside erosion by debris flows. Longer term fluctuations are evident at historic and paleontologic time scales.

Average rates, as measured between two dated horizons, tend to decrease systematically as we consider dates that are farther apart and the rates are averaged over longer and longer time spans. The strong dependence of mean rate on averaging time is shown by empirical power laws in which the logarithm of rate decreases steeply and nearly linearly, as the logarithm of averaging time increases (Sadler, 1993) up to 1,000 – 10,000 yr intervals, then less steeply beyond 100,000 yr intervals. Notice that the rollover in behavior of rates tends to separate the time scales of stratigraphy from sedimentology, paleontology from biology. Only the combination of insights from these disciplines reveals the full picture.

Short term pulses of rapid change have obvious consequence for human activities. They also appear to be a influential agents in Earth history; e.g., tsunamis and hurricanes for coastal change, “punctuations” and mass extinctions in evolution, and short term climate pulses for ecological change. Short term extremes are most easily assessed for the recent past, where radioisotopic dates are relatively plentiful and dating techniques achieve the smallest uncertainties in age. In order to estimate the most intense plausible events and the full scope of Earth processes, however, it is

desirable to extend the sampled interval to older parts of geologic history. Our practical ability to calibrate abrupt changes in the distant past is more often limited by the interval between successive dates in stratigraphic sequence, than by the precision of the individual dates; this is because interpolation between widely spaced dates loses the hard-won precision in the dates themselves.

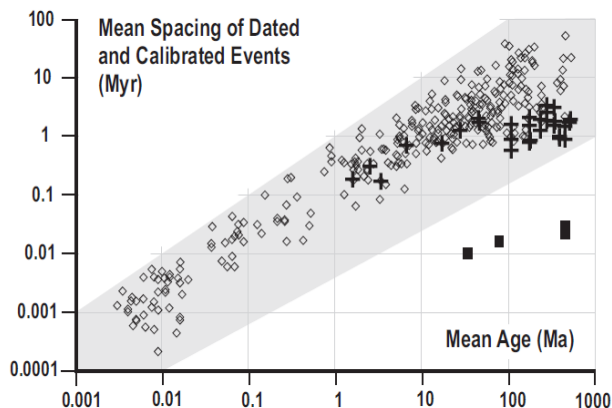
### II.3 The Need for Composite Sections

In order to measure the unsteadiness of organic evolution and calibrate abrupt environmental changes, it is evidently desirable to have closely-spaced dated events in the same stratigraphic section. Ideally, the spacing between dated events should approach the analytical uncertainty of the dates themselves. For practical stratigraphic problems, of course, this ideal condition is far from realized, especially in Mesozoic and Paleozoic strata. Even global compilations for whole periods find few radioisotopic dates with high precision and robust stratigraphic context (e.g., Gradstein et al., 2004; summarized in Fig. II.2).

Consider the graptolite and conodont clades of the Ordovician and Silurian periods. The world-wide fossil record of this 70 to 80 million year interval has produced at least 500 published stratigraphic sections with range charts for more than 2000 graptolite species and another 300 or more sections with ranges for over 1000 conodont species. That is a total of about 6000 different species’ first- and last-appearance events, recorded by more than 28,00 locally observed range ends (Sadler and Cooper, 2003; Cooper and Sadler, 2004; Melchin et al., 2004; Sadler unpublished compilation). For the same time interval, only 20 to 25 reliable radioisotopic dates can be linked by association with a conodont or graptolite fauna. Of these, less than ten are in sections that have yielded a rich, well-constrained set of graptolite or conodont ranges and only one section contains more than one dated event. On average, more than 3 million years separate the useful dated events. The spacing is very uneven. At best, there are four dated bentonites within a 5 million year interval that can be summarized in a composite stratigraphic section.

The situation improves for younger clades. Campanian and Maastrichtian ammonites have been reported from at least 193 individual sections or beds, world-wide, which have yielded 562 species; that is 1124 events recorded as 3476 observed range ends





**Figure II.3** – Mean time intervals between dated and calibrated events (open diamonds) in examples of the most intensively calibrated single stratigraphic sections (calcareous ooze and carbonate platform databases; Sadler, 1993, 1999), compared with the resolving power of biostratigraphic zones (crosses). The gray zone encompasses the uncertainty intervals ( $\pm 2$  sigma) of typical published radioisotopic dates in stratigraphic studies. The black squares show the potential resolving power of long sequences of biostratigraphic events (mammals, ammonites, and graptolites are the examples used).

(Sadler, unpubl. compilation). For this 18 million year interval, 21 dated events can be associated directly with an ammonite fauna. For North American Cenozoic land mammals, Alroy (1992, 1994, 2000) reports 3248 species from 4978 collections associated with 186 dated events.

Figure II.2 plots these three examples together with the average spacing of those radioisotopic dates considered suitable for calibrating other periods in a recently compiled geologic time scale (Gradstein et al., 2004). The data of figure II.2 are global compilations, not single sections. Figure II.3 illustrates the average spacing of dated and calibrated events that has been achieved for examples of the very best calibrated single stratigraphic sections – those in fossiliferous shallow marine limestones. In addition to directly dated beds, these sections include the calibrated boundaries of units defined by biostratigraphy, stable isotopes, and paleomagnetic polarity. For Mesozoic and Paleozoic time, the average spacing of these control points for the calibration of single sections approaches neither the duration of biostratigraphic zones (Fig. 2), nor the best analytical uncertainty of the dating methods (Fig. II. 3).

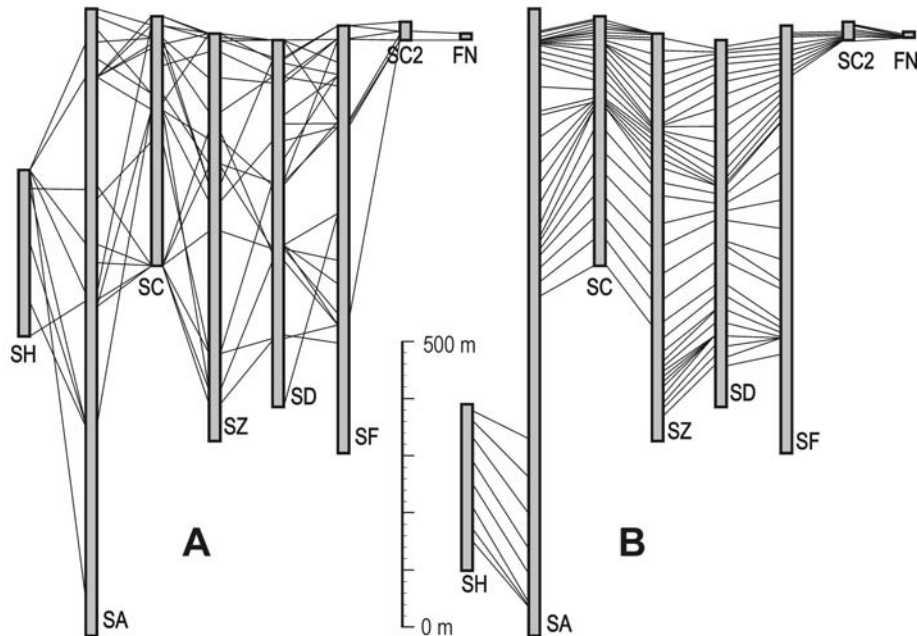
To preserve the full potential of the dated events shown in figure II.2, it is necessary to correlate all the local stratigraphic information into a global composite section. If the correlation proceeds by traditional biostratigraphic practice, the result will be a time scale with units whose resolving power is limited to the duration of standard biozones. This is better than the current average spacing of dated events (Fig. 2), but coarser than the best analytical uncertainty. The potential resolving power of biostratigraphy can be better matched to the most precise geochronometry by attempting to sequence all the range ends of all available species into a single composite time-line.

For this sequencing task, the advantage of radioisotopically dated events is that their relative ages may be evident without superposition in a single section. If the uncertainty intervals of two radioisotopic dates from different localities do not overlap, the order of the dated events is not in doubt and may constrain the correlation and sequencing processes. The composite time line can bring more dated events into a single sequence than are observed in a single measured section, thus reducing the intervals over which interpolation must be attempted. It also deals with the nuisance that the locations of an observed range end in different local stratigraphic sections are unlikely to correspond exactly in age with one another or the age of the true evolutionary origination and extinction events – the FAD and LAD. This discrepancy is the essence of the biostratigraphic sequencing problem.

#### II.4 The Biostratigraphic Sequencing Problem

Biostratigraphic zones are a ploy to deal with the problem that, for many reasons, observed range ends are not everywhere of the same age. The distribution of a living species is likely to be dynamic and patchy; the limits of its habitat may have complex shape and shift with time. Add to this the effects of patchy preservation and the inefficient collection of fossils. It is to be expected that the observed local ranges of species will be shorter than the true global or regional ranges. These discrepancies may be partly biogeographic fact and partly a fiction introduced by preservation and collection; but both factors combine to make locally observed range ends less reliable for correlation and time scales.

For the same reasons, the sequence of observed range ends is likely to vary from section to section. This is a nuisance, but it is also the most compelling



**Figure II.4** – Fence diagrams correlating 8 stratigraphic sections on Seymour Island, Antarctic Peninsula, using ammonite and nautiloid species. A: direct connection of observed species range ends. B: the same data after parsimonious range extension of observed ranges to match all sections to same sequence and insertion of missing events. Optimal solution found by constrained optimization using CONOP software; resulting range end positions subject to 7-point smoothing to eliminate spurious concentrations at sample horizons (see text for more explanation).

evidence that local range ends cannot be treated directly as if they were FADs and LADs. Global range-end events (true FADs and LADs) that are closer in age are more likely to be observed in the wrong order than those whose true ages are farther apart. Thus, the long-term progression of evolutionary faunas is not in doubt, (trilobites surely appeared before dinosaurs, which appeared before hominids) even though the biostratigraphic sequencing problem is real. The problem arises and intensifies as we seek greater precision in correlation and the sequencing of species.

To expose the extent of the problem, simply attempt correlation by drawing a fence diagram that directly connects the observed range ends for each species (Fig. II.4A). The usual result is a tangle of crossed lines that cannot possibly all represent time correlations. In order to understand how the problem will be solved, consider how we might refute an unpalatable proposition that “all species originated before the Cambrian and none has ever gone extinct.” Given that true species ranges may be longer than their observed ranges, we cannot directly conclude that the

notion is absolutely impossible. Rather, we reject the idea because it implies such colossal, systematic failures of preservation and observation.

Similarly we can “solve” the sequencing problem by accepting that sequence of events which implies the least failures of preservation and observation. Before considering practical ways to find such a sequence, let us examine how traditional biozones solved the problem. It was a sensible way to simplify the task before computers were readily available.

#### II.4 Traditional Biozones as a Solution to the Sequencing Problem

In effect, the interval zones of traditional biostratigraphy cull correlation lines selectively from the tangled fence diagram (Fig. II.4) until the smaller number that remain do not cross one another; these remaining lines become the definitive boundary events for interval zones within which several potentially separable range end events are grouped into an



unresolved cluster. The culling process solves the problem by sacrificing potential resolving power. Unfortunately it does not correct the problem that none of the events, even those that survive the cull, is likely to be preserved at horizons of the same age everywhere. Sadler et al. (2009) have demonstrated that traditional biozones may be recognized in time lines that have been developed without recourse to zonation.

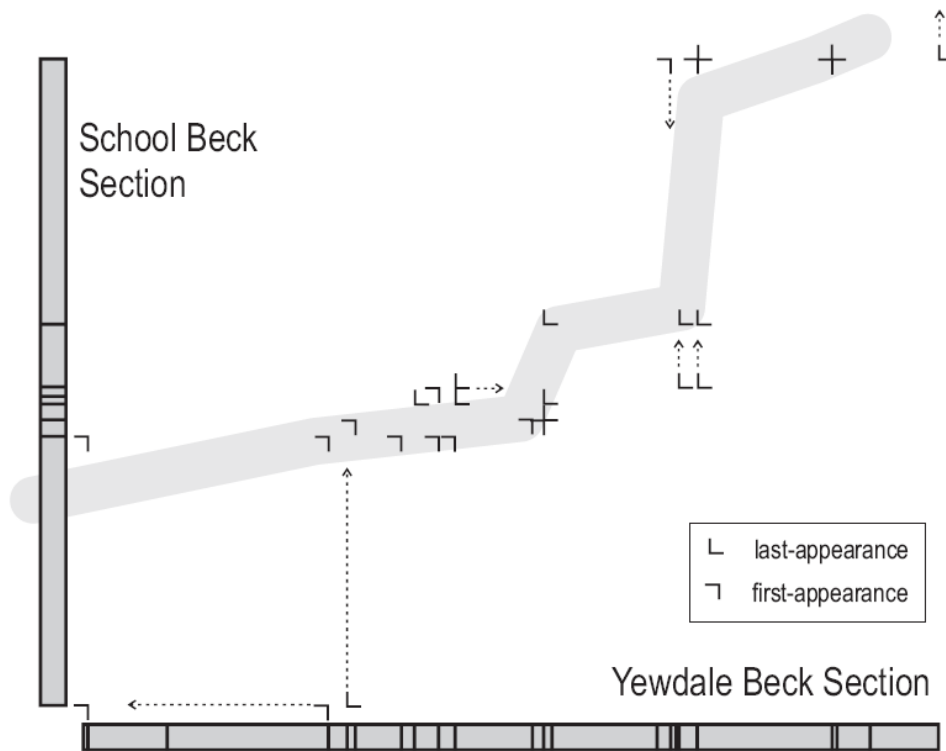
Alternatively, there are traditional assemblage zone boundaries. These are variously defined on the appearance of a number of characteristic taxa. Joint occurrences of a whole or partial association may be more likely to maintain the same order from place to place. It is also possible that, while individual first appearances may be diachronous, the first appearance of a subset of taxa will be more stable. In developing the method of "Unitary Association" Guex (1977, 1991) makes the case for computer assisted recognition of assemblage zones as another means to mitigate the problem of diachronism inherent in the boundary

definitions of interval zones. Both approaches sacrifice the resolving power of a time-line. Faced with any assemblage of taxa, the time-line identifies that interval in which they may be expected to co-occur. Named zones and defined boundaries are useful for communication but not necessary for correlation and relative dating, if a time-line is available.

The duration of biozones becomes the tacit, but often ignored, uncertainty interval for correlations built upon them. Perhaps the most unfortunate consequence of biozonation is that expert biostratigraphers often report the age of a faunas to non-experts as a biozone. Subsequent publications by the recipients of this derived information, perhaps full of vital isotopic or paleomagnetic data, combine this information with purported zone boundaries rather than range charts or faunal lists. As Sweet (2005) has complained, failure to publish the underlying raw data limits the opportunity to build fully informed and fully resolved time-lines.

### III. Graphic Correlation

Introduces Alan Shaw's graphs as a means to visualize the paleobiologic sequencing problem.



**Figure III.1** – Graphic cross-plot of two early Llandovery (Silurian) sections (data from Hutt, 1975) showing range-end coordinates of all graptolite species found in both sections. Thick gray line of correlation (LOC) is fit not only to taxa on this chart but also with guidance from a full graptolite data set of 506 sections and 2090 taxa. The first- and last-appearance event symbols are designed to indicate that they are merely the corners of larger boxes through which the line of correlation must pass (Kemple et al. 1995); thin dashed arrows show examples of implied range-end adjustments, which are always a continuation of one of the open ends of the event symbol and represent an extension of the observed range in one of the two sections.

#### III.1 Graphic Correlation: a Better Solution

Shaw's (1964) graphic correlation method introduced a better solution to the correlation problem: apply to the observed range ends the smallest net adjustments necessary to bring all measured sections and their range charts into agreement with a uniform sequence

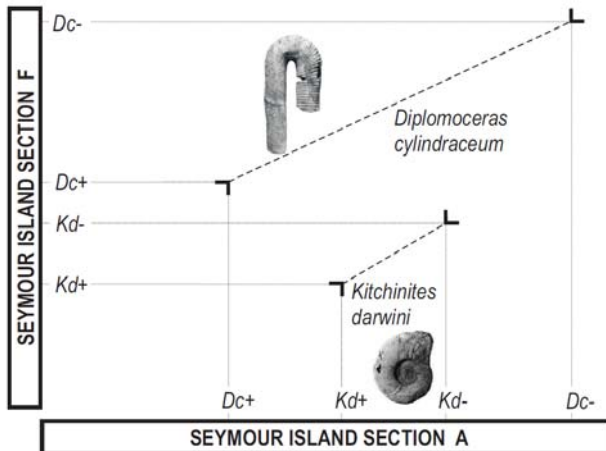
of events (Mann and Lane, 1995). Some events may occur in very different positions and with very different species associations from section to section; they will require such large adjustments, that they are clearly unsuitable for correlation. But the number of events requiring only small adjustments typically outnumbers the number of traditional zone boundaries. Graphic

correlation is a total-data approach that does deal with the real problem – that none of the observed range ends is likely to coincide with the true FAD or LAD.

When adjusting the observed range ends, graphic correlation assumes that observed ranges are most likely shorter than the true ranges, not longer (Edwards, 1982). It limits range adjustments to extensions. In effect, it searches for extreme events: the earliest of the local first appearances and the youngest of the last appearances. Other range ends are extended to match the extremes. Thus, graphic correlation is both truly time-stratigraphic and also susceptible to bias by erroneous taxonomy. Fortunately, its graphic nature facilitates the recognition of outliers and anomalies. It enables quality control and correlation at the same time.

Graphic correlation proceeds by rather straightforward manipulation of two-dimensional graphs. It is worthwhile to consider its simple rules because they can be transformed into algorithms that automate the sequencing of very large data sets. They may be illustrated with the simple 2-section, 2-taxon problem as shown in figure II.2.

The two sections are drawn up as the X- and Y-axes of a graph in which sections young away from the origin. Range-end events seen in both sections are plotted as (x,y) points, where x and y are the levels at which the event is observed in the two sections. The plotting symbols are chosen here to indicate that they are the corners of error boxes. Adjustments to their positions are permitted in the direction of the open ends of the symbols. First appearance coordinates,



**Figure II.2.** Graphic cross-plot of two ammonite taxon ranges from two of the stratigraphic sections on Seymour Island, Antarctica (Macellari, 1985).

for example, may be adjusted only downward or to the left; i.e., stratigraphically downward in one section or the other to extend an observed range.

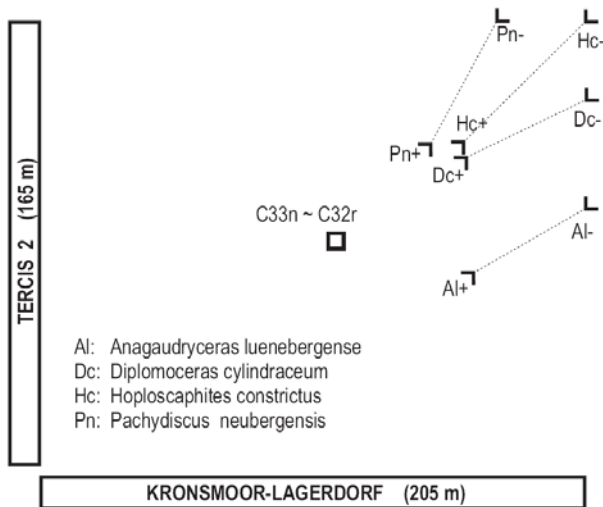
The purpose of adjusting the points is to bring them all onto a single line of correlation (LOC; MacLeod and Sadler, 1995). All points on the LOC on such a graph indicate a pair of horizons, one in each section, that were deposited at the same time. The LOC may be piecewise linear. Hiatuses introduce vertical and horizontal segments in the LOC – multiple horizons in one section collapse to a single horizon in the other – but no segment in the LOC may have a negative slope (unless part of one section is upside-down!).

As in figure II.2, these logical limitations on the LOC typically prevent it from passing through the coordinates of every point on the graph. The LOC is fitted to as many as possible of the most reliable range ends; others are adjusted onto the line. These adjustments represent range extensions in one of the sections. FAD events are adjusted downward. LAD events are adjusted upward only, unless reworking is suspected. In well-behaved instances of the correlation problem, most of the points on the graph form a narrow band that dictates the general position of the LOC. Obvious outliers identify those events that are least consistent in position and least reliable for time correlation.

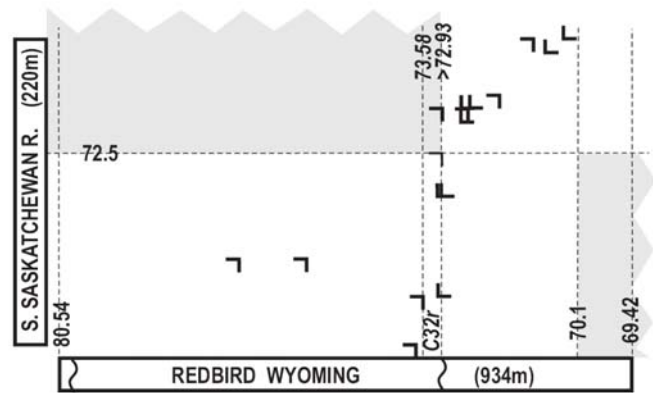
In figure II.2 LOCs passing through Dc+ can intercept only one more range end coordinate without a negatively sloping segment. Better LOCs pass through Kd+ and pick up two more range-end coordinates, leaving Dc+ to be adjusted downward in section F. Notice that Section A recorded a coexistence of the two taxa. Section F did not. We will see later that observed coexistences are powerful constraints on LOCs and time-lines. In this regard, the information in Section A is superior to that provided by Section F. The two sets of observations are compatible because observed ranges are assumed to underestimate true ranges.

### III.3 Incorporating Other Events

Time correlation may be aided considerably by events that allow no freedom to adjust their position in stratigraphic sections. Typical examples include marker beds, isotopic excursions and paleomagnetic reversals. They must truly occur where they are observed. The LOC must pass through the shared coordinates of these events. Ideally, all LOCs would be determined entirely by such events but, in practice, they are much less readily available than taxon range ends.



**Figure III.3.** Graphic correlation of two European late Cretaceous sections that share a paleomagnetic reversal (open box) and four ammonite taxa.



**Figure III.3.** Graphic correlation of the North American late Cretaceous Redbird and South Saskatchewan River sections, using ammonite range ends and dated events (four radio-isotopically dated ash fall tuffs and two calibrated paleomagnetic reversals.)

Figure III.3 shows how a magnetic reversal can constrain the placement of LOCs when the range end information is conflicted. Because the LOC must pass through the square symbol in the center of the graph, it cannot reach range end Al+. That would require a negative slope. The cluster of the other three first-appearance events (Pn+, Hc+ and Dc+) would have attracted the LOC away from Al+ anyway. This example encourages the view that majority of range end information support a good LOC.

To be useful, marker beds and paleomagnetic reversals must occur in both sections on the graph. Numerically dated or calibrated events are more powerful, in the sense that they know their relative ages independent of local superposition. Figure III.3 shows how dated events block the LOC from entering parts of the graph (gray in Fig. III.3). A pair of dated events, one in each section, divides the graph into four quadrants. One of the three quadrants will be out of bounds because its coordinates would imply an age older than the older date and younger than the younger date! With two pairs of dates, two blocked zones may generate a “gate” or “pass” through which the LOC must pass (Fig. III.3). The best LOC in figure III.3 will include a vertical segment that passes through the left hand limit of the gate. A vertical LOC segment indicates a hiatus in the x-axis section. A hiatus has indeed been identified in the Redbird section at this level.

### III.3 Shortcomings of Graphic Correlation

Kemple et al. (1995) have reviewed the strengths and weaknesses of graphic correlation. The main strength is the ability to see explicitly the consequences of every decision. All the weaknesses arise because the graphic procedure deals with only two sections at a time. Information from the second and each subsequently added section is projected in turn into the first section using the LOCs. Ranges observed in the first section are extended as needed and missing taxa can be projected into it. Thus, the first section is progressively corrected and augmented until it becomes a composite time-line with a local stratigraphic thickness scale.

As the number of taxa and sections increases, manual graphical correlation becomes increasingly unwieldy and time-consuming. With more taxa, the number of possible LOCs may increase rapidly; and the optimal LOC is harder to recognize; several LOCs may be equally good. The addition of each section requires a new graph and the outcome is very sensitive to the order in which sections are composited. To remove this bias, the compositing process is repeated through multiple rounds, to the point of diminishing returns (see Shaw 1964 for a detailed example). Some practitioners use only a single round of graphic correlation. This is a fine way to examine the quality of the available information and identify conflicts, but it stops short of becoming a rigorous correlation exercise.

For efficient compositing of global data sets, we resort to computers. Computer arrays are not limited to two dimensions; they allow all sections to be considered simultaneously. Thus, one form of computer-assisted sequencing may usefully be described as multidimensional graphic correlation, with one dimension for every section. There are clever computational strategies for searching among the huge number of possible LOCs by trial and error or “inversion.” These should not obscure the fact that the rules for choosing between LOCs follow the same simple stratigraphic logic of the kind that Shaw pioneered. The basic biostratigraphic rules set

out by Shaw can be augmented to include more constraints, especially those provided by geochronometry. They can also be modified to mimic different correlation strategies. For example, the size of range extensions can be measured by counting event-levels rather than rock thickness; this adopts the strategy of Edwards’ (1982) no-space graphs, to avoid any bias toward the thickest section. If the set of rules captures sufficient expert insight, these inversion techniques provide the best solutions to the biostratigraphic sequencing problem.

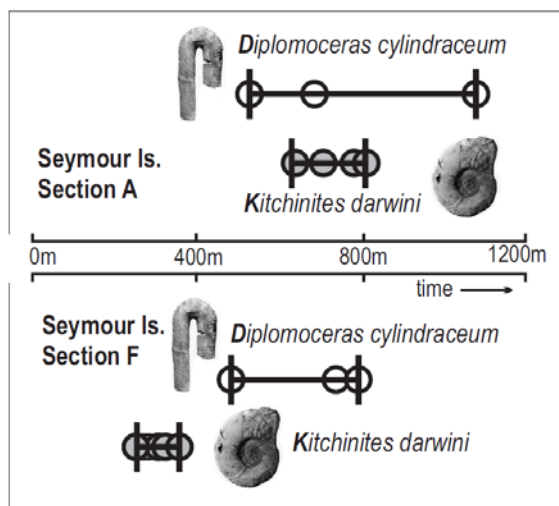
# IV. Constrained Optimization

Illustrates one option for extending graphic correlation to many dimensions.

## IV.1 The Simplest Case

To take advantage of computer speed and memory the graphic correlation problem is inverted. This means that, instead of working piecemeal from the field data to a solution, as in graphic correlation, possible answers are examined in turn for their fit to the data (Kemple et al. 1995). Computers have enough speed to solve problems by trial and error in this fashion. They exploit the “brute force” of time and reap the advantage that all the information is available all the time. It should not surprise paleontologists to learn that several of the heuristic inversion techniques are classified as “evolutionary programming.” Unintelligent mutation combined with fitness tests can solve very complex problems, given enough trials.

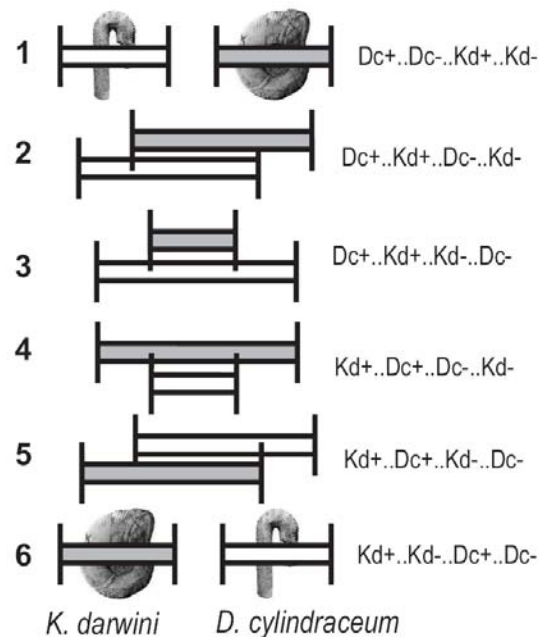
In order to appreciate the computer scientists’ approach to the sequencing problem, re-consider the simplest case from figure II.2 in the form of parallel range charts (Fig. IV.1) of the same two taxa and two sections. The data set is so small that the full set of possible event sequences can be examined in this format.



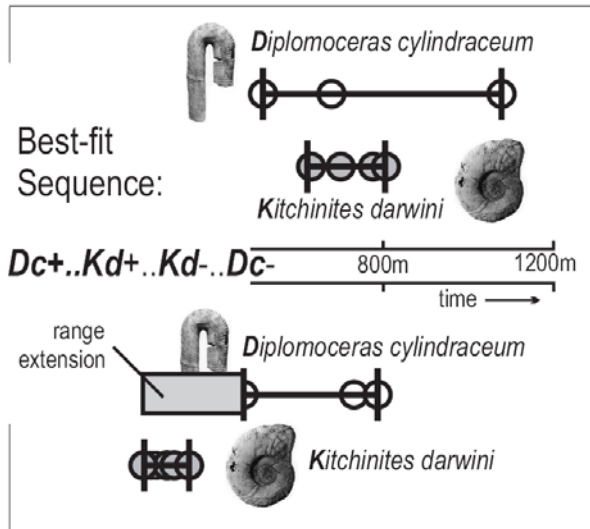
**Figure IV.1** Range charts for two shared ammonite taxa in two sections from Seymour Island, Antarctic. The same data as figure II.2

Two taxon ranges generate four range-end events that may be arranged in six different sequences – ties are not recognized and each taxon must appear before it disappears, of course (Fig. IV.2). There are several reasons not to worry about ties. First, apparent ties (e.g. simultaneous extinction) might often disappear with sufficiently precise time-keeping. Second, we are solving a sequencing problem and may reserve ties until a subsequent spacing task in which ties may be recognized as intervals between events that have zero length. Thirdly, we may find several equally well fit sequences. They will differ in the order of pairs of events whose sequence cannot be resolved; these too are ties.

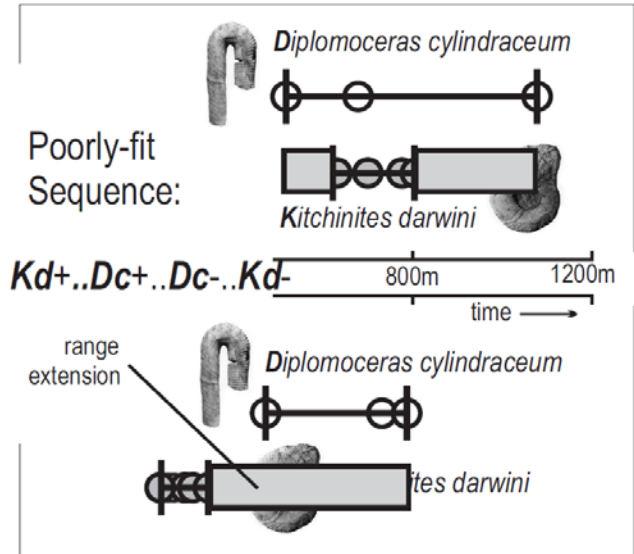
The six possible sequences in figure IV.2 arise as follows. One taxon range may overlap both ends of the other range (sequences 3 and 4), may overlap only one end of the other (sequences 2 and 5), or may be entirely



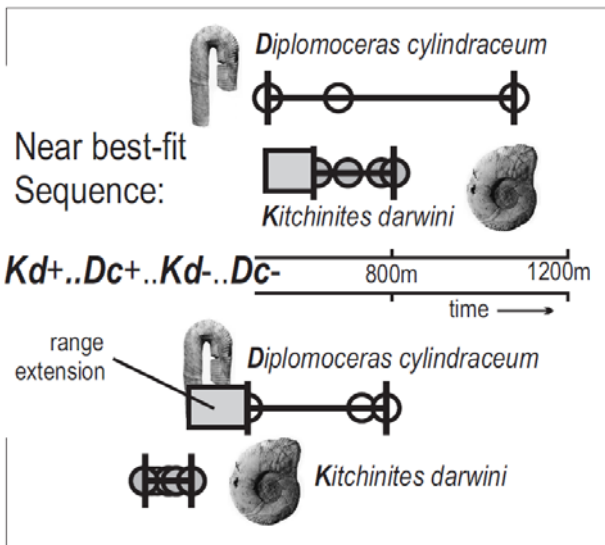
**Figure IV.2.** Six possible ways to order the range end events of two taxa, disregarding the possibility of ties.



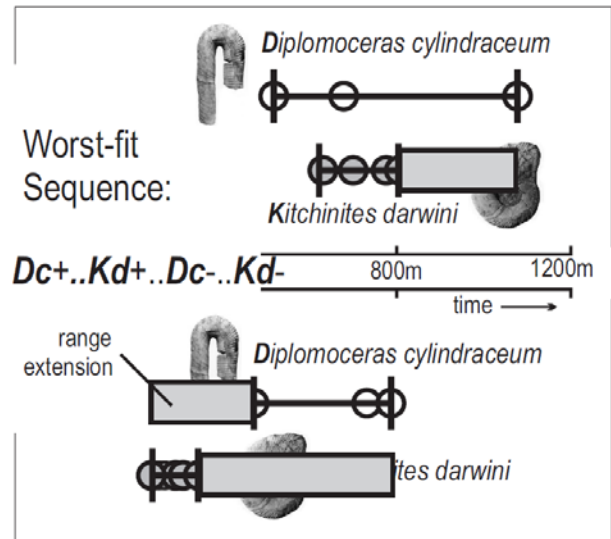
**Figure IV.3** The sequence of events that requires the smallest net extension of the four locally observed ranges.



**Figure IV.5** The sequence of events that requires a large net extension of the four locally observed ranges.



**Figure IV.4** A sequence of events that requires a small net extension of the four locally observed ranges.



**Figure IV.6** The sequence of events that requires the greatest net extension of the four locally observed ranges.

older or younger than the other range (sequences 1 and 6). For convenience in print, we could dispense with the range-chart drawings and represent the solutions as a series of event codes. Each taxon might be identified by its genus and species initials; appearance and disappearance events could have plus and minus suffixes, respectively. Thus, solution 6b becomes (Kd+, Dc+, Dc-, Kd-). For computer storage a single vector or pointer array, or spreadsheet column would

suffice: each event can be assigned to a cell that contains the event rank, or the cells may be physically sorted. It must suffice here to say that there are many simple options for computers to store possible solutions for large data sets.

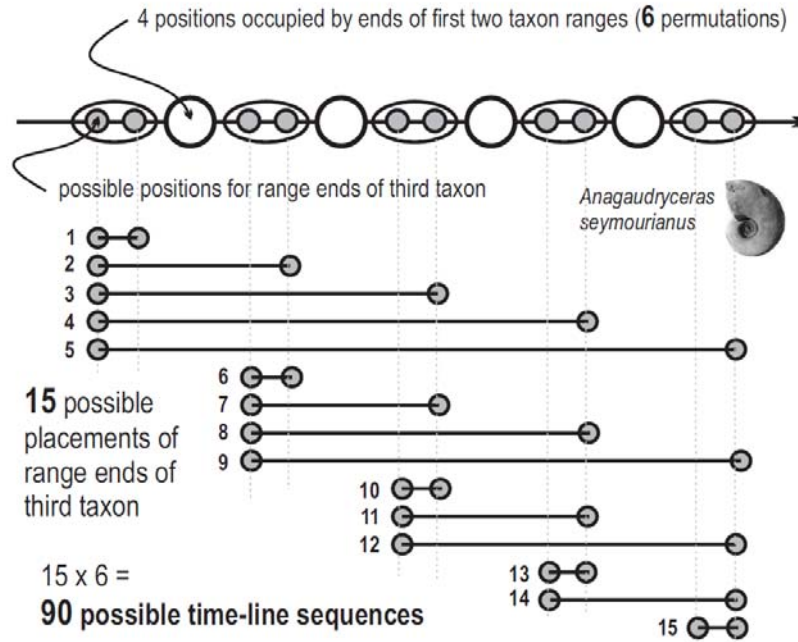
The best solution for the trivial case in figure IV.1 emerges by inversion as follows. Because there is field evidence that *D. cylindraceum* and *K. darwini* coexist, we rule out cases 6e (Dc+, Dc-, Kd+, Kd-) and 6f (Kd+,



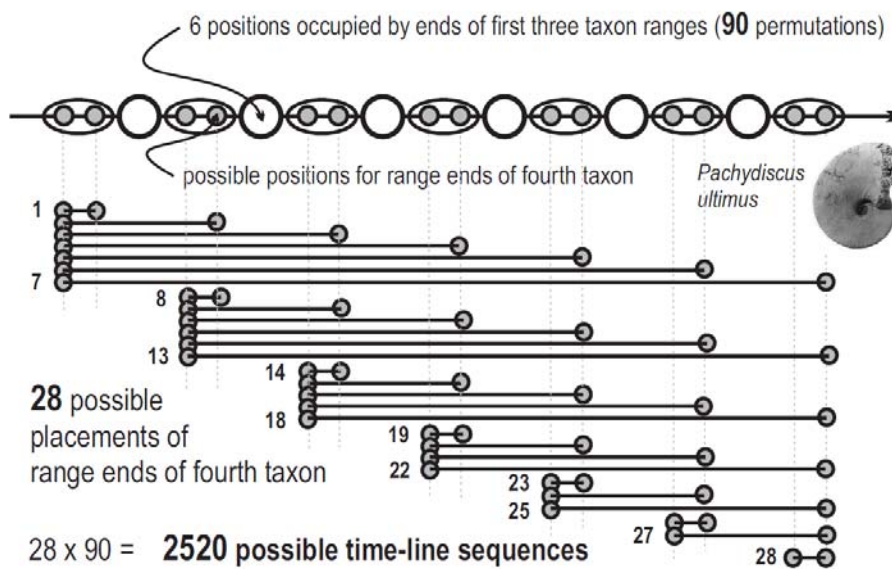
Kd-, Dc+, Dc-). For each of the remaining plausible solutions, determine by how much the observed ranges must be extended to fit the corresponding sequence of events. Solution 6a (Dc+, Kd+, Kd-, Dc-) generates the minimum misfit with the data (Fig. IV.3). It is the optimal, or most parsimonious, or best-fit solution. The process we have followed is a

constrained optimization: eliminate impossible solutions and search among the others for the one with the least misfit to the data. It is implemented by the CONOP9 programs (Sadler, 2003-6)

Note that we have measured the misfit by stratigraphic thickness, as in traditional graphic correlation. This tends to favor the sequence preserved



**Figure IV. 7:** Ninety possible time lines after the introduction of a third taxon



**Figure IV. 8:** Over twenty-five thousand possible time lines after addition of a fourth taxon



in the thickest sections. The size of a range extension can also be measured in terms of the number of fossiliferous horizons or observed events that fall within it. These measures favor the sequence of events preserved in the most richly fossiliferous and intensely sampled sections. They are to be preferred, especially for problems that involve a variety of sedimentary facies or highly variable accumulation rates. Different measures of misfit must be expected to lead to different solutions. Shaw's graphic correlation used only the stratigraphic thickness as a measure of misfit. Edward's (1978) "no-space graphs" counted the event horizons in the range

Number of Taxa:	1	2	3	4
Number of Possible Time-line Sequences:	1	6	90	2,520

5	6	7
113,400	7,484,400	681,080,400

**Table IV.1:** Rapid increase in the number of possible sequences as the taxon count climbs to seven

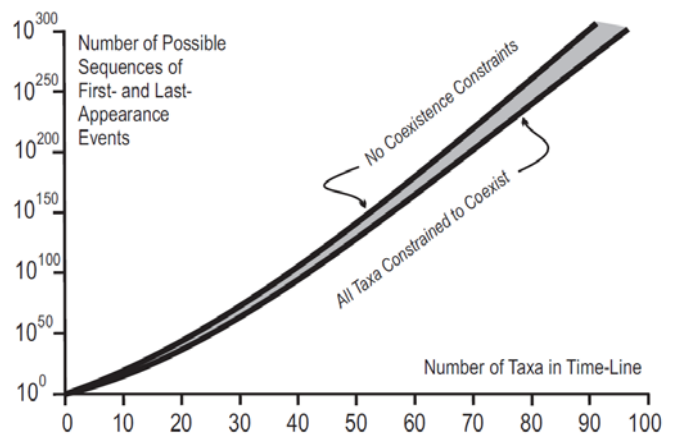
extension. With computer assistance, it is a simple matter to calculate optimal sequences for several measures of fit. Either the best measure is selected to suit the problem, or the robustness of a solution can be assessed in terms of its sensitivity to the measure of misfit.

Compositing usefully large numbers of sections and taxa is much more computationally challenging than our trivial example. Nevertheless, the rules for resolving conflicting information remain as straightforward as in these very small examples.

## IV.2 Adding More Taxa

Large stratigraphic data sets introduce two complications. First, the number of possible sequences increases factorially with the number of taxa (Figs. IV.7-8). Adding a third taxon would enlarge our set of possible sequences to ninety. For seven taxa, more than half a billion sequences are

possible (Table IV.1). For usefully large data sets, the number of possible sequences soon exceeds the number of milli-seconds in geologic time. For the graptolite, ammonite, and mammal examples given earlier, the number far exceeds the counting capacity of popular desk-top spreadsheet programs, which may develop fatal seizures at about  $10^{300}$ . Clearly, exhaustive searches are not feasible. Fortunately, computer science has developed heuristic search strategies that find very good solutions to this class of problem (so called "non-deterministic polynomial-time complete" or "NP hard" problems, Dell et al. 1992) without evaluating every possible sequence.



**Figure IV.9:** As the number of taxa approaches practical values, an exhaustive search of all possible sequences is out of the question.

The second complication is that there may not be a unique best-fit solution. This situation may arise even in the smallest data sets and the number of equally most parsimonious solutions is likely to grow as the data set grows. From the set of non-unique solutions, we may develop a very simple measure of the resolving power of each event: the range of positions in sequence that the event may take without increasing the misfit (Sadler and Cooper, 2003).

## IV.3 Which Information is Trustworthy?

We resort to constrained optimization because local biostratigraphic information is never fully reliable. Observations that are simply wrong (e.g. false identifications) always need to be purged. Here we consider how much reliable information can be extracted from the observed positions of first- and last-

appearances. Three properties are potentially useful: the order and spacing of range-end events and the overlap of taxon ranges. The most reliable, incontrovertible observations may be applied as a constraint. Some or all of the others, which are subject to adjustment, may be incorporated into measures of misfit. These measures are variously known as the objective functions, costs, or penalty functions to be minimized (optimized).

Expert opinions differ about the scope of stratigraphic information that may be considered reliable. Where alternative constraints and objective functions can be justified, this may exacerbate the non-uniqueness of the results and enlarge the uncertainty intervals in the composite succession (Sadler et al, 2008). This is a proper reflection of the degree to which inadequate information leaves any instance of the problem underdetermined.

#### **IV.4 Observed Stratigraphic Spacing and Order**

Shaw's graphic correlation uses both order and stratigraphic spacing of local range-end events. Hood's (1986) Graphcor program automates Shaw's procedure. The equivalent optimization in Conop9 uses the "INTERVAL" measure of misfit, which totals the local stratigraphic thicknesses of all range end-adjustments. Both order and spacing are influenced by migration and preservation. Stratigraphic thickness suffers from the additional complication that accumulation rates are neither steady from time to time, nor uniform from section to section. Accordingly it is usually preferable to use a different spacing scale or drop the spacing information altogether. Edwards' (1978) "no-space graphs" resemble graphic correlation but use ordinal stratigraphic scales with uniform spacing of event horizons. In Conop9, the "LEVEL" function measures range extensions in this way, counting only the number of event levels by which the range-end is moved. The "EVENTUAL" function is similar but scores each event level according to the number of events at that level.

All three of these Conop9 penalty functions are used to determine the minimum cost of adjusting all local successions to fit the trial time-lines. Other penalty functions, discussed below, simply compare the order of events in the trial time-line with the

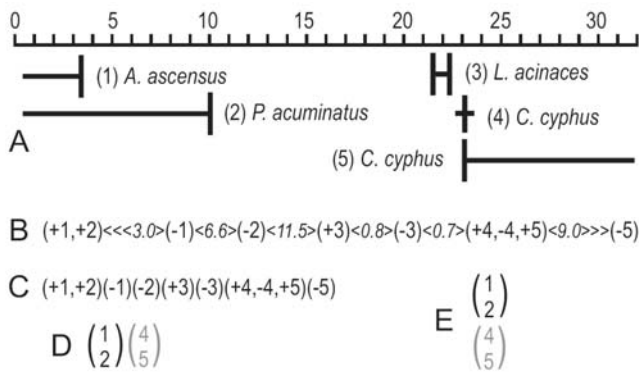
ordering information in the field data. Unlike graphic correlation, they do not fit the local sections to the time-line.

#### **IV.5 Observed Stratigraphic Overlap.**

Guex (1991) has long advocated that neither the local spacing nor the detailed local ordering of range-end events is sufficiently reliable information for biochronology. Rather than fitting solutions to the order of range ends, his method of Unitary Associations (implemented in the Biograph program) uses only the observed overlap of ranges. Alroy's (1992) Conjoint program and the seriation practices of archaeologists use a similar strategy. These programs are applicable to isolated fossil collections and sites that lack the superpositional information necessary for graphic correlation and no-space graphs.

Extended ranges may create new coexistences but cannot falsify observed coexistences. Thus, observed (as opposed to implied) coexistences may be used as a measure of misfit in constrained optimization. When constrained optimization is set up to mimic graphic correlation or no-space graphs (INTERVAL, LEVEL or EVENTUAL objective functions in CONOP), observed coexistences are recognized as powerful constraints that must be honored by the composite sequences. The best fit sequences almost inevitably include additional coexistences, not observed anywhere. This is especially prevalent where the data include coeval but separate biotic provinces. The number of these additional coexistences is another measure of misfit with the local information and may be used as the objective function to minimize (ROYAL option in CONOP).

Rules that permit only range extensions, actually force more constraints on the solution than those derived from observed coexistences: no observation of a first appearance before (below) a last appearance can be falsified. Alroy (1994) realized that this constraint (FORCEFB4L='ON' in CONOP) applies equally to individual taxa and to pairs of taxa whose ranges do not overlap. His "Appearance Event Ordination" thus leverages more power than coexistences alone. The same strategy is implemented as the SEQUEL objective function in Conop9, which counts excess first-before-last pairs in the time-line but not required by the local sections. It is particularly useful for time-lines that exceed the durations of the individual sections.



**Figure IV.10:** Different levels of biostratigraphic information extracted for 5 of the 56 taxa observed in the Llandoveryan (Silurian) Yewdale Beck section (Hutt, 1975). 1: the range chart; the most commonly published graphical summaries of biostratigraphic data; range ends are plotted against a thickness scale and linked by continuous lines. 2-5: progressive reduction of the information to its potentially more reliable core; the notation is invented here; there is no standard practice. 2: range-end order and spacing; numbers correspond to taxa in range chart; positive numbers in parentheses are first-appearance range-ends; negative numbers are last-appearances; events at same level are placed in same parentheses; relative spacing in italic numbers. 3: range-end order only. 4: overlap and order; coexisting taxa share vertical parenthesis; black parentheses for certain overlap (*sensu stricto* evidence of coexistence); gray parentheses for coexistence at a single level only (*sensu lato* evidence). 5: overlap information only. Genus abbreviations: A. Akidograptus; P. Parakidograptus; L. Lagarograptus; C. Coronograptus.

The data for long time-lines will include pairs of sections that share no taxa because they do not overlap in age. The ROYAL and SEQUEL objective functions will tend to keep these taxa apart in the time-line to avoid implying excess coexistences. Thus, the two functions are not simply a conservative focus on the least falsifiable information, they also carry the implicit assumption that faunal dissimilarity more likely arises from differences in age than differences between coeval provinces (Sadler and Sabado, 2009). For the same data sets, traditional economy-of-fit adds

no penalty for interleaving events from dissimilar biotas; that is, graphic correlation effectively prefers to attribute dissimilar biotas to provinciality. Clearly, the decision to use all evidence of sequence or only that which cannot be falsified by range extension must also be guided by the temporal and biogeographic scope of the data set.

#### IV.6 Reworked Fossils

Do we ever have reason to shrink locally observed ranges? Yes. The observed range may be too long, or entirely out of place, for four reasons. First, fossils are sedimentary particles that may survive a cycle of erosion and re-deposition above the true last appearance horizon. This “reworking” can create either a last appearance or an entire local range that is too late. Second, benthic mixing may lower a first appearance by the thickness of the mixing zone. Fossil occurrences may appear to be anomalously old in well cuttings if the bore-hole walls collapse (“caving”) or the mud circulation is inefficient. Finally, mis-identification can generate totally spurious ranges.

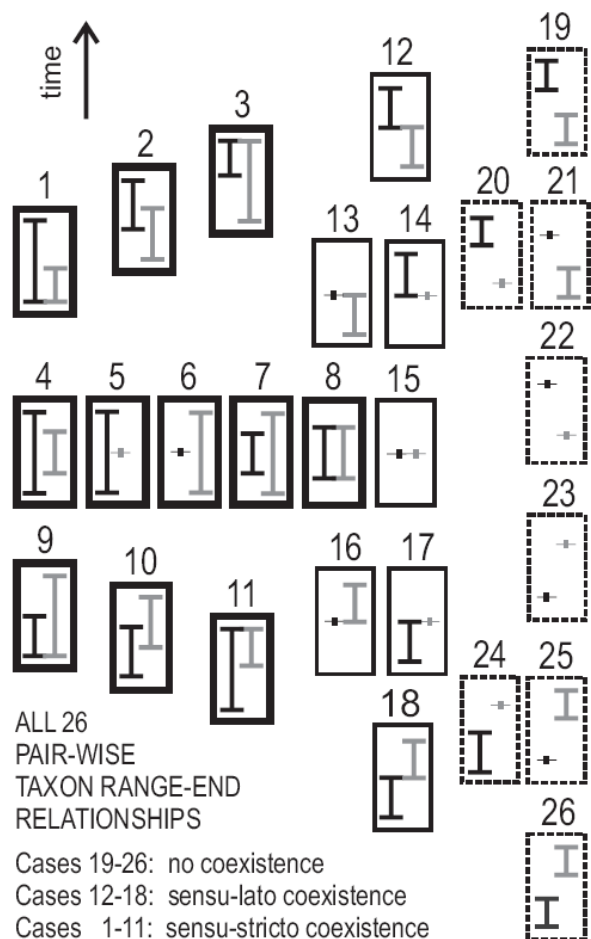
Bad taxonomy, which can ruin any time-line, is a routine issue of quality control. Bore-hole caving can be avoided by using only last occurrences. Reworking, and to a lesser extent downward mixing, are more troubling, especially for robust microfossils, such as palynofloras in particular. Where the observed range can be longer than the true range, all the objective functions discussed so far are invalid.

The RASC algorithms (Agterberg and Nel 1982, Agterberg 1990; Gradstein and Agterberg 1998; Agterberg and Gradstein, 1999) fit the composite sequence to the most commonly observed pairwise ordering of range-end events. In Conop9 the ORDINAL, SPATIAL, and RASCAL objective functions are comparable. They differ by considering the number, the spacing, or the fraction of observed pairs, respectively, that contradict the trial time-line. While other objective functions assume that reworked occurrences have been eliminated from the data, these functions over-simplify in the opposite direction; they treat local ranges as if they are as likely to over-estimate the true range as underestimate it. Although this is not the most satisfactory strategy, it proves very difficult to develop computer algorithms that recognize reworked fossils. Where this is easy for the human operator, perhaps because the specimens are so worn or the taxa clearly out of context, the occurrences are

purged from the range determinations. Where obvious signs are lacking, safe and logical rules are equally difficult to teach to humans and computers.

#### IV.7 Assemblage Zones and Interval Zones.

Sometimes the decision whether to use all sequencing information or only evidence of coexistence is the difference between sequencing range-end events and sequencing taxa. The latter offers fewer permutations. Two taxa have three possible ordering outcomes: the two ranges overlap or not; if not, either taxon may be the older. This treatment is sufficient to build time-lines of taxon



**Figure IV.11.** Twenty six possible arrangements of the four range ends of two taxa when “ties” are considered (from a standard CONMAN screen analysis).

assemblages that can be used to define biostratigraphic assemblage zones (Guex, 1991).

Alternatively, as in graphic correlation, two taxa are treated as four range-end events for which there are 26 possible sequences, including ties (Fig. IV.11). Time-lines of events are a more fully resolved estimate of paleobiologic history and can be used to select a series of first- or last-appearance events to define biostratigraphic interval zones. But there is a logical inconsistency with some uses of traditional interval zones. Zone boundaries use only a culled subset of range-end events to avoid those that can be observed in the wrong order. This perfectly reasonable practice acknowledges that local first- and last-appearance events are diachronous horizons. So the local event horizons should *not* be assumed to correlate from place to place. Graphic correlation and the constrained optimizations based upon it estimate what local event horizons might better approximate the age of true origination and extinction – the range extensions based on information from all other localities.

#### IV.8 Constraints from Non-Paleobiologic Events

In the trial-and-error approach, other types of chronostratigraphic event may readily be added to time-lines. They need different rules for determining their feasible positions and mutations. To be useful, these rules greatly reduce the number of feasible sequences that the search must consider.

*Correlative horizons (marker beds).* – These are layers that can be matched or traced exactly between localities. Examples include ash-fall tuffs with geochemically “finger-printed” compositions. Their position in measured sections is not in doubt and, unlike taxon range ends, may not be adjusted. Their permissible positions in the time-line must honor observed superpositional relationships relative to events that cannot be adjusted toward them. They do not move relative to one another. First-appearances seen below them anywhere must be older; last-appearances seen above them must be younger.

*Correlative intervals.* – These are events that can be matched from section to section but whose location is limited to a finite uncertainty interval. Examples include seismic reflectors whose position cannot be resolved to better than a few meters in a core or stable isotope anomalies whose peaks or limits are not

everywhere well-defined. We can say only that a correlative horizon does exist that runs through each section somewhere inside a locally specified interval. These uncertainty intervals may be adjusted but, unlike taxon ranges, they shrink to fit rather than stretch. They serve to guide the coarse early stages of the search for best-fit timelines or augment weak paleobiologic constraints.

*Dated events* – The order of events that have been dated directly or indirectly by radioisotopic methods is established without recourse to superposition. In sufficient numbers, dated events are the most powerful independent constraints on time-lines (Sadler, 2006). Incorporated into the sequencing task from the outset, with any associated faunas, they acquire positions in the time-line and uncertainty intervals that facilitate numerical time calibration without prior recourse to biostratigraphic zonation (Sadler et al., 2009).

---

**Table IV.2** Properties and Information Content of Time-Stratigraphic Events

	Uniqueness of Event			Freedom to adjust stratigraphic position				Nature of change				May occur at surface of hiatus	Reliable ordering information
	Unique	Binary	Multiple	Up only	Up or down	Down only	None	Present-absent	Alternation	Repetition	Transient		
<b>TAXON RANGE-ENDS:</b>													
<b>Reworked last app.</b>	X				X			X				Y/N	above FAD
<b>Last appearance</b>	X			X				X				Y/N	super-position only <sup>2</sup>
<b>First appearance</b>	X					X		X				Y/N	sub-position only <sup>1</sup>
<b>Caved first app.</b>	X				X			X				Y/N	below LAD
<b>REVERSALS:</b>													
<b>Paleomag reversal</b>		X					X	X				Y/N	position <sup>3</sup>
<b>Isotope cycle boundary</b>		X					X	X				Y/N	position <sup>3</sup>
<b>RESETTINGS:</b>													
<b>Sequence boundary</b>			X			X		X				Y	position <sup>3</sup>
<b>Flooding surface</b>			X			X		X				Y	position <sup>3</sup>
<b>SPIKES:</b>													
<b>Marker bed</b>	?	?				X		X				N	position <sup>3</sup>
<b>Isotopic transient</b>	?	?				X		X				N	position <sup>3</sup>
<b>TAXON ACMES</b>			X		X			X				N	between FAD and LAD
<b>DATED ASH FALLS</b>						X		X				N	position <sup>3</sup> and age value <sup>4</sup>

<sup>1</sup> must be older than type <sup>2</sup> and <sup>3</sup> events higher in section, but might be shifted with respect to lower events.  
<sup>2</sup> must be younger than type <sup>1</sup> and <sup>3</sup> events lower in section, but might be shifted with respect to higher events.  
<sup>3</sup> must be younger than type <sup>1</sup> and <sup>3</sup> events lower in section and older than type <sup>1</sup> and <sup>3</sup> events higher in section.  
<sup>4</sup> age relative to other type <sup>4</sup> events determined by value in numerical age scale; does not require occurrence in same section.

## V. The Traveling Salesman's Solution

**Explains how to solve the biochronological sequencing problem by analogy.**

### V.1 Traveling Salesman Problems

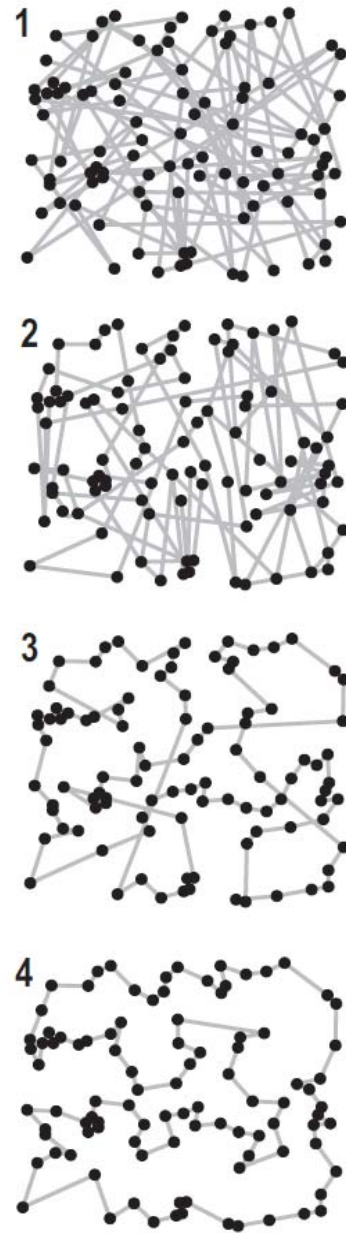
The traveling salesman problem (TSP) is easily stated

“find the shortest tour that visits each city once and only once”

but it is one of a difficult class of combinatorial optimization problems termed NP-Complete. If an exact solution to these problems is sought by searching all possible sequences, the running time increases exponentially, or worse, with the size of the problem – in this case the number of cities.

After the investigation of time lines in the previous chapter, it probably comes as no surprise that the paleobiologic time-line problem has been attributed to the NP-Complete class (Dell et al., 1999). These problems include many practical applications and have been investigated for decades. Once we map the paleobiologic time-line problem into the terms of the TSP, it can be solved by the same methods. The mapping is very straightforward: the taxon range-end events become the cities and net range adjustment becomes the travel distance. The best-fit sequence is the shortest route; it requires the least adjustment (Kemple et al., 1995). Compared with the number of range ends, the number of stratigraphic sections adds trivially to the computation time; it lengthens the calculation of the range adjustment, but the increase is linear.

Exact methods for solving the TSP become impractical before the number of range ends reaches a useful value. Fortunately, there are heuristic methods that lead to good approximations in more acceptable running times. Heuristics are strategies that lead to the “discovery” of an answer, often by analogy with biological or physical systems, that are simpler to understand and implement than direct mathematical determination (the N in NP-Complete stands for “Non Deterministic anyway!”). We need a heuristic or trial-and-error strategy that can lead from an initial guess to

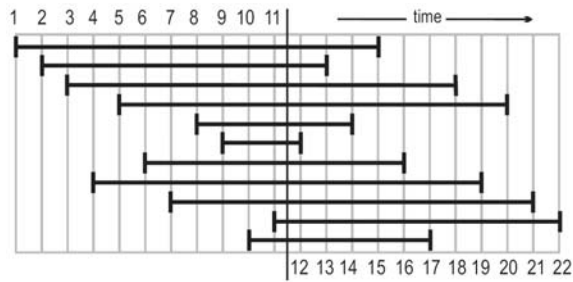


**Figure V.1:** Four snapshots from a trial and error solution of a traveling sales man problem from a random starting tour (1) through numerous random path changes (. . . 3 . . 4 . .) to the shortest tour (4).

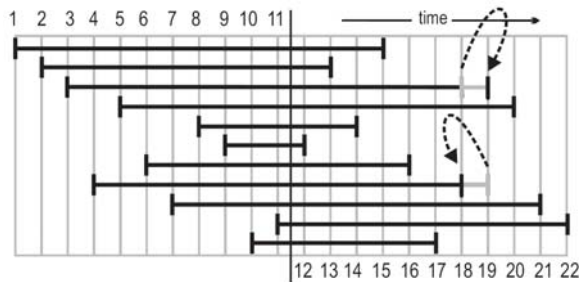


a very well fit time-line while examining only a fraction of the possible sequences.

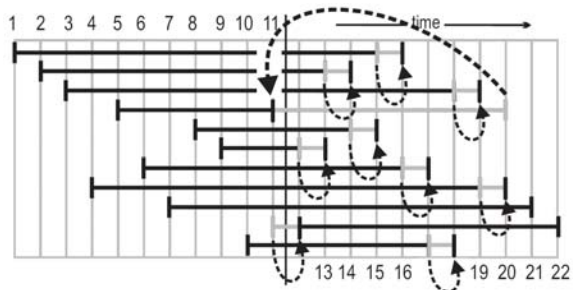
The TSP can be solved as follows (Fig. V.1): choose an initial “feasible” tour at random and measure



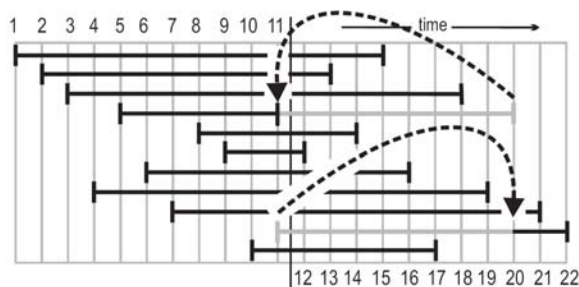
**Figure V.2:** In this simple and sufficient initial sequence all taxa coexist.



**Figure V.3:** The smallest mutation option.



**Figure V.4:** The most efficient mutation option moves a randomly chosen event to a random position.



**Figure V.5:** A more radical mutation protocol switches two events chosen at random.

its length; then randomly select a neighborhood of four cities and make a new tour by rearranging their links; determine the length of the new tour; if it is shorter or not “too much” longer than the last, make another rearrangement; if it is “too much” longer, undo the rearrangement and try another; repeat until a “reasonable” number of rearrangements no longer shortens the route. Whether for the salesman or the biostratigrapher, operational definitions must be provided for “too much” and “reasonable.” Different heuristics build different procedures from these definitions. We will examine how CONOP implements the simulated annealing heuristic. But first, we need a biochronological definition for “feasible” and an equivalent for the tour rearrangements, which we will call mutations (Figs. V.2-5).

The range-end events in a feasible time line must satisfy the constraint that every taxon appears before it disappears. We have also seen that expert biostratigraphers may wish to preserve every observed coexistence or every observation of a first-appearance before a last appearance. For the initial guess (Fig. V.2) these constraints can all be satisfied by placing all first-appearance events at random in the first half of the sequence followed by a randomized set of all last-appearances.

Non-biological events are placed in order in the middle of the initial sequence. The order of dated events is constrained to match their numerical ages. The order of undated marker events, such as chemically fingerprinted ash fall deposits, may be constrained by superposition; unlike observed range ends, they are not subject to adjustment up and down section.

CONOP offers three options for mutating the sequence. The smallest and simplest mutation strategy swaps a randomly selected event with the next in sequence (Fig. V.3). Such a mutation simplifies the re-calculation of fitness after mutation, but becomes progressively more inefficient as the size of the problem increases. The most generally efficient and practical mutation moves an event at random to a new position anywhere in the sequence (Fig. V.4); all events leapfrogged by this mutation move one place in the opposite direction. A more extreme mutation switches two events chosen at random anywhere in the sequence (Fig. V.5); such mutations rapidly change the character of a time line but are ill-suited to making fine adjustments.



## V.2 Heuristic Search Strategies

For a concrete visualization of heuristic searches, imagine a landscape on which each point corresponds to a different, feasible time-line of events and elevation is the total misfit between that time-line and the field data. The best-fit sequence lies at the lowest point on the landscape. Neighboring points on the landscape are sequences that differ by a single mutation; e.g. two events that swap positions. Topography varies, sometimes considerably, with the measure of misfit and the kind of mutation that is permitted.

The search strategies string together steps (mutations) across the landscape to make tracks from random starting positions to the lowest point. The landscape is so extensive, even for small datasets, that an exhaustive survey is not feasible (Sadler and Cooper, 1993). The searcher feels its way blindly one step at a time, with no map beyond its memory of previous steps. At every step, the heuristic algorithms loop through numerous iterations of a few very simple operations: use a random number generator to choose the next trial mutation; total the change in fit caused by that mutation; and compare the change with a threshold that determines whether the mutation is retained or removed. The threshold determines which mutations make the fit “too much” worse to be worth keeping..

Memorizing the full map of prior steps is optional; only the coordinates (time-line) of the lowest and current position are required. Memory management slows every step in a search path that might take millions of steps. The landscape is complex and most likely “lunar” in the sense that it includes many closed depressions that are not as deep as the best-fit point. The search strategy must be designed to recognize or escape from such local minima.

The landscape might be searched in the manner of a rain drop as follows: from a random starting position, and at each subsequent position, map the changes in elevation (fitness) for one small step in every possible direction (single mutations); take the steepest downhill step (adopt the mutation with the biggest improvement in fit); then scrap the old neighborhood map and start again at the new position. Such a search most likely plunges uselessly into the nearest local minimum; that is, a sequence whose fitness cannot be improved by any one mutation, but is not the best possible. The neighborhood mapping is not essential. Similar suboptimal destinations result from choosing the first downhill step (beneficial mutation) discovered at each location. With enough raindrops, some will find the true

minimum elevation, but it is difficult to know how many raindrops will be “enough.”

Genetic algorithms start a handful of rain-drop searches and then periodically restart with hybrids of the fittest of them. The hybridization process avoids traps set by local minima. Zhang and Plotnick (2001) employed a genetic algorithm to fit the LOC in graphic correlation. For time-lines, however, viable hybrids are time-consuming to implement. Instead, Kemple et al. (1995) recommended the simulated annealing heuristic (Kirkpatrick et al., 1983).

## V.3 Simulated Annealing

The simulated annealing heuristic uses only a single landscape searcher but adds the ability to climb up-slope. The searcher is programmed to be more likely to accept a gentle uphill step than a steep one (prefers less bad mutations) and steadily loses its tolerance for uphill steps as the search progresses. In effect, it searches coarsely at first and more finely as the search progresses; i.e. climbs progressively smaller hills. The formula that determines the probability of accepting an uphill step uses the Boltzmann equation from thermodynamics and the analogy of growing a perfect crystal by annealing and slow quenching (Kemple et al., 1995). The waning energy for uphill steps on the landscape now becomes the quenching process – a cooling curve.

A few parameters must be chosen to ensure that there are sufficient steps in the search to reach the lowest point. They are analogous to the starting temperature and the cooling rate. Most of this can be handled by the CONOP software. More critical decisions concern how the field information will be used to measure the fitness after each mutation. Fortunately, the simulated annealing heuristic imposes almost no limits on the mathematical properties of the fitness formulations and constraints (Ingber, 1993); so, the decisions can be purely biostratigraphical. The price of this flexibility can be a long computation time. The benefit is that we can try to build an artificial stratigraphic intelligence into the landscape searcher.

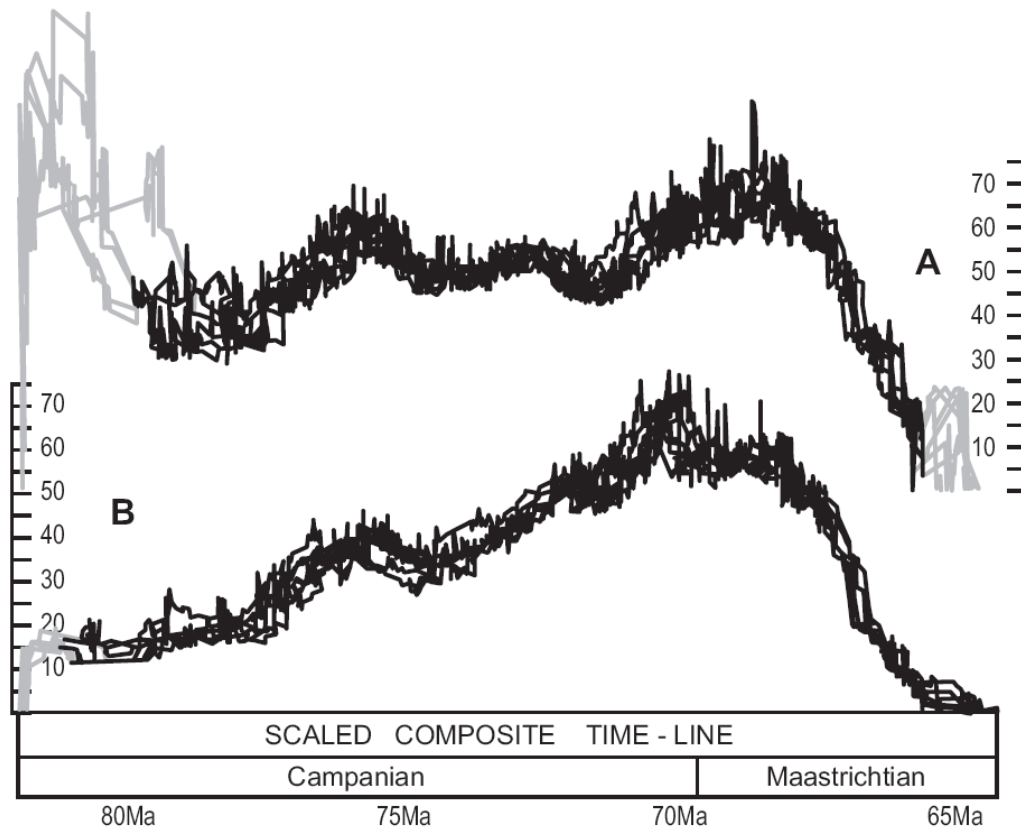
Although the search is random, good mutations tend to be found in clusters – when the searcher happens upon a downhill slope. On a complex, multi-dimensional landscape, the searcher will repeatedly wander onto steep descending slopes and it is tempting to give the searcher a little more intelligence

than the minimal simulated annealing rules. The tabu search heuristic (Glover, 1989), for example, adds intelligence via dynamic “first-in-first-out” shortlists of deleterious mutations that should be remembered and not repeated for a while (really bad mutations go onto longer lists and persist longer). Dell et al. (1992) were unable to design tabu lists that could make the tabu search more efficient than simulated annealing. Time spent managing memory seems to be better used trying more mutations.

A number of simple strategies might be added to increasing the rate of beneficial mutations. Unfortunately, these merely gets the searcher caught in local optima as surely as cooling too rapidly. A large proportion of temporarily deleterious mutations are essential to ultimately improving fitness; i.e. either keep the brute very simple-minded or switch to an altogether more specialized approach to the problem.

## VI. Inside Composite Time Lines

**Demonstrates the use of composite time lines as a tool for quality control and paleobiologic interpretation.**



**Figure VI.1** Late Cretaceous ammonite taxon richness. A: richness corrected for differences in section coverage. B: raw richness. All curves based on time lines within 2% of the best-known fit to a global data set of 3789 local observations of 605 taxa and 58 other events in 201 sections. Curves are grayed where compromised by artificial truncation or very low section coverage.

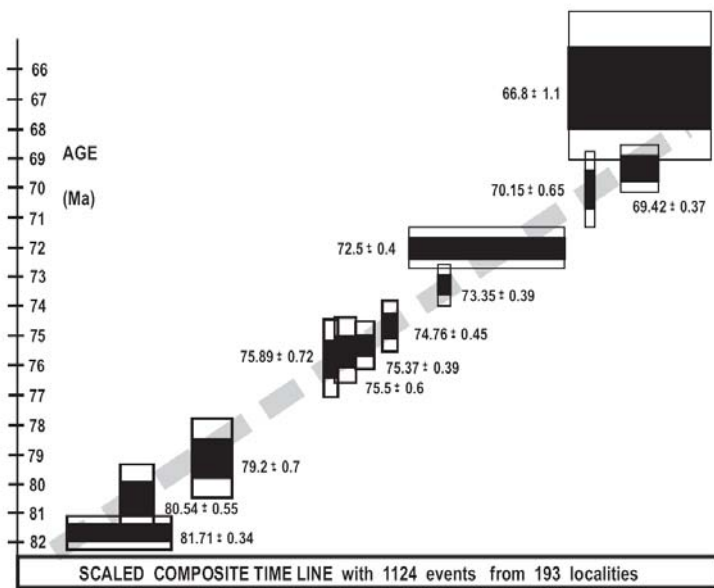
### VI.1. Taxon Richness Curves

Because paleobiologic time-lines include an ordered series of first- and last-appearance events, it is a very simple matter to develop an unbinned taxon richness history. The required time series is a running total incremented by one for every FAD and decremented by one for every LAD. Of course, the precise position of each event in the time line is not uniquely determined. This uncertainty can be transferred to the taxon richness curve by superimposing curves from a set of equally well-fit time lines (Fig. VI.1).

The time-lines also readily support time series of taxon longevity, origination rate and extinction rate. Traditional measures assume that data will be binned into zones or stages. The time-line data are not binned; this causes trouble if conventional formulae are applied at each event. For example, the number of originations at each event is either one or zero. Consequently, origination and extinction rate must be expressed for moving average windows. A mean longevity may be calculated for all taxa extant at each event horizon (i.e. ranging through). Note, however, that mean longevity and taxon richness are

mathematically interrelated. Mean longevity is the sum of taxon range lengths divided by the number of taxa. The denominator is richness and potentially a much smaller number than the numerator. For rich clades, the origination or extinction of one short lived taxon changes the richness by 1 while hardly altering large sums of longevity. Consequently, segments of plots of mean longevity against taxon richness (one of many standard graphical outputs from CONOP) in which rare short-lived taxa come or go in long runs form hollow curves whose form approximates a relationship in which mean longevity is the reciprocal of richness multiplied by a near-constant.

The taxon richness history in figure VI.1a has two enhancements. First, it is plotted on a time scale – a scaled time-line in which events are spaced according to the time intervals between them. Second the taxon richness includes a correction for differences in the number of stratigraphic sections that provided information at each horizon in the time line. The next sections consider how to add these features.



**Figure VI.2** Calibration of a scaled composite time line, using radioisotopically dated events included in the compositing process. Biostratigraphic data are Campanian and Maastrichtian ammonite ranges. Rectangle dimensions show analytical uncertainties (height) and range of possible positions in the composite time line (width).

## VI.2. Calibrated Time Scales

Taxon richness curves are most informative when plotted on a scaled time-line. The slopes of the curve then begin to reflect rates of radiation and extinction. If the time line includes dated events, it is possible to calibrate the age of all events by interpolation. For this purpose, examine a cross-plot of the composite sequence against a time scale (Fig. VI.2).

In order to be able to make such a cross-plot within CONOP, add a “pseudo-section” that uses age in place of thickness and includes all the dated events at the appropriate level. Mark this section “no” in the last column of the section dictionary, to indicate that it is not used to generate a scaled composite from average thicknesses. The CONOP graphic (2-section LOC) will plot the mean age. Analytical and stratigraphic uncertainty must be added elsewhere.

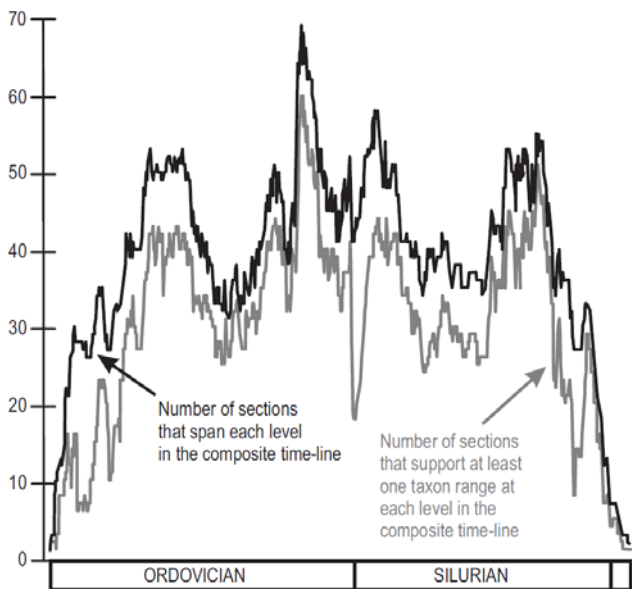
The main purpose of this cross-plot is to determine which of the ten scaled composite sequences generated by CONOP, if any, has a near-linear regression of position in sequence against time. It can be used directly as a proxy time scale (Sadler et al., 2009). Otherwise the timeline scaling may need to proceed by piecewise linear interpolation between adjacent pairs of dated events.

The scaling options in CONOP include an unscaled (ORD) time-line in which events are uniformly spaced. Other scaled time lines are based on minimum or maximum thicknesses between adjusted range ends and on mean thickness between adjusted range ends. The mean values can be calculated with or without the zero values. Inclusion of zero values may compensate for clustering of events at mass extinctions. Thicknesses may be as measured or after the rescaling the total thicknesses of the sections to their span of events in the composite sequence. The rescaling attempts to compensate for differences in long term accumulation rate. Refer to the CONOP Reference Manual for all options and Sadler et al. (2009), for discussion of their application to time calibration.

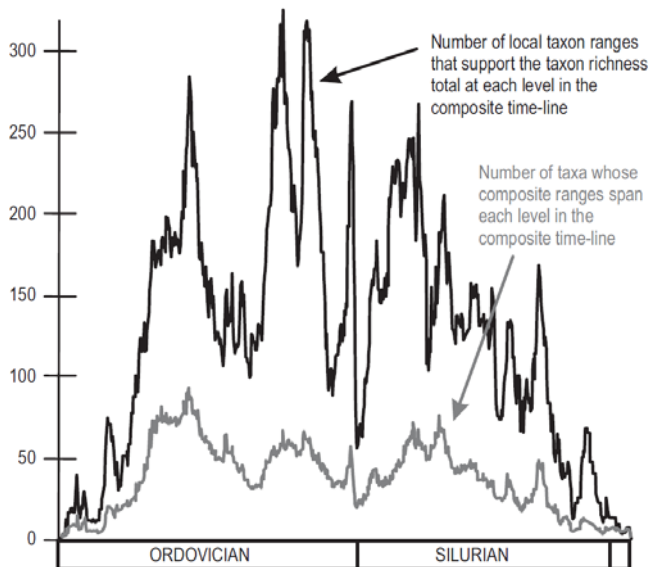
## VI.3. Section Coverage

The taxon richness history determined from the composite order of events (Fig. VI.1b) will be distorted by the differing number of sections that support different parts of the time line and by

variations in the richness of the various sections. CONOP offers several ways to explore and mitigate these effects. Two simple graphs of section coverage are available (Fig. VI.3). One plots the number of sections that span the time of each event horizon in the time-line, whether or not the sections contribute a local range for any of the taxa extant at that time. The other plots the smaller number of sections that span



**Figure VI.3.** Section coverage for a time line of the total graptolite clade, based on over 500 sections world-wide

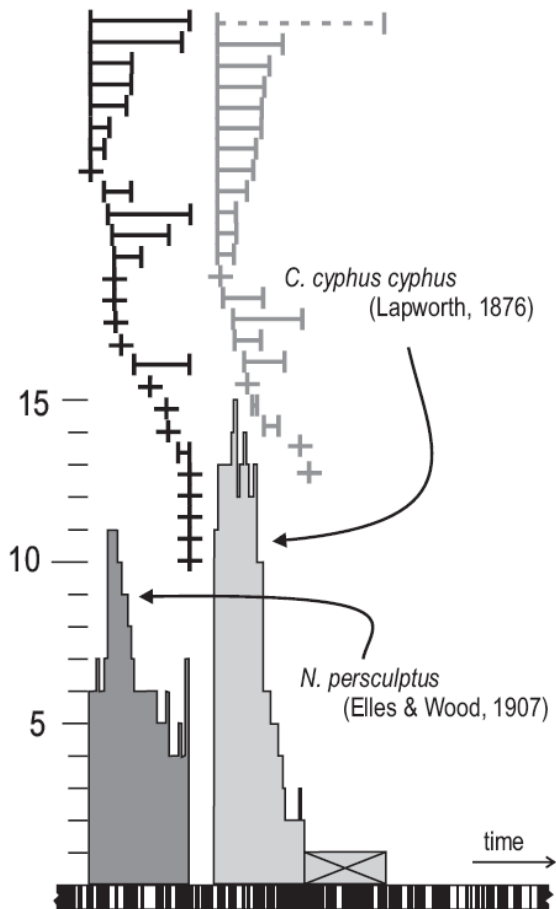


**Figure VI.4.** Local taxon support for a richness time series from a time line of the total graptolite clade, based on over 500 sections world wide.

the event horizon in the time-line contribute at least one local range to each level in the time line.

The distance between the two curves in figure VI.3 is driven by sections with sampling gaps that are not filled by range-through taxa. The closer the two curves the better, because closeness indicates strength in the underlying data.

In order to make the second variant of the section-coverage graph, CONOP must map the locally observed range ends back into the time-line. This



**Figure VI.5** Fragment of composite time-line for whole graptolite clade (506 sections and 2090 taxa) showing support for composite ranges of two taxa that are never observed to coexist. Horizontal “bar-code” at base is scaled composite section (Sadler et al., 2009) with event levels represented by black bars. Grey histograms show number of supporting local ranges (horizontal lines at top of diagram). Local ranges are arranged in order of first appearance. Dashed local range and crossed portion of histogram identify an erroneously long range.

same information can also be used to draw a graph of the local support for the taxon richness history (Fig. VI.4). Again two graphs are drawn. At each level in the best-fit composite time line, CONOP can plot the number of locally observed ranges (one or more for each extant taxon) and the smaller number of different taxa represented.

The latter curve is the raw taxon richness history; the former is the global support for the richness history. The species richness history can be represented on a per-section basis. If the global support is treated as the sample size for each level, however, as if the whole world were the ecologist's sample plot at each time horizon, then CONOP can perform a rarefaction on the global taxon richness history using the routine equations from field ecology. These formulae allow CONOP to draw an expected richness curve with standard deviations for a chosen level of uniform sampling; i.e. reducing the support to a lower common denominator.

In figure IV.4 a wide separation of the two curves is desirable – the wider the separation of the curves the more voluminous the support for the taxon richness curve. Rarefaction asks what the lower curve (richness history) would look like if the upper curve (sample size) were flat. It answers the question by randomly reducing the peaks of the upper curve – losing information but smoothing out the sampling.

As an alternative and conceptually superior approach to rarefying the composite time-line, one might imagine randomly reducing the local information to a uniform level *prior* to performing the constrained optimization. There are two snags with this noble idea. First, a complete optimization run is needed to determine which slices of which sections are likely of the same age. Then, for each of the time slices in the composite, a large number of random re-sampling exercises will be needed to generate a representative population of uniformly (under)sampled faunas. For each of member of this population, an optimization run is needed. For large data sets each run may take several days and the quality of the outcomes will be compromised by the culling of information. The object of rarefaction is not the time-line itself, but the richness history it implies. It is much better to build one composite time line using all the relevant information and perform the rarefaction on it.

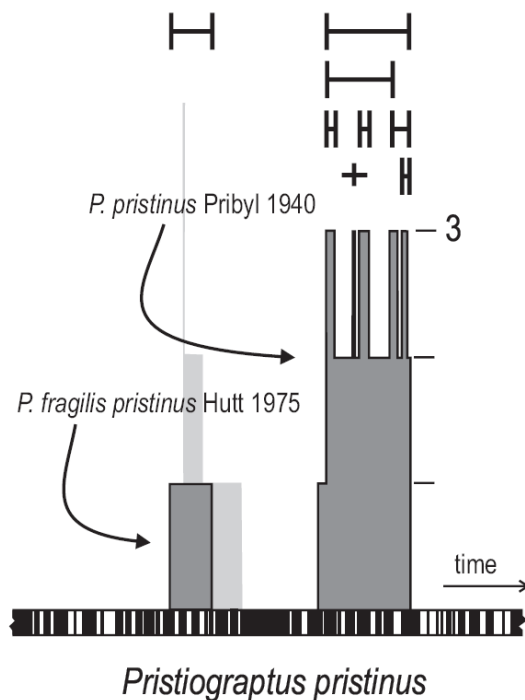
Having seen the advantages of examining the distribution of all local ranges in the time-line, we can

now try a similar exercise within each composite taxon range.

#### VI.4. Support Within the Composite Range

The locally observed ranges that have been mapped back into the time line combine to generate the composite taxon ranges. To reveal the amount and distribution of support for a composite taxon range, CONOP9 draws a histogram that shows the number of supporting local ranges as a function of position within each of the composite ranges (Fig. VI.5-8).

Notice that the ends of the local ranges do not all correlate. This confirms that they vary in age. The optimization algorithms would align all local range ends, if possible (whether or not this is correct), to avoid the penalty of adjusting range ends (especially for the frequently encountered, short-ranged taxa illustrated here). It is not possible. Accordingly, it is



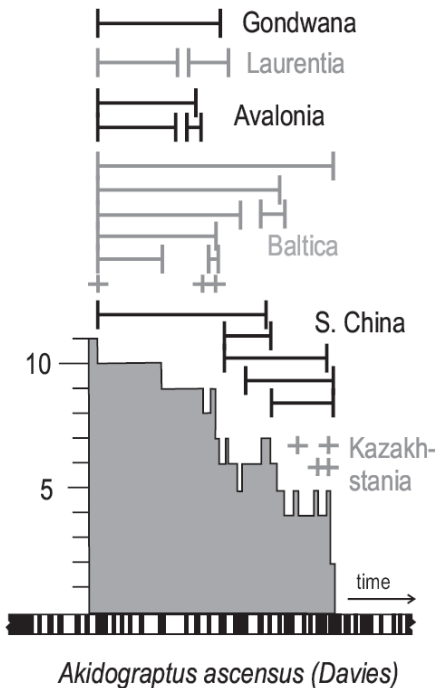
**Figure VI.6:** Fragment of composite time-line for whole graptolite clade showing disjunct support (dark gray histogram) for composite range indicative of a taxonomic irregularity - two different taxa listed as *P. pristinus*. Arrows identify the correct names in full. Pale gray histogram superimposes sum of local ranges identified as *P. fragilis pristinus* in 4 sections.

not reasonable to assume that local range ends correlate – the data cannot be forced to fit this ideal, especially for the taxon ranges supported by the most local observations and are commonly used to define zone boundaries. This is to be expected from a consideration of biogeography. But the unrealizable ideal is a tacit assumption of some uses of zonal biostratigraphy.

Knowledge of the most parsimonious arrangement of local ranges, as shown by these histograms, serves three purposes: to discover likely data errors; to seek geographic differences in the timing and duration of the local ranges and to assess the reliability of the composite range ends

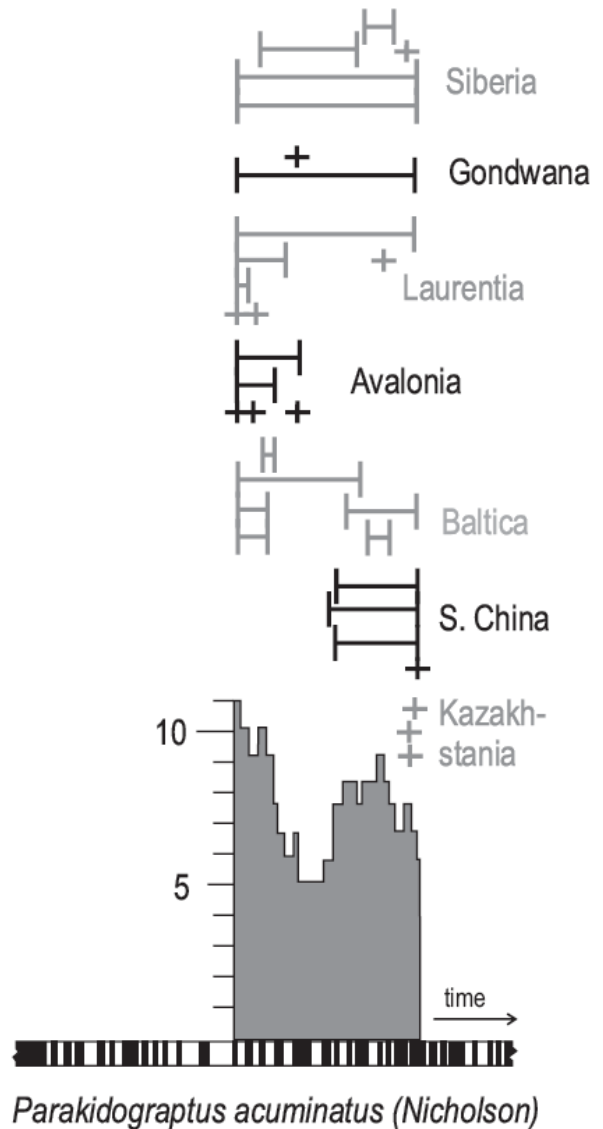
*Disjunct Ranges.* – Disjunct composite ranges are those that include intervals not covered by any local ranges (Fig. VI.7). This is not a feasible biological outcome. Perhaps the biogeographic range of the taxon has been inadequately sampled. Alternatively, a disjunct range may indicate misidentifications or regional differences in taxonomic practice.

One taxonomic problem leading to disjunct ranges arises when the same name appears in range charts for two different species within a genus. The name may



**Figure VI.7** Strong local range support for the composite range of *Akidograptus ascensus*, differentiated by paleocontinent. Symbols as in figure VI.5.

have initially been applied to one species and later the subspecies of another. Unaware of the two usages, perhaps because the two taxa have quite different ages, later writers treat the subspecies as a species. The result is one name for two species. One example (Fig. VI.6) has arisen in the genus *Pristiograptus*, a graptolite. *Pristiograptus pristinus* Pribyl 1940 and *Pristiograptus fragilis pristinus* Hutt 1975 have both been recorded in published range charts as *P. pristinus*. The inadvertent conflation of *Monograptus vulgaris* Wood 1900 and *Monograptus austerus vulgaris* Hutt 1975 was similarly revealed



**Figure VI.8:** Strong local range support for the composite range of *Parakidograptus acuminatus*, differentiated by paleocontinent. Symbols as in figure VI.5.



by a highly disjunct composite range.

*Geographic Support.* – Ideally, the support for the composite range includes many local sections with wide geographic distribution. If the taxon appears diachronously across different provinces, with different associates, this may be revealed by clusters of local ranges. The regional clusters compensate for local sampling deficiencies and yet different regions support different parts of the composite range (Figs. VI.7-8).

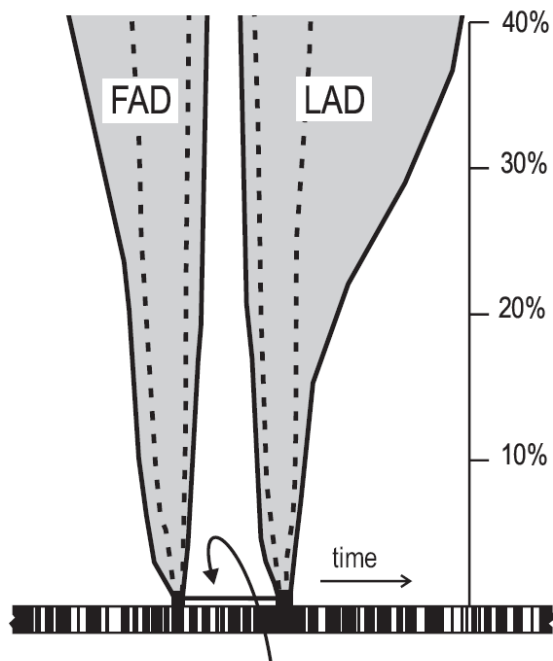
The patterns not only inform biogeographic reconstructions, they also enable inter-provincial correlations to improve upon literal correlation of observed range ends, and they make it possible to extract shorter composite ranges for correlation within provinces. These patterns will not emerge unless other elements of the biota differ between provinces or there are non-biologic events to constrain the correlations.

*Range End Uncertainty.* – The support histogram is another indicator of the relative uncertainty at the range ends. We may have more confidence in composite range ends if many local sections map into the terminal portions of the composite range. If the range end is determined by an outlier occurrence (Fig. VI.5), the source of that information may deserve re-examination. It might be preferable to prune the composite range back to the continuously supported portion.

### VI.5. Composite Range Ends

The previous paragraphs explained how to examine the support for a composite range end in the outcome of a single optimization run. More quantitative insight can be gained by combining the results of additional runs that are designed to reach the best-fit solution less directly. The goal is to discover other equally well-fit sequences and the many near-best-fit solutions. From these, CONOP maps out best-fit intervals and relaxed fit curves (Fig. VI.9).

Best-fit intervals are the more easily reached objective. They show the span of positions that an event may take in sequences that have the best-known fit. They require finding as many of the equally best-fit solutions as possible. Because the simulated annealing heuristic progressively reduces the scope of mutations that it accepts, the thoroughness of the



*Coronograptus cyphus cyphus* (Lapworth, 1876)

**Figure VI.9:** Composite range for *Coronograptus cyphus* (arrowed bar) with best-fit intervals (black rectangles at end of composite range bar) and U-shaped relaxed-fit curves (gray) for position of range-ends in composite time-line. Scale for relaxed fit curves shows misfit as percentage increase over best fit. Solid relaxed fit curve is for misfit measured by sum of range extensions in event levels; dashed relaxed fit curves is for additional implied coexistences. The composite range was optimized on a weight sum of these two measures. Horizontal “bar-code” at base is scaled composite section (Sadler et al., 2009) with event levels represented by black bars.

search always increases as it narrows down to the vicinity of the best-fit solution. Thus, a single run may provide a reasonable estimate of the best-fit interval. To fill out the interval more thoroughly, restart a search from the previous best outcome. If the tolerance for deleterious mutations is increased a little above its final value, the search will first move a short distance away from the optimal position on the misfit landscape and then re-try the final stages of the search. This is a good strategy, anyway, to allay worries that a marginally better solution was missed by the first run.



Larger areas may be searched around the optimum sequence by further increasing the tolerance for deleterious mutations and restarting runs with less efficient cooling rates and/or mutations. With large numbers of these restarts, it is possible to map out relaxed-fit curves (Sadler and Cooper, 1993). These funnel-shaped curves (Fig. VI.9) show how the best achievable misfit increases in timelines that place an event progressively farther outside its best-fit interval.

They indicate the rate of loss of parsimony that accompanies any subjective assertion that the “true” position of an event lies outside the best fit interval. For well-constrained events the relaxed fit funnels are narrow: the event remains confined to a narrow span of positions, even if significantly sub-optimal sequences are tolerated. These relatively “immobile” events are those known from many richly fossiliferous locations and with the same associated taxa. The relaxed-fit funnels of poorly constrained events, by contrast, widen rapidly. These events may be asserted to lie outside the best-fit interval with relatively little loss in parsimony. The relaxed fit curves are typically asymmetrical. The FAD and LAD curves flare more rapidly away from each other than toward each other. This happens because each is a constraint on the mobility of the other.

Where misfit is measured in stratigraphic thickness, natural limits to precision result from measurement error and sample spacing. Sample size influences the chances of finding taxa. It would seem reasonable not to fine-tune the optimization to a point that discriminates differences of fit more finely than the field data can justify. This line of reasoning suggests accepting both best-fit and near-best-fit solutions as a measure of the reliability of composite range ends.

During the trial-and-error process, misfit decreases and each range-end event can be imagined to be descending irregularly inside its relaxed fit funnel (gray in Fig. VI.9) as the solution slowly quenches. Thus, during a run the motion of the animated range chart becomes progressively less agitated and as the options for mutation are throttled by the narrowing cone and falling temperature.

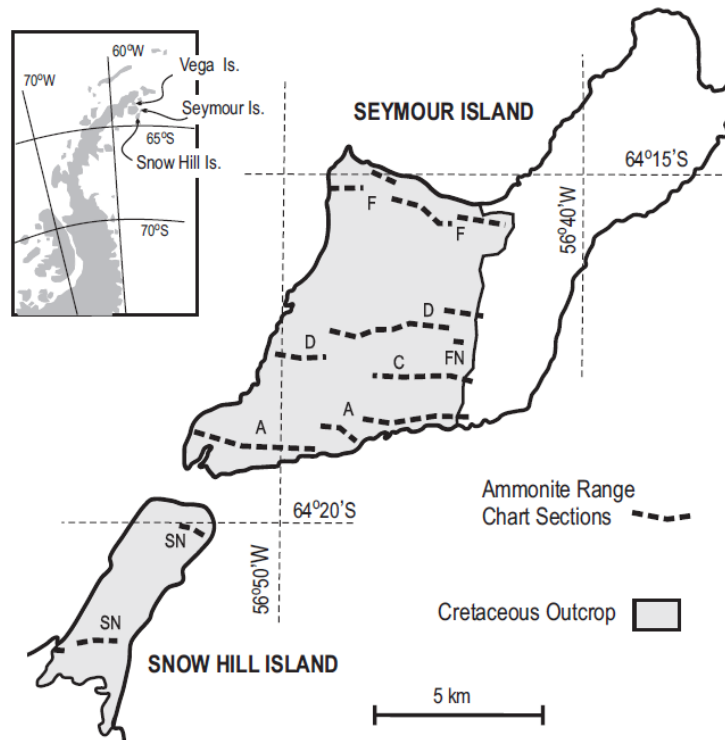
CONOP can be set up to continually repeat searches and keep a tally of the best fit discovered for each event in each possible position; i.e. collect a set of points that must lie inside or at the edge of the funnel. Remember, however, that heuristic searches are needed because exhaustive searches are not feasible. It is also not feasible to map exactly the entire relaxed fit funnel. An inadequate search will likely underestimate the funnel width and the feasibility falls dramatically as one tries to search larger areas of the misfit landscape to map up into the wider parts of the funnel. The more realistic goal is a good approximation of the lower, narrower parts of the funnel.

So far, this section has considered the range of time-line outcomes from a single data set and a single measure of fit between the data and the time-line. Another approach examines the range of outcomes that emerge when different measures of misfit are applied. The purpose may be to develop more conservative estimates of uncertainty (e.g. Sadler et al., 2009) or to tease out evidence of provinciality from the way it impacts different measures of misfit (e.g. Sadler and Sabado, 2009). This is one example of sensitivity analysis. Another example watches how the outcomes respond to the removal of individual taxa and localities.

CONOP also permits individual observed range ends to be weighted differently in the measurement of adjustments. As reasonable as this may seem, as a means to incorporate more expert judgments, it opens unmanageably many degrees of freedom. Very good solutions routinely emerge without any weight differentials, probably because many expert judgment are based on a knowledge of all the local ranges and associations; CONOP has this information (Kemple et al. 1995). Sadler and Kemple (1995) used the weights not to get a “better” time lines but to explore what subjective evaluations of the data might be needed to justify a particular preferred or expected outcome.

## VII. Latest Cretaceous Ammonites from Antarctica

A small case study in the use of CONOP9 for seriation and CONMAN9 for data management.



**Figure VII.1** Location of measured sections on Seymour Island, Antarctic Peninsula. Sections A, C, D, and F from Macellari (1986). Section FN (“Filo Negro”) from Elliot et al. (1994). Section SN (“Snow H”) from Pirrie et al. (1997). Section CL (Cape Lamb) is on Vega Island.

### VII.1 A Standard Data Set for Biostratigraphy

Macellari’s (1986) observed ranges for the latest Cretaceous ammonites from Seymour Island, Antarctica (Fig. VII.1) have become a standard data set for illustrating methods that place confidence intervals on taxon range end and the likely position of mass extinction horizons. The same data are amenable to graphic correlation and constrained optimization. The Macellari data have been augmented here with additional sections from the Antarctic Peninsula (Elliot et al., 1994; Pirrie et al., 1997; Cramer et al., 2004) and Chile, and Zinsmeister’s (2001) composite section for Seymour Island. The additional information in these papers includes two types of non-paleobiologic events –

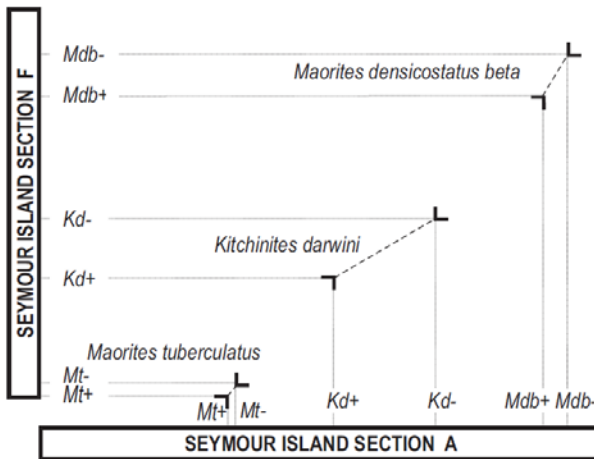
strontium-isotope estimates of geologic age and the stratigraphic position of the terminal-Cretaceous iridium anomaly. The extra sections also introduce a different suite of ammonite taxa.

The case study is part of a comprehensive late Cretaceous ammonite database in CONMAN9 with at least 201 sections, 605 taxa, and 58 other events. A complete run would generate 3789 local observations. Our part of it contains 12 sections, 49 taxa and 22 other events; it generates an input file with 275 lines – one for each locally observed event. We will examine small parts of the data in the format of graphic correlation. After this, the recommended exercises will require running the data set in CONOP to explore the

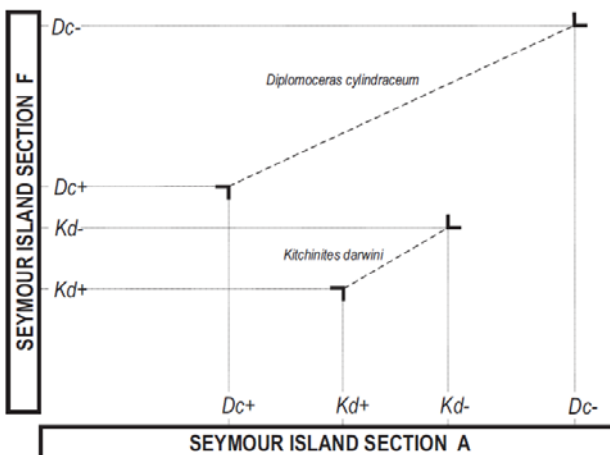
versatility of the program and generate screen graphics and output files.

## VII.2 Graphic Correlation

Consider first the graphic correlation of Macellari's sections A and F. In the manner of traditional biostratigraphy, taxa may be selected that appear and disappear in the same order in both sections. They



**Figure VII.2** Graphic correlation of the observed ranges of three taxa in two sections from Seymour Island. Range end abbreviations use the initials of the genus and species, suffixed with “plus” for first appearances and “minus” for last appearances.



**Figure VII.3** Graphic correlation of two more taxon ranges from sections A and F. These conflict in the sense that no line of correlation can be drawn to connect all four symbols without including a segment with negative slope (down to the right).

present no challenge for graphic correlation (Fig. VII.2). For this selection, a piece-wise linear line of correlation (LOC) can be drawn through all range end coordinates without violating the one critical property of all feasible LOCs:

*The LOC must be a continuous line with non-negative slope.*

Vertical and horizontal segments are permitted; they represent surfaces of hiatus in one or other of the sections. Changes in the slope of the LOC represent changes in the ratio of accumulation rate in the two sections.

**EXERCISE VII.2a:**  
Although a piece-wise linear LOC can be fit to all the points in figure VII.2, it is also permissible to draw a single straight line of correlation that comes close to all the points. In his original explanation of the method, Shaw (1964) preferred straight LOCs. Do this; then show how the range-end coordinates must be adjusted to the LOC. Remember that the symbols move so that one or both local ranges are extended. Sequence stratigraphers now routinely draw LOCs with horizontal or vertical segments at sequence boundary hiatuses.

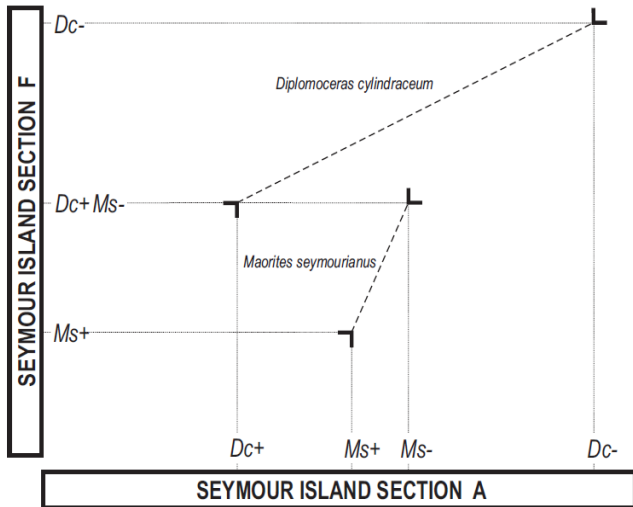
Other taxon range ends in these two sections are not amenable to LOCs that pass through all points, or even close to them all (Fig. VII.3-4). Figure VII.3 displays the case examined in the chapter about constrained optimization. One or more range ends must be adjusted to fit a line of correlation.

**EXERCISE VII.2b:**  
The chapter on constrained optimization determined the best-fit sequence for the four events in these two sections:

Dc+ . . . Kd+ . . . Kd- . . . Dc-

Draw an LOC on figure VII.3 that corresponds to this sequence. Remember that this sequence must describe the range-end sequence in both sections, *after* ranges are adjusted to the LOC.

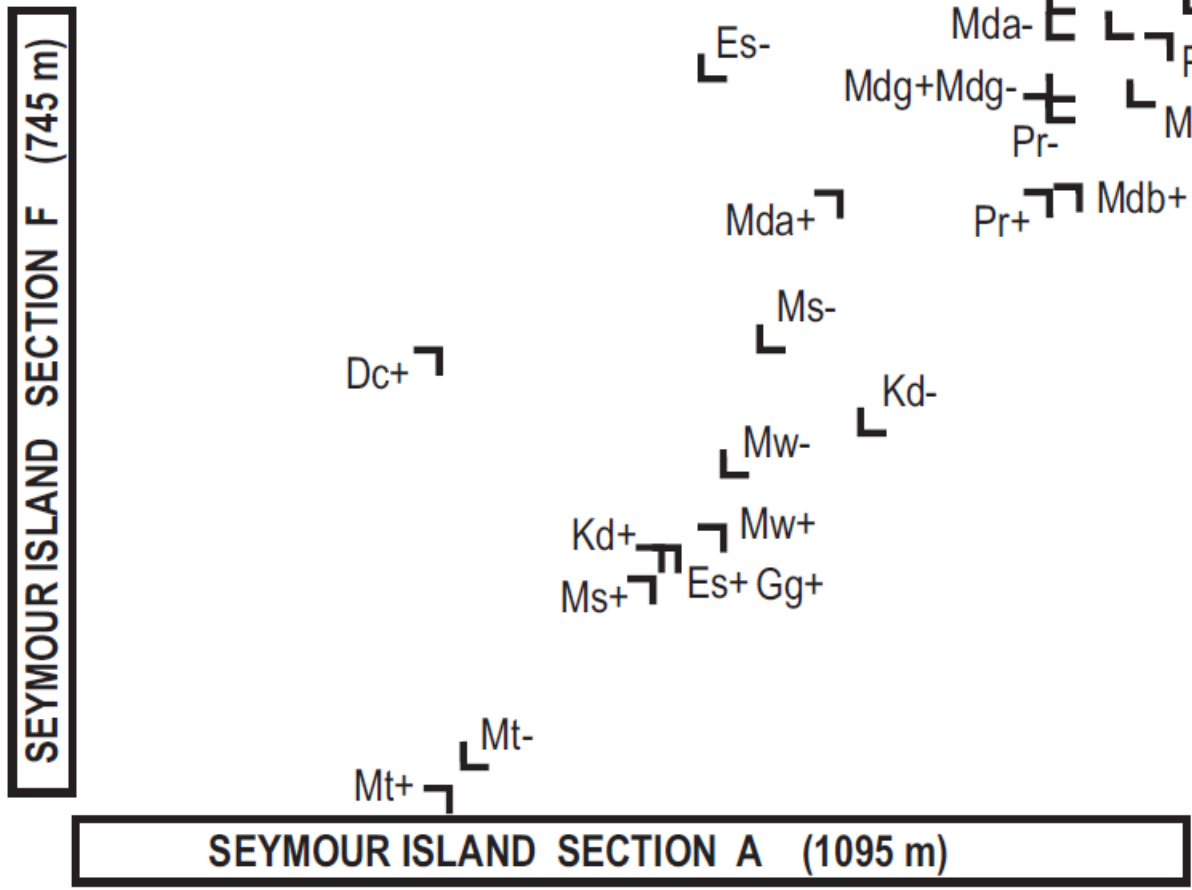
Note that range-ends coincident with the top or bottom limits of a section are not likely to be trustworthy. These range ends are artificial truncations. CONOP automatically discounts them because they accumulate no penalty for moving beyond the section limits; i.e. off the penalty scale.



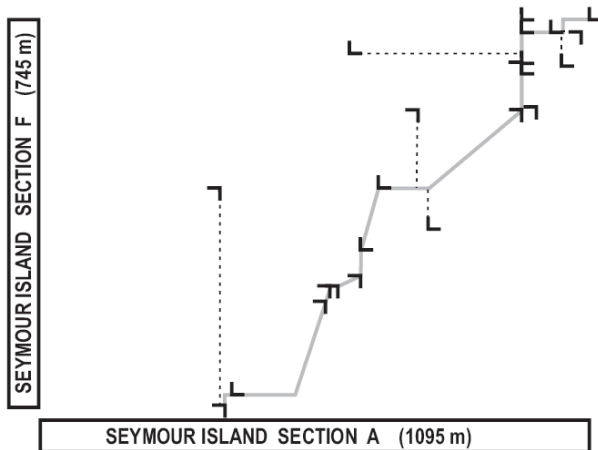
**Figure VII.4** Another conflicted pair of taxa from sections A and F.

Figure VII.4 depicts another pair of taxa with a similar conflict. Examination of Figures VII.2-4 indicates that *D. cylindraceum* is likely the least trustworthy observed range.

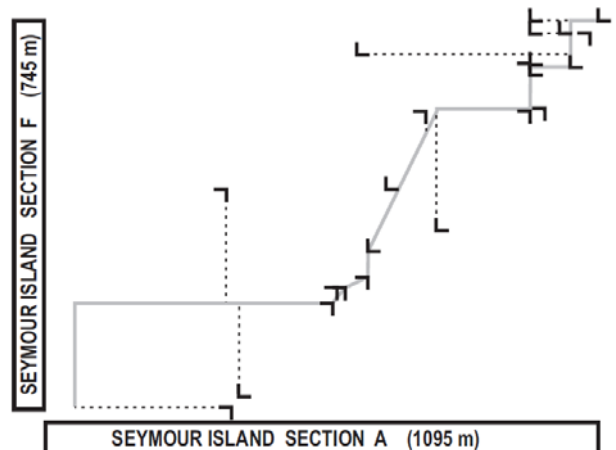
**EXERCISE VII.2c**  
 Attempt to draw a suitable LOC for the total data set depicted in figure V.III.5. It should come as close as possible to as many of the range-end coordinates as possible.  
 Indicate a zone or band across the diagram in which the true LOC most likely lies. Keep this uncertainty interval reasonably narrow.



**Figure VII.5** A graphical plot of all taxon range ends shared by sections A and F.



**Figure VII.6:** The LOC for sections A and F determined by constrained optimization using CONOP9 for a data set that included only these two sections.



**Figure VII.7** The LOC for sections A and F determined by constrained optimization using CONOP9 for a data set that included 12 sections from Antarctica and South America.

### VII.3 The CONOP Run-Time Animation

Compare your LOC in figure VII.5 with the two machine-generated LOCs in figures VII.6-7. We will attempt to reproduce these results using CONOP9. First, let us re-run the CONOP files used for figure VII.7. We can use CONMAN later to develop input files for the 2-section data set used in figure VII.6. During the CONOP run, the screen will resemble figures VII.8-9.

#### EXERCISE VII. 3

Examine a ready-made 21-section CONOP run for late Cretaceous ammonite ranges in Antarctica and South America.

3.1 Open the computer folder labeled 12x49x22

3.2 Run the file CONOP9.exe (and repeat as many times as necessary to examine all the features mentioned below.

3.3 Read the start-up screen then click Enter to start a run-time animation. The white lines are individual taxon range zones. The yellow ticks are other events. The red line shows the diminishing tolerance for bad mutations. The green squares show the progress of the smallest

misfit. The gray squares show the misfit after every accepted mutation. Notice the following during the animation

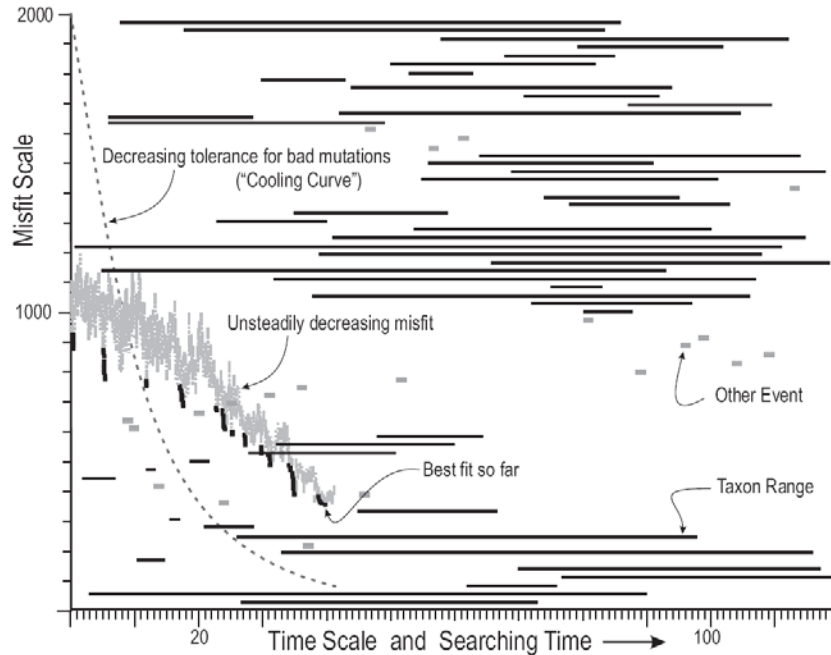
- \* the white ranges start long and finish shorter
- \* the white ranges change more often at first, then change less frequently
- \* the white ranges change by large amounts at first, then by smaller amounts
- \* the amplitude of the fluctuations in the grey curve become smaller and eventually disappear.
- \* all these progressive changes occur with the lowering of the red line

Discuss what this means in terms of mutations and the coarseness of the search.

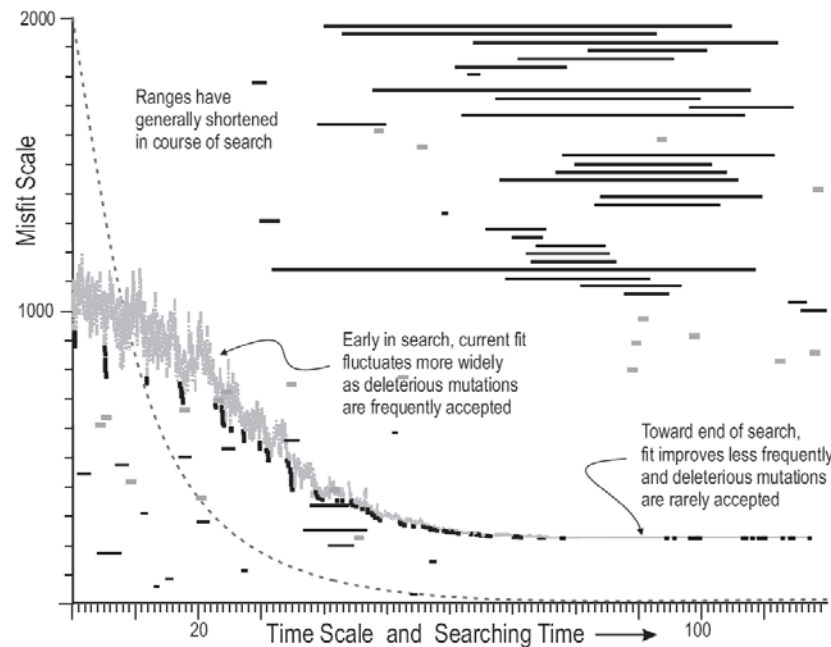
### VII.4 CONOP's Red "Temperature" Curve

During simulated annealing a "temperature" parameter is progressively lowered to reduce the tolerance for bad mutations of the time-line. The CONOP9.CFG file controls the form of this curve and the length of the run using four parameters:

- STEPS sets the number of cooling steps undertaken before the run stops



**Figure VII.8.** Components of the CONOP screen mid-way through search.



**Figure VII.9.** Components of CONOP screen at end of search

- **TRIALS** sets the number of mutations that will be tried at each step. Thus the total number of mutations is the product of steps times trials.
- **STARTTEMP** determines the initial temperature.
- **RATIO** determines the amount by which the temperature is lowered at each step.

#### EXERCISE VII. 4

Change the cooling rate for the simulated annealing

4.1 In any text editor, open the file CONOP9.CFG in the folder labeled 12x49x22.

4.2 Change the combination of steps, trials, ratio and starting temperature to make a less thorough but more rapid search. There are many ways to do this: a lower starting temperature, a smaller ratio, fewer steps, fewer trials, or combinations of these changes.

4.3 Re-run the search and note the impact on the trajectory of the search and the solution.

4.4 Now adjust the parameters to achieve the opposite effect – a very conservative run that is more likely to find the best solution. Do this in the opposite way: a higher starting temperature, a larger ratio (but less than 1.00!), more steps, more trials, or combinations of these changes.

4.5 Re-run the search and note the impact on the trajectory of the search and the solution (open the file `outmain.txt` or whatever file-name appears after the `UNLOADMAIN` parameter in the CFG file).

#### VII.5 Mutating the Time-Line

CONOP offers three mutation options

`HOODSIZE= 'BIG'` Moves one event at random to a new position. If the random selection violates any constraints, it is immediately abandoned and another is tried. This is the default mutation and the only effective option for most data sets.

`HOODSIZE= 'SML'` (Small) Moves one event at random to the next position in sequence. It requires many more mutations and is not efficient for most data sets.

`HOODSIZE= 'DBL'` (Double) Moves swaps two events chosen at random. It makes bigger changes to the sequence but as the search progresses it becomes more difficult to find two events that can switch positions without violating a constraint

Although the “small” mutation is rarely effective, it provides a motive for adjusting the cooling curve.

#### EXERCISE VII.5:

Change the mutation option

5.1 In any text editor, open the file CONOP9.CFG in the folder labeled 12x49x22.

5.2 Change the `HOODSIZE=` parameter (neighborhood size) to `'SML'`. Save and close the CFG file then re-run CONOP.

5.3 Notice the changes in the animation and then alter the cooling rate as necessary to reach a good solution.

#### VII.6 The Measure of Misfit

CONOP offers several significantly different options for measuring the misfit between the time-line and the data. This is the objective function that the optimization will minimize. It is set by the `PENALTY=` line in the CFG file. Here are three of the more useful settings.

`PENALTY= 'INTERVAL'` measures the sum of all range extensions in stratigraphic rock thickness. This setting favors the sequence of events seen in the thickest sections. It most resembles graphic correlation, but is not recommended.

`PENALTY= 'LEVEL'` measures the sum of all range extensions in terms of event levels. It favors the sequences found in the most taxon rich and intensely sampled sections. It is recommended as the default setting.

`PENALTY= 'SEQUEL'` does not measure range extensions. It counts the number of FAD-before-LAD event pairs seen in the time line and subtracts from this total the number that have been observed in a real stratigraphic section. It resembles Alroy’s “appearance event ordination” and is faster than the other two options.

Sequel and level differ dramatically in their treatment of faunal dissimilarity (Sadler and Sabado, 2009). Given a data set that contains faunas that share no taxa and could lead to separate time lines, `LEVEL` treats them as coeval provinces and arbitrarily interleaves the two partial time lines in parallel. `SEQUEL` separates the two and stacks them in series,



as if the faunas differed by reason of age rather than provinciality.

**EXERCISE VII.6:**

Change the mutation option

6.1 In any text editor, open the file CONOP9.CFG in the folder labeled 12x49x22.

6.2 Change the "PENALTY=" parameter (objective function or cost) to 'SEQUEL'. Save and close the CFG file then re-run CONOP.

6.3 Notice the changes in the animation and then alter the cooling rate if necessary to reach a good solution.

6.4 Notice that the final value of the misfit function is different and no longer comparable.

**VII.7 Adding Composite Time-Lines to the CONMAN9 database as New Sections**

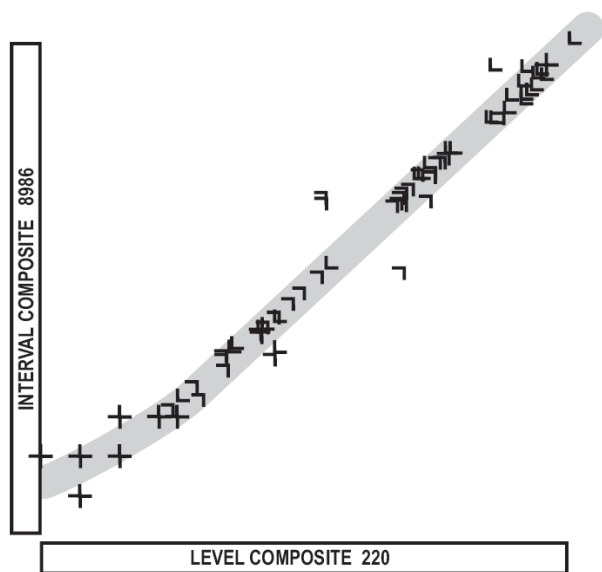
At the end of a CONOP run, the menus offer many graphical analyses of the outcome and the opportunity to write additional text-files to disk. One of the text-file options writes out the best-fit sequence in the

CONMAN format as if it were a stratigraphic section. This enables the outcomes of different runs to be compared, using the CONMAN9 analysis menu. The most insightful of these lists the shared taxa for selected pairs of sections and then plots a graphic correlation diagram for them.

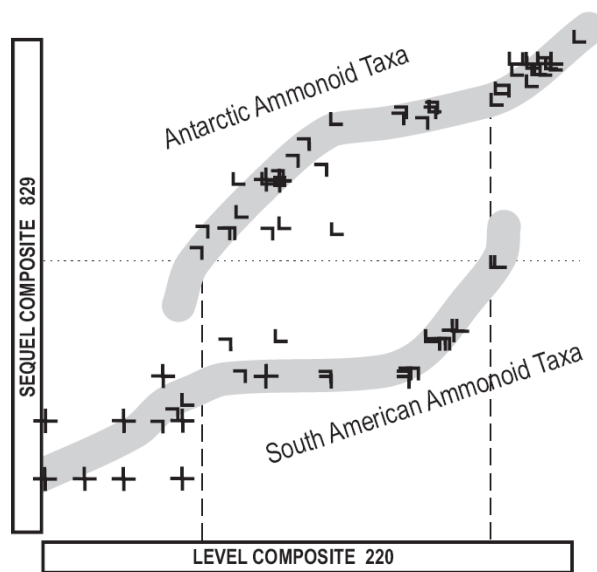
**VII.8 Cross-Plots of Two Best-Fit Sequences**

Cross-plots are the best ways to examine differences between two best-fit sequences from the same data. These might have been run with the same settings in the configuration file. They will then differ only due to random chance and the cross-plot should show the LOC as a more or less diffuse band (Sadler and Sabado, 2009) along the main diagonal of the plot. The random differences will be greater if the runs were too short or too fast and did not find any of the optimal fits. If the fits are both optimized, the differences will be greater for data sets with more internal discrepancies between sections.

Cross plots for the same data but with different objective functions (PENALTY parameter) settings may both be optimized, but to different criteria. For a very rich and minimally under-determined data set, the differences may be slight. Often they increase toward



**Figure VII.10** A cross-plot of two very similar best fit sequences for the 12x49x22 data set of late Cretaceous southern hemisphere ammonite sections.



**Figure VII.11** Cross plot of two very dissimilar best-fit solutions for the 12x49x22 data set of late Cretaceous southern hemisphere ammonite sections.

the ends of the LOC where constraint is typically compromised by the limits of the measured sections.

Differences between the LEVEL and SEQUEL objective functions will be most extreme if the data include runs of faunas that have no taxa in common. These dissimilarities might have arisen because the data set includes two groups of sections of quite different geologic age. Alternatively, the two groups might be contemporaneous biotic provinces. In the absence of information to confirm one or other cause (overlapping taxa, longer ranging sections and dated events), the LEVEL and INTERVAL functions treat the two groups as if they were from distinct geographic or ecologic provinces of the same age. This is simply a consequence of minimizing range extensions. The two functions may generate very similar outcomes (Fig. VII.10). The SEQUEL function minimizes the number of implied coexistences that have not been observed in the field. As a result it tends to separate the two groups in time, so that their included assemblages of taxa do not coexist.

When the LEVEL and SEQUEL outcomes are plotted against one another (Fig. VII.11), the LOC develops two distinct branches in the interval occupied by the dissimilar faunas (Sadler and Sabado, 2009). This occurs in the ammonite data set because it mixes faunas from Antarctica with rather different assemblages from Chile. A similar branched LOC would have arisen if we had added a few early Cretaceous ammonite-bearing sections to the data set. Like human experts, CONOP cannot reasonably correlate sections with nothing in common. Unlike the human operator, the computer might try – if asked!

#### EXERCISE VII.7

Moving best-fit sequences from CONOP9 to the CONMAN9 database.

7.1 Re-run the 12x49x22 data set with the `PENALTY= 'LEVEL'` option. At the end of the run go to the Text-File menu and choose the “with guessed abbreviations” options for writing out the best-fit sequence in CONMAN format.

7.2 The file will be written to the folder with the input files and named “CONOPCOMP.SEC” Rename the file to something more distinct; e.g. “12x49x22-Level.sec” and copy it to the folder with the CONMAN ammonite database.

7.3 Repeat the exercise with the `PENALTY= 'SEQUEL'` option and ensure that the second file has a different name from that generated for the LEVEL function.

7.4 Run the CONMAN9.EXE program. In the “Append” menu, choose the “New Sections” options and enter information for the two files migrated from CONOP. Make sure that the filenames match exactly the migrated files.

#### EXERCISE VII.8

Cross-plotting two best-fit sequences from the same data set.

8.1 Complete exercise VII.7 and remain in the main CONMAN9 dialog box.

8.2 From the “Analysis Menu” choose the “– list for section pairs” option under “SHARED TAXA.” In the dialog box, select the two new composite sections entered during exercise VII.7 and observe the list of shared taxa and events. The “Close and View Graph” button will lead to a graphic correlation plot of the two sections. Hit {Esc} two or three times to exit the graph.

8.3 Look for cross-plots that reproduce Figures VII.10-11.

8.4 Examine plots of individual sections against the composites. Those used to generate the composites will share all their taxa with the composite; others may share few or none.

8.5 For other CONOP runs of the same data set, try importing and comparing the composite sequences.

8.6 To see written listings of the composite time-lines and analyses of the fit, open the file `outmain.txt` or whatever file-name appears after the `UNLOADMAIN` parameter in the CFG file.

### EXERCISE VII.9

Duplicating figure VII.6. This will require exporting a new set of CONOP input files from CONMAN – one that includes only two sections. Refer to the “Starting” manual as necessary.

9.1 Reopen the main CONMAN dialog, click the “Export” menu and choose “- select section individually.”

9.2 In the new dialog box add Seymour F and Seymour A to the right hand column; then clip accept.

9.3 Return to the Export menu and click the option to select all taxa that occur in at least one section. Other options for taxon selection lead to smaller data sets. There follow options to omit taxa in various forms of open nomenclature. Leave the default settings and click buttons to proceed and, after the progress bars stop, to exit.

9.4 Returning to the Export menu each time, select all other events that occur in at least one section, write the CONOP9 input files, and write the CONOP9.CFG file. For the CFG file, choose no-editing, a fixed cooling schedule, and the animated range chart.

9.5 In the CONMAN database folder, list files in order of last modification and select conop9.cfg, loadfile.dat, prefile.dis, sections.sct and events.evt. Move these files to a new folder.

9.6 Add a copy of CONOP9.EXE to the new folder and run the program. (The program file is not large, so it is always simplest to keep a separate copy with each run and keep each run in a different folder.)

9.7 When the animated run has finished, proceed to the Graphical Output menu and select “2 section LOC”

9.8 Change the X and Y axes of the cross-plots by typing Y or X followed by the plus (+) or minus (-) keys. The axes will display the two sections in the run and up to ten versions of the composite sequence, each with the same order but a different scaling (spacing)

9.9 Striking the + or – keys without a preceding X or Y will change the labeled event on the graph. Exit the graphic by striking the {Esc} key 2 or 3 times.

For more about the post-run options in CONOP9, either experiment with the menus or read the Users’ Guide and Reference Manual!

## VII.9 CONOP9 Input files for the Antarctic Examples

---

### CONOP SECTION DICTIONARY FILE

There is one row for each section. The rows have five fields:

- section number,
- section abbreviation for use in some graphics
- order of sections in fence diagrams
- section name
- “1” to indicate that the section has a thickness scale for use in scaling the composite

*This file is optional. If it is omitted, CONOP simply refers to sections by number.*

1	'SyA'	1	'Seymour Island A'	1
2	'SyC'	2	'Seymour Island C'	1
3	'SyD'	3	'Seymour Island D'	1
4	'SyF'	4	'Seymour Island F'	1
5	'Fln'	5	'Seymour Filo Negro'	1
6	'Sy2'	6	'Seymour Island C2'	1
7	'Qrq'	7	'Quiriquina Island'	1
8	'O97'	8	'Obrovich Time 93/75'	0
9	'Lmb'	9	'Cape Lamb'	1
10	'Snw'	10	'Snow Hill Island'	1
11	'Xch'	11	'Chilean Composite'	1
12	'SyZ'	12	'Seymour Isl. Z composit'	1

---

### CONOP EVENT DICTIONARY FILE

There is one row for each event. The rows have three fields:

- event number,
- event abbreviation or the code from an external database
- event name

*This file is optional. If it is omitted, CONOP simply refers to events by number.*

1	'1051'	'Diplomoceras cylindraceum (Defrance) 1816'
2	'1183'	'Kitchinites darwini (Steinmann) 1895'
3	'1184'	'Maorites densicostatus alpha (Killian & Reboul)'
4	'1186'	'Maorites densicostatus gamma (Killian & Reboul)'
5	'1185'	'Maorites densicostatus beta (Killian & Reboul)'
6	'1187'	'Maorites seymourianus (Killian & Reboul) 1909'
7	'1188'	'Maorites wedelliensis Macellari 1986'

8 '1189' 'Gunnarites bhavaniformis (Killian & Reboul) 1909'  
9 '1190' 'Grossouvrites gemmatus (Hupe) 1854'  
10 '1191' 'Pachydiscus riccardi Macellari 1986'  
11 '1192' 'Pachydiscus ultimus Macellari 1986'  
12 '1193' 'Eutrephoceras simile Spath'  
13 '1194' 'Maorites tuberculatus Howarth 1958'  
14 '0117' 'Hasl'  
15 '0532' 'Strontium 65.8'  
16 '0533' 'Strontium 67.5'  
17 '1195' 'Diplomoceras cylindraceum maximum Olivero & Zins. 1989'  
18 '1198' 'Anagaudryceras seymouriense Macellari 1986'  
19 '1196' 'Pseudophyllites loryi (Killian & Reboul) 1909'  
20 '1197' 'Kitchinites laurae Macellari 1986'  
21 '0100' 'KT iridium/tektites'  
22 '1107' 'Zelandites varuna (Forbes) 1846'  
23 '1045' 'Gaudryceras kayei (Forbes) 1846'  
24 '1200' 'Pachydiscus ootacodensis (Stoliczka) 1865'  
25 '1275' 'Baculites vincentei'  
26 '1276' 'Anapachydiscus fresvillensis quirquinae Stinnesbeck 1986'  
27 '1277' 'Baculites hunickeni'  
28 '1278' 'Phylloceras\_(H.) inflatum'  
29 '1279' 'Eutrephoceras subplicatum'  
30 '1280' 'Pachydiscus\_(P.) jacquoti chilensis'  
31 '1130' 'Eubaculites lyelli (d Orbigny) 1847'  
32 '1281' 'Kossmaticeras\_(N.) erbeni'  
33 '1100' 'Hypophylloceras hetonaiense (Matsumoto) 1942'  
34 '1293' 'Neophylloceras ramosum (Meek) 1857'  
35 '1285' 'Hoploscaphites quiriquiniensis'  
36 '1044' 'Neophylloceras surya (Forbes) 1846'  
37 '0114' 'C30n~C29r'  
38 '0529' 'Obrad93 10\_81.71 S.hippoc'  
39 '0518' 'Obrad93 1a\_65.35'  
40 '0519' 'Obrad93 1b\_65.43'  
41 '0520' 'Obrad93 2a\_64.97'  
42 '0521' 'Obrad93 2b\_65.19'  
43 '0522' 'Obrad93 3\_66.8 KneeHills'  
44 '0523' 'Obrad93 4\_69.42 B.clinolo'  
45 '0524' 'Obrad93 5\_73.35 B.compres'  
46 '0525' 'Obrad93 6\_74.76 E.jenneyi'  
47 '0526' 'Obrad93 7\_75.37 G.calcara'  
48 '0527' 'Obrad93 8\_75.89 B.scotti'  
49 '0528' 'Obrad93 9\_80.54 Ardmore'  
50 '0102' 'Ardmore Bentonite'  
51 '1059' 'Baculites clinolobatus Elias 1933'  
52 '1249' 'Baculites compressus Say 1820'  
53 '1073' 'Exiteloceras jenneyi (Whitfield) 1877'  
54 '1078' 'Baculites scotti Cobban 1958'  
55 '1088' 'Baculites obtusus Meek 1876'  
56 '1393' 'Scaphites\_(S.) hippocrepis (DeKay) 1828'  
57 '0506' 'Obrad88 8a\_79.2'  
58 '0501' 'Obrad88 3\_70.1 B.grandis'  
59 '0505' 'Obrad88 7\_75.5'

```

60 '1062' 'Baculites grandis Hall & Meek 1855'
61 '1087' 'Baculites mclearnii Landes 1940'
62 '1077' 'Didymoceras nebrascense (Meek & Hayden) 1856'
63 '1545' 'Gunnarites antarcticus (Weller)'
64 '0531' 'Strontium 71.0 +/-0.2'
65 '1546' 'Jacobites crofti'
66 '1660' 'Neograhamites taylori Spath'
67 '1369' 'Gunnarites kalika (Stoliczka) 1865'
68 '1662' 'Kossmaticeras bhavani'
69 '1663' 'Kossmaticeras guentheri'
70 '1664' 'Kossmaticeras theobaldianus'
71 '1669' 'Dimitobelus seymouriensis'

```

---

## CONOP INPUT DATA FILE

The file has one row for each locally observed range end. Each row has eight fields:

- the event number (integer from event dictionary)
- the event code (integer code)
  - 2 = unpaired LAD
  - 1 = unpaired FAD
  - 1 = paired FAD
  - 2 = paired LAD
  - 3 = mid range event
  - 4 = marker bed
  - 5 = numerical age
  - 11 = paired limit of uncertainty interval
  - 12 = paired limit of uncertainty interval
- the section number (integer from section dictionary)
- the level in the section measured in thickness above a datum (decimal;)
- the level in the section counted from the oldest (integer)
- the freedom to adjust the level (integer code)
  - 0 = none
  - 1 = down only
  - 2 = up only
  - 3 = up or down
- the weight for upward adjustment (decimal)
- the weight for downward adjustment (decimal)

**Bold** entries indicate that a comment has been added and would not appear in the actual entry file

The rows in this file must be sorted on the first three columns.

```

1    1    1    400.0000    3    1    1.000    1.000    {type 1 events are FADs}
1    1    2    50.0000    3    1    1.000    1.000

```

1	1	3	255.0000	1	1	1.000	1.000
1	1	4	460.0000	8	1	1.000	1.000
1	1	9	100.0000	2	1	1.000	1.000
1	1	10	-330.0000	2	1	1.000	1.000
1	<b>2</b>	1	1060.0000	17	2	1.000	1.000
1	2	2	190.0000	11	2	1.000	1.000
1	2	3	820.0000	14	2	1.000	1.000
1	2	4	760.0000	18	2	1.000	1.000
1	2	9	180.0000	4	2	1.000	1.000
1	2	10	-100.0000	6	2	1.000	1.000
2	1	1	620.0000	8	1	1.000	1.000
2	1	2	-200.0000	1	1	1.000	1.000
2	1	3	375.0000	2	1	1.000	1.000
2	1	4	270.0000	4	1	1.000	1.000
2	1	7	17.0000	7	1	1.000	1.000
2	1	10	-390.0000	1	1	1.000	1.000
2	1	12	408.9700	2	1	1.000	1.000
2	2	1	815.0000	14	2	1.000	1.000
2	2	2	-200.0000	1	2	1.000	1.000
2	2	3	530.0000	5	2	1.000	1.000
2	2	4	380.0000	7	2	1.000	1.000
2	2	7	44.0000	11	2	1.000	1.000
2	2	10	-225.0000	4	2	1.000	1.000
2	2	12	884.9000	12	2	1.000	1.000
3	1	1	795.0000	13	1	1.000	1.000
3	1	2	10.0000	2	1	1.000	1.000
3	1	3	650.0000	8	1	1.000	1.000
3	1	4	610.0000	9	1	1.000	1.000
3	1	5	0.0000	1	1	1.000	1.000
3	1	6	220.0000	1	1	1.000	1.000
3	1	11	7.0000	7	1	1.000	1.000
3	1	12	495.0800	5	1	1.000	1.000
3	2	1	1000.0000	15	2	1.000	1.000
3	2	2	228.0000	13	2	1.000	1.000
3	2	3	895.0000	19	2	1.000	1.000
3	2	4	760.0000	18	2	1.000	1.000
3	2	5	6.0000	2	2	1.000	1.000
3	2	6	235.0000	3	2	1.000	1.000
3	2	11	7.0000	7	2	1.000	1.000
3	2	12	1054.3400	20	2	1.000	1.000
4	1	1	1000.0000	15	1	1.000	1.000
4	1	2	173.0000	9	1	1.000	1.000
4	1	3	690.0000	9	1	1.000	1.000
4	1	4	700.0000	16	1	1.000	1.000
4	2	1	1000.0000	15	2	1.000	1.000
4	2	2	173.0000	9	2	1.000	1.000
4	2	3	855.0000	16	2	1.000	1.000
4	2	4	700.0000	16	2	1.000	1.000
5	1	1	1030.0000	16	1	1.000	1.000
5	1	2	100.0000	7	1	1.000	1.000
5	1	3	750.0000	11	1	1.000	1.000
5	1	4	615.0000	10	1	1.000	1.000
5	1	7	5.0000	3	1	1.000	1.000
5	2	1	1080.0000	18	2	1.000	1.000
5	2	2	100.0000	7	2	1.000	1.000
5	2	3	765.0000	12	2	1.000	1.000
5	2	4	695.0000	15	2	1.000	1.000

{type 2 events are LADs}



5	2	7	44.0000	11	2	1.000	1.000
6	1	1	610.0000	7	1	1.000	1.000
6	1	2	-200.0000	1	1	1.000	1.000
6	1	3	380.0000	3	1	1.000	1.000
6	1	4	240.0000	3	1	1.000	1.000
6	1	10	-100.0000	6	1	1.000	1.000
6	1	12	495.0800	5	1	1.000	1.000
6	2	1	715.0000	12	2	1.000	1.000
6	2	2	-200.0000	1	2	1.000	1.000
6	2	3	535.0000	6	2	1.000	1.000
6	2	4	460.0000	8	2	1.000	1.000
6	2	10	-100.0000	6	2	1.000	1.000
6	2	12	821.7100	9	2	1.000	1.000
7	1	1	680.0000	11	1	1.000	1.000
7	1	2	-200.0000	1	1	1.000	1.000
7	1	3	535.0000	6	1	1.000	1.000
7	1	4	290.0000	5	1	1.000	1.000
7	1	12	708.5300	6	1	1.000	1.000
7	2	1	680.0000	11	2	1.000	1.000
7	2	2	-200.0000	1	2	1.000	1.000
7	2	3	535.0000	6	2	1.000	1.000
7	2	4	340.0000	6	2	1.000	1.000
7	2	12	708.5300	6	2	1.000	1.000
8	1	1	40.0000	1	1	1.000	1.000
8	2	1	40.0000	1	2	1.000	1.000
9	1	1	635.0000	9	1	1.000	1.000
9	1	2	70.0000	4	1	1.000	1.000
9	1	3	510.0000	4	1	1.000	1.000
9	1	4	270.0000	4	1	1.000	1.000
9	1	6	220.0000	1	1	1.000	1.000
9	1	7	5.0000	3	1	1.000	1.000
9	1	10	-225.0000	4	1	1.000	1.000
9	1	12	430.9400	3	1	1.000	1.000
9	2	1	1000.0000	15	2	1.000	1.000
9	2	2	215.0000	12	2	1.000	1.000
9	2	3	860.0000	17	2	1.000	1.000
9	2	4	785.0000	19	2	1.000	1.000
9	2	6	235.0000	3	2	1.000	1.000
9	2	7	16.0000	6	2	1.000	1.000
9	2	10	-100.0000	6	2	1.000	1.000
9	2	12	1054.3400	20	2	1.000	1.000
10	1	1	1000.0000	15	1	1.000	1.000
10	1	2	70.0000	4	1	1.000	1.000
10	1	3	720.0000	10	1	1.000	1.000
10	1	4	610.0000	9	1	1.000	1.000
10	1	12	760.2900	7	1	1.000	1.000
10	2	1	1000.0000	15	2	1.000	1.000
10	2	2	180.0000	10	2	1.000	1.000
10	2	3	830.0000	15	2	1.000	1.000
10	2	4	680.0000	14	2	1.000	1.000
10	2	12	915.3000	13	2	1.000	1.000
11	1	1	1120.0000	19	1	1.000	1.000
11	1	2	190.0000	11	1	1.000	1.000
11	1	3	820.0000	14	1	1.000	1.000
11	1	4	760.0000	18	1	1.000	1.000
11	1	6	220.0000	1	1	1.000	1.000
11	1	12	974.7500	14	1	1.000	1.000

11	2	1	1135.0000	20	2	1.000	1.000
11	2	2	215.0000	12	2	1.000	1.000
11	2	3	865.0000	18	2	1.000	1.000
11	2	4	785.0000	19	2	1.000	1.000
11	2	6	252.0000	5	2	1.000	1.000
11	2	12	1042.0900	16	2	1.000	1.000
12	1	1	635.0000	9	1	1.000	1.000
12	1	2	215.0000	12	1	1.000	1.000
12	1	3	530.0000	5	1	1.000	1.000
12	1	4	270.0000	4	1	1.000	1.000
12	2	1	658.0000	10	2	1.000	1.000
12	2	2	215.0000	12	2	1.000	1.000
12	2	3	895.0000	19	2	1.000	1.000
12	2	4	720.0000	17	2	1.000	1.000
13	1	1	410.0000	4	1	1.000	1.000
13	1	4	40.0000	1	1	1.000	1.000
13	1	10	-330.0000	2	1	1.000	1.000
13	1	12	344.3900	1	1	1.000	1.000
13	2	1	425.0000	5	2	1.000	1.000
13	2	4	60.0000	2	2	1.000	1.000
13	2	10	-100.0000	6	2	1.000	1.000
13	2	12	430.9400	3	2	1.000	1.000
14	<b>4</b>	1	100.0000	2	0	1.000	1.000
14	4	10	-135.0000	5	0	1.000	1.000
15	5	1	1000.0000	15	0	65.400	66.200
16	5	1	550.0000	6	0	67.100	67.900
17	1	2	190.0000	11	1	1.000	1.000
17	1	3	855.0000	16	1	1.000	1.000
17	1	6	220.0000	1	1	1.000	1.000
17	1	12	788.6200	8	1	1.000	1.000
17	2	2	215.0000	12	2	1.000	1.000
17	2	3	895.0000	19	2	1.000	1.000
17	2	6	246.0000	4	2	1.000	1.000
17	2	12	1054.3400	20	2	1.000	1.000
18	1	2	90.0000	6	1	1.000	1.000
18	1	3	650.0000	8	1	1.000	1.000
18	1	4	760.0000	18	1	1.000	1.000
18	1	12	825.1400	10	1	1.000	1.000
18	2	2	145.0000	8	2	1.000	1.000
18	2	3	830.0000	15	2	1.000	1.000
18	2	4	760.0000	18	2	1.000	1.000
18	2	12	1020.6100	15	2	1.000	1.000
19	1	2	180.0000	10	1	1.000	1.000
19	1	4	640.0000	12	1	1.000	1.000
19	1	7	17.0000	7	1	1.000	1.000
19	1	12	788.6200	8	1	1.000	1.000
19	2	2	180.0000	10	2	1.000	1.000
19	2	4	670.0000	13	2	1.000	1.000
19	2	7	44.0000	11	2	1.000	1.000
19	2	12	1054.3400	20	2	1.000	1.000
20	1	2	70.0000	4	1	1.000	1.000
20	1	3	645.0000	7	1	1.000	1.000
20	1	4	615.0000	10	1	1.000	1.000
20	1	6	220.0000	1	1	1.000	1.000
20	1	12	869.1600	11	1	1.000	1.000
20	2	2	80.0000	5	2	1.000	1.000

{type 4 events are  
marker beds}

20	2	3	895.0000	19	2	1.000	1.000
20	2	4	620.0000	11	2	1.000	1.000
20	2	6	232.0000	2	2	1.000	1.000
20	2	12	1054.2400	19	2	1.000	1.000
21	4	2	230.3000	14	0	1.000	1.000
21	4	5	9.4000	3	0	1.000	1.000
21	4	8	-64.9700	35	0	1.000	1.000
22	1	2	228.0000	13	1	1.000	1.000
22	1	3	895.0000	19	1	1.000	1.000
22	1	7	17.0000	7	1	1.000	1.000
22	1	12	1043.5699	17	1	1.000	1.000
22	2	2	228.0000	13	2	1.000	1.000
22	2	3	895.0000	19	2	1.000	1.000
22	2	7	44.0000	11	2	1.000	1.000
22	2	12	1049.7800	18	2	1.000	1.000
23	1	3	785.0000	13	1	1.000	1.000
23	1	7	17.0000	7	1	1.000	1.000
23	2	3	855.0000	16	2	1.000	1.000
23	2	7	44.0000	11	2	1.000	1.000
24	1	3	380.0000	3	1	1.000	1.000
24	2	3	380.0000	3	2	1.000	1.000
25	1	7	1.0000	1	1	1.000	1.000
25	2	7	5.0000	3	2	1.000	1.000
26	1	7	5.0000	3	1	1.000	1.000
26	2	7	16.0000	6	2	1.000	1.000
27	1	7	5.0000	3	1	1.000	1.000
27	2	7	16.0000	6	2	1.000	1.000
28	1	7	11.0000	5	1	1.000	1.000
28	2	7	19.0000	8	2	1.000	1.000
29	1	7	5.0000	3	1	1.000	1.000
29	2	7	20.0000	9	2	1.000	1.000
30	1	7	6.0000	4	1	1.000	1.000
30	2	7	20.0000	9	2	1.000	1.000
31	1	7	4.0000	2	1	1.000	1.000
31	2	7	44.0000	11	2	1.000	1.000
32	1	7	5.0000	3	1	1.000	1.000
32	2	7	44.0000	11	2	1.000	1.000
33	1	7	17.0000	7	1	1.000	1.000
33	2	7	44.0000	11	2	1.000	1.000
34	1	7	20.0000	9	1	1.000	1.000
34	2	7	44.0000	11	2	1.000	1.000
35	1	7	44.0000	11	1	1.000	1.000
35	2	7	46.0000	12	2	1.000	1.000
36	1	7	44.0000	11	1	1.000	1.000
36	2	7	46.0000	12	2	1.000	1.000
37	4	7	38.0000	10	0	1.000	1.000
38	5	8	-81.7100	2	0	<b>79.990</b>	<b>81.090</b>
39	5	8	-65.3500	33	0	65.040	65.660
40	5	8	-65.4300	32	0	65.220	65.640
41	5	8	-64.9700	35	0	64.680	65.260
42	5	8	-65.1900	34	0	64.740	65.640
43	5	8	-66.8000	31	0	65.700	67.900
44	5	8	-69.4200	30	0	69.790	69.050
45	5	8	-73.3500	23	0	72.960	73.740
46	5	8	-74.7600	20	0	74.310	75.210
47	5	8	-75.3700	18	0	74.980	75.760

{range of numerical age  
for a dated event  
type = 5;  
weights are not needed  
because these events  
cannot be adjusted  
up or down section.}

48	5	8	-75.8900	13	0	75.170	76.610
49	5	8	-80.5400	4	0	79.990	81.090
50	4	8	-80.5400	4	0	1.000	1.000
51	1	8	-69.8000	28	1	1.000	1.000
51	2	8	-69.6000	29	2	1.000	1.000
52	1	8	<b>-73.4000</b>	22	1	1.000	1.000
52	2	8	-73.3000	24	2	1.000	1.000
53	1	8	-74.9000	19	1	1.000	1.000
53	2	8	-74.5000	21	2	1.000	1.000
54	1	8	-76.0000	12	1	1.000	1.000
54	2	8	-75.8000	14	2	1.000	1.000
55	1	8	-79.5000	7	1	1.000	1.000
55	1	8	-80.5000	5	1	1.000	1.000
55	2	8	-80.1000	6	2	1.000	1.000
55	2	8	-79.4000	8	2	1.000	1.000
56	1	8	-81.9000	1	1	1.000	1.000
56	2	8	-81.5000	3	2	1.000	1.000
57	5	8	-79.2000	9	0	77.800	80.600
58	5	8	-70.1000	26	0	68.700	71.500
59	5	8	-75.5000	16	0	74.300	76.700
60	1	8	-70.2000	25	1	1.000	1.000
60	2	8	-70.0000	27	2	1.000	1.000
61	1	8	-79.1000	10	1	1.000	1.000
61	2	8	-79.0000	11	2	1.000	1.000
62	1	8	-75.6000	15	1	1.000	1.000
62	2	8	-75.4000	17	2	1.000	1.000
63	1	9	87.0000	1	1	1.000	1.000
63	1	10	-390.0000	1	1	1.000	1.000
63	1	11	5.0000	5	1	1.000	1.000
63	2	9	287.0000	5	2	1.000	1.000
63	2	10	-225.0000	4	2	1.000	1.000
63	2	11	5.0000	5	2	1.000	1.000
64	5	9	168.0000	3	0	70.600	71.400
65	1	10	-390.0000	1	1	1.000	1.000
65	2	10	-300.0000	3	2	1.000	1.000
66	1	11	<b>1.0000</b>	1	1	1.000	1.000
66	2	11	1.0000	1	2	1.000	1.000
67	1	11	6.0000	6	1	1.000	1.000
67	2	11	6.0000	6	2	1.000	1.000
68	1	11	4.0000	4	1	1.000	1.000
68	2	11	4.0000	4	2	1.000	1.000
69	1	11	3.0000	3	1	1.000	1.000
69	2	11	3.0000	3	2	1.000	1.000
70	1	11	2.0000	2	1	1.000	1.000
70	2	11	2.0000	2	2	1.000	1.000
71	1	12	483.5600	4	1	1.000	1.000
71	2	12	483.5600	4	2	1.000	1.000

{negative values are permitted for stratigraphic position, as in a core log measured from younger down to older levels

{positive values for stratigraphic position are measured upward from older to younger levels, as in an outcrop.}

---

## CONOP RUN-CONFIGURATION FILE (CONOP9.CFG)

CONMAN9 will build a suitable CONOP.CFG file to match selected sections and events.  
Use CONMAN “Export” menu.

Must be named “CONOP9.CFG” because this name is written into the program code.  
Only items to the right of the equals signs (“=”) should be edited.  
The CONOP9 Reference Manual documents all the permitted code words.

Curly brackets {} indicate comments that do NOT appear in run-time file.

```
&getinn                                {This block determines input file names and sizes}
PROJECT='SMALL 12x49x22'                {any brief text is OK}
SECTIONS=12                             {integer; must be correct value}
TAXA=49                                 {integer; must be correct value}
EVENTS=22                               {integer; must be correct value}
MAX_LEVELS=1                            {integer; must exceed number of levels in any one section}
MAX_LABELS=15                            {ignore this line}
LOADFILE='SML12x71.dat'                 {must match data file name}
PREPFILE='SML12x71.dis'                  {ignore this line}
SECTFILE='SML12.sct'                     {must match section dictionary name}
SECTTAGFILE='OFF.tag'                     {ignore this line}
SECTTAGS='OFF.dat'                       {ignore this line}
LABELFILE='OFF.lbl'                       {ignore this line}
EVENTFILE='SML71.evt'                     {must match event dictionary name}
EVENTTAGFILE='OFF.tag'                    {ignore this line}
EVENTTAGS='OFF.dat'                       {ignore this line}
BESTKNOWN=0.000                          {ignore this line}
/
```

```
&getans                                {This block determines misfit measures to optimize}
PENALTY='LEVEL'
LETCONTRACT='OFF'
WEIGHTING='ON'
USENEGATIVE='OFF'
NEARENOUGH=5.000
EXCLUSIVES='NO'
FORCECOEX='SS'
FORCEf4L='ON'
HOMERANGE='SL'
SMOOTHER=0.000
SQUEEZER=0.000
SHRINKER=0.000
TEASER=0.010
STACKER='COEX'
/
```

```

&getrun
SOLVER='anneal'
STEPS=600
TRIALS=400
STARTEMP=250.00
RATIO=0.980
HOODSIZE='BIG'
STARTYPE='RAND'
STARTSECT=1
STARTEVENT=0
SHOWMOVIES='CHT'
TRAJECTORY='ALL'
VIDEOMODE='SVGA'
PAUSES='ON'
CURVFILE='grid.grd'
CRV2FILE='grd2.gr2'
/

```

**{This block determines the run time parameters}**

{larger numbers generate longer searches}  
 {larger numbers generate longer searches}  
 {larger numbers needed for larger data sets}  
 {must be less than 1.00; best not altered; smaller values cause coarser searches}  
 {experiment with 'SML' and 'DBL' if you must}  
 {change to 'FILE' to restart from last solution}  
 {ignore this line}  
 {ignore this line}  
 {change to 'DIV' for large data sets}  
 {ignore this line}  
 {ignore this line}

```

&getout
COLUMNS=7
UNLOADMAIN='SMLmain.txt'
FITS_OUT='OFF'
CNFG_OUT='OFF'
SEQN_OUT='OFF'
INCR_OUT='OFF'
LOC_OUT='OFF'
OBS_OUT='OFF'
COMP_OUT='OFF'
UNLOADSECT='SMLsect.txt'
SECT_OUT='MIN'
UNLOADEVNT='SMLevnt.txt'
EVNT_OUT='OFF'
COEX_OUT='OFF'
RUNLOGFILE='SMLlog.txt'
CULLFILE='OFFcull.txt'
SOLNLIST='OFFsolution.sln'
STARTFILE='SMLsoln.dat'
STEPFILE='OFFSMLstep.dat'
BESTARTFILE='SMLbest.dat'
COMPOFILE='SMLcmpst.dat'
COMPOSNMBR=1
COMPOSTYPE='ZST'
OBSDFILE='OFFSMLab.dat'
PLCDFILE='OFFSMLalbet.dat'
EXTNFILE='OFFSMLdelta.dat'
COEXISTFILE='SMLcoex.dat'
FAD_LADFILE='SMLfb4l.dat'
ORDERFILE='SMLordr.dat'
/

```

**{This block determines output file names and sizes}**

{'ON' will generate larger output files}  
 {'ON' will generate larger output files}  
 {'ON' will generate larger output files}  
 {'ON' will generate larger output files}  
 {'ON' will generate larger output files}  
 {'ON' will generate larger output files}  
 {'ON' will generate larger output files}  
 {'MAX' will generate larger output files}  
 {'ON' will generate larger output files}  
 {'ON' will generate larger output files}

## VII.10 Sample lines from CONMAN9 Data Structures

Sample lines from files that are designed to be read and written by the CONMAN9 Fortran routines, but to be readable and editable in word-processing and spreadsheet programs that read and write simple ASCII files (e.g. NotePad, WordPad and Excel).

Asterisks (\*) are reserved as place-holders for empty fields.

Inverted commas ( ' ' ) terminate fields that may contain blank spaces and should not be used within names.

---

### CONMAN9 TAXON DICTIONARY

- *unique* code [one string; no spaces; may prefix clade abbreviation]
- abbreviated name [not unique]
- clade [one name from clade dictionary]
- Genus [one word; no spaces]
- Species [one word; no spaces; *suffix* \_aff or \_cf]
- Subspecies [one word; no spaces]
- Morphotypes [one word; no spaces]
- sensu (etc) [one word; no spaces]
- author [words and spaces; list aff author but not cf author]
- year [four integers]
- *unique* code for preferred synonym}

1030	'Jel.nodo'	ammonite	Jeletzkytes	nodosus	*
	* '(Owen)'		1852 * *		
1031	'Pse.indr'	ammonite	Pseudophyllites	indra	*
	* '(Forbes)'		1846 * *		
1032	'Pac.neub'	ammonite	Pachydiscus_(P.)	neubergicus	neubergicus
	* '(Hauer)'		1858 * *		
1091	'Hau.remb'	ammonite	Hauericeras	rembda	*
	* '(Forbes)'		1846 * *		
1092	'Hau.recf'	ammonite	Hauericeras	rembda_cf	*
	* '(Forbes)'		1846 * *		
1771	'Coi.newe'	ammonite	Coilopoceras	newelli	*
	* 'Benavides-Caleres'		1956 * *		
1772	'Coi.neaf'	ammonite	Coilopoceras	newelli_aff	*
	* '_Zaborsky'		1985 * *		
1783	'Pac.egaf'	ammonite	Pachydiscus	eggertoni_aff	*
	* '_Yazykova'		2002 * *		
1755	'But.butt'	ammonite	Butticeras	buttense	*
	* 'Anderson'		* 1738 *		

---

## CONMAN9 EVENT DICTIONARY

- *unique* event code
- event type code [CONOP reserved words: ASH, AGE, MID, APP . . .]
- abbreviated event name
- event name
- minimum age [0.00 for events other than AGEs]
- maximum age [0.00 for events other than AGEs]

0100	ASH	KT_irid	'KT iridium/tektites'	0.00	0.00
0101	ASH	KT_GSSP	'KT GSSP El Kef'	0.00	0.00
0116	ASH	34n~33r	'C34n~C33r'	0.00	0.00
0531	AGE	Sr71.0	'Strontium 71.0 +_0.2'	70.6	71.4
0500	AGE	Ob88_2	'Obad88 2_69.0'	67.6	71.4

---

## CONMAN9 SECTION DICTIONARY

- section name
- abbreviated name
- *unique* filename
- geographic location [for convenience; not used]
- geologic a age [for convenience; not used]
- unique code for reference 1 [links to reference dictionary]
- unique code for reference 2 [links to reference dictionary]
- unique code for reference 3 [links to reference dictionary]
- use in scaling? yes/no

'Seymour Island A' 'SyA' 'ANTSeymA.sec' 'Antarctic Peninsula'  
 'Maastrichtian' 'MACE8499' 'CRAM0411' 'PIRR9709' 'yes'

'Seymour Island C' 'SyC' 'ANTSeymC.sec' 'Antarctic Peninsula'  
 'Maastrichtian' 'MACE8499' 'OLIV8926' '\*' 'yes'

'Seymour Island D' 'SyD' 'ANTSeymD.sec' 'Antarctic Peninsula'  
 'Maastrichtian' 'MACE8499' 'OLIV8926' '\*' 'yes'



---

## CONMAN9 REFERENCE DICTIONARY

- unique code [boldface in listing shows source of code characters]
- author(s)
- year
- paper/chapter title
- editors
- book/volume title
- journal name
- publisher
- volume number
- page(s)

'EMER9439'     '**Emerson**, B. L., Emerson, J. H., Akers, R. E., and Akers, T. J.'  
'1994'     'Texas Cretaceous ammonoids and nautiloids'     '\*'     '\*'     'Texas  
Paleontology Series'     '\*'     'v. 5'     '439 p.'

'MATS6004'     '**Matsumoto**, T.'     '1960'     'Upper Cretaceous Ammonites of  
California, Part III'     '\*'     '\*'     'Memoirs of the Faculty of Science , Kyushu  
University, Series D, Geology, Special Volume'     '\*'     'v. 2'     'p. 1-204.'

KUCH8991'     '**Kuchler**, T., and Kutz, A.'     '1989'     'Biostratigraphie des Campan  
bis Unter-Maastricht der E-Barranca und des Urdiroz/Imiscoz-Gebietes (Navarra, N.-  
Spanien)'     'in Wiedmann, J. (ed.)'     'Cretaceous of the Western Tethys'  
'Proceedings of the Third International Cretaceous Symposium, Stuttgart'     'E.  
Schweizerbartsche Verlagsbuchhandlung'     '\*'     'p. 191-213'

'KUCH0123'     '**Kuchler**, T., Kutz, A., and Wagereich, M.'     '2001'     'The  
Campanian-Maastrichtian boundary in northern Spain (Navarra Province): the Imiscoz  
and Erro sections.'     'in Odin, G. S. (ed.)'     'The Campanian-Maastrichtian Stage  
Boundary'     '\*'     'Elsevier'     '\*'     'p. 723-744.'

'LARS9746'     '**Larson**, N. L., Jurgensen, S. D., Farrar, R. A., and Larson, P. L.'  
'1997'     'Ammonites and other cephalopods of the Pierre Seaway: an identification  
guide'     '\*'     '\*'     '\*'     'Geoscience Press Inc, Tucson, AZ'     '\*'     '146  
p.'

---

## CONMAN9 RANGE CHART FILES

### Sample lines for paired events:

e.g. taxon range (FAD-LAD) or uncertainty interval (MAX-MIN)

- *unique* code for preferred name [links to taxon dictionary; may be updated with synonym]
- abbreviated preferred name [a convenience for the human reader only]
- first appearance level
- last appearance level
- unique code for name as published [links to taxon dictionary; never updated]
- abbreviated name as published [a convenience for the human reader only]
- source [links to reference dictionary]
- date of entry
- 

1197	'Kit.laur'	220.00	232.00	1197	'Kit.laur'	ZINS8931	14-FEB-05
1184	'Mao.dena'	220.00	235.00	1184	'Mao.dena'	ZINS8931	14-FEB-05
1190	'Gro.gemm'	220.00	235.00	1190	'Gro.gemm'	ZINS8931	14-FEB-05
1195	'Dip.maxi'	220.00	246.00	1195	'Dip.maxi'	ZINS8931	14-FEB-05
1192	'Pac.ulti'	220.00	252.00	1192	'Pac.ulti'	ZINS8931	14-FEB-05

### Sample line for unpaired events

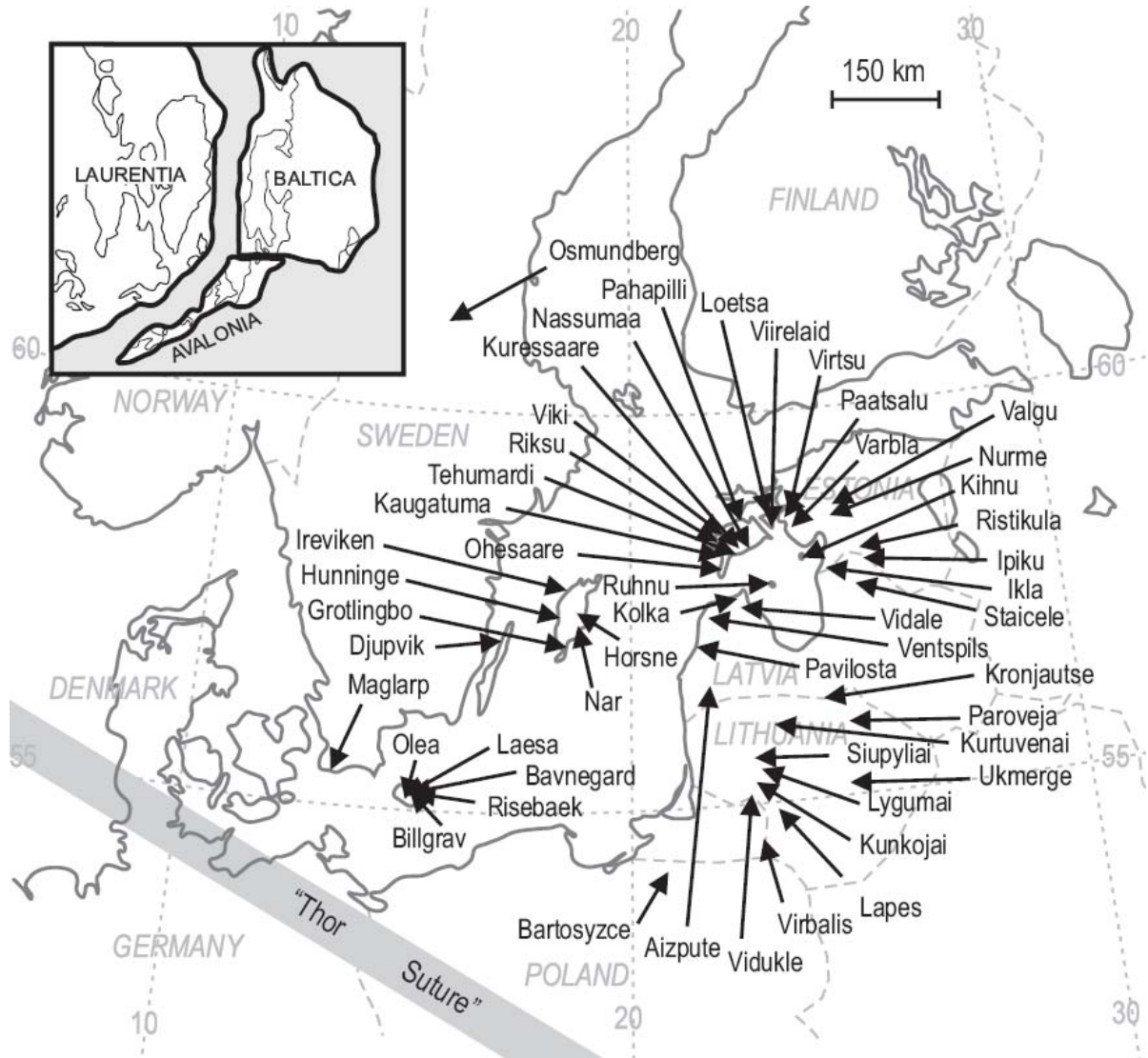
e.g.. marker bed (ASH), dated event (AGE), acme event (MID), appearance (APP)

- *unique* code for preferred event [links to event dictionary]
- abbreviated name [a convenience for the human reader only]
- event type [a CONOP code word – ASH, AGE, MID, APP, DIS . . . ]
- stratigraphic level
- code for event as published [links to event dictionary]
- abbreviated name as published [a convenience for the human reader only]
- source [links to reference dictionary]
- date of entry
- 

0100	'KT_irid'	ASH	9.40	0100	'KT_irid'	ELLI9475	14-FEB-05
0531	'Sr71.0'	AGE	168.00	0531	'Sr71.0'	CRAM0411	03-MAR-05

## VIII. Silurian of Baltica

A case study in the analysis of a composite time line and the incorporation of carbon isotope excursions.

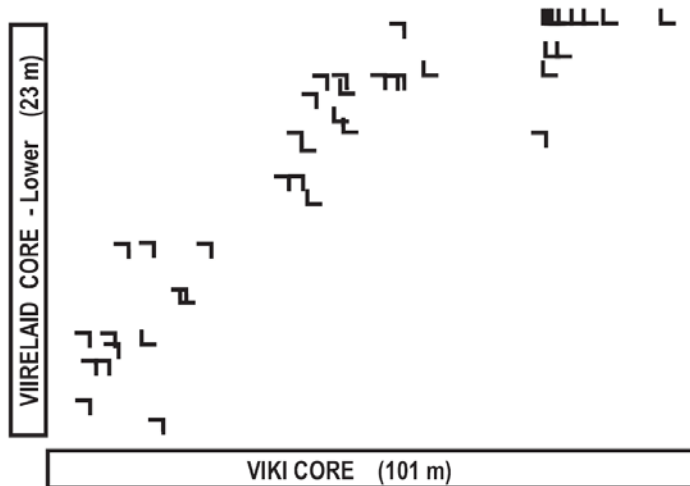


**Figure VIII.1:** Location of cores and sections in the Silurian of Baltica, shown on the present day map. Inset: early Silurian paleogeography.

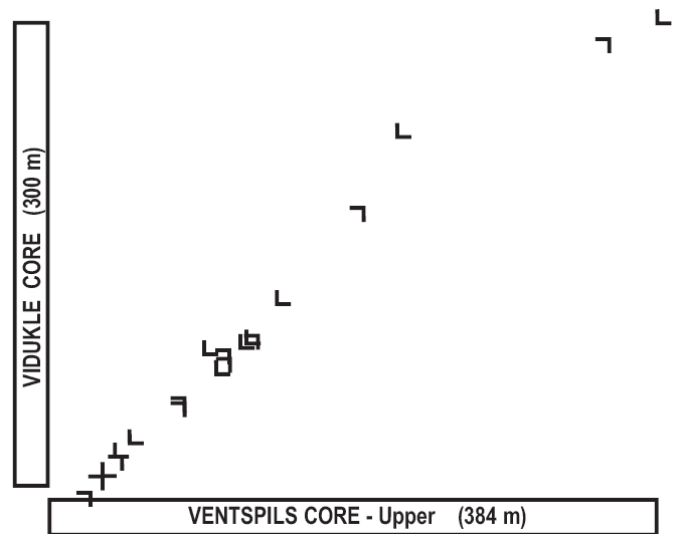
### VIII.1. State-of-the-Art Integrated Stratigraphy

The early Paleozoic strata on the ancient continent of Baltica (Fig. VIII.1) have become the subject of some of the best integrated biochronology. For the Silurian, detailed range charts of chitinozoan,

conodont, and graptolite taxa have been published for numerous well cores and some surface outcrops. These excellent data have been augmented by carbon isotope stratigraphy and placed in the context of a robust framework of bentonite correlations, based upon chemical fingerprinting. The Silurian



**Figure VIII.2** Graphic plot of shared Silurian range ends in two Baltic wells.



**Figure VIII.3:** Graphic plot of shared Silurian range ends in two more Baltic wells.

data set includes about 40 sections and 600 taxa, plus more than 80 other horizons of time stratigraphic value – K-bentonites, isotope excursions and hiatuses. On typical desktop computers, CONOP takes a few hours to find a best-fit time line for a data set of this size. We shall use a time line that has already been run in order to explore the post-run analyses that CONOP performs. First, however, let us examine some pairs of sections via graphic correlation.

### VIII.3: Graphic Correlation

*Marker Beds* – In graphic correlation a shared marker bed, like a bentonite matched by chemical fingerprinting, is a point on the graph through which the line of correlation (LOC) must pass. CONOP simply implements marker beds as events that may not be adjusted up- or down-section.

#### EXERCISE VIII.3a

Correlation augmented with bentonites.

3a.1 Draw a line of correlation on the basis of the range-end data in Figure VIII.2

3a.2 Compare the results after consideration of the shared bentonites plotted in figure VIII.5 and with the CONOP-generated LOC (Fig. VIII.7)

*Uncertainty Intervals:* A shared peak in carbon isotopic ration might be considered as another kind

of marker bed. Most often, however, the shape of the isotopic anomaly varies significantly from section to section. The variation might be a real effect or an artifact of the material sampled or the sample spacing. In such situations, we have a less precise marker and ought to place an uncertainty interval on the position of the isotope peak in each section. Then we can say that a correlative level occurs somewhere (not precisely known) in that interval in each section.

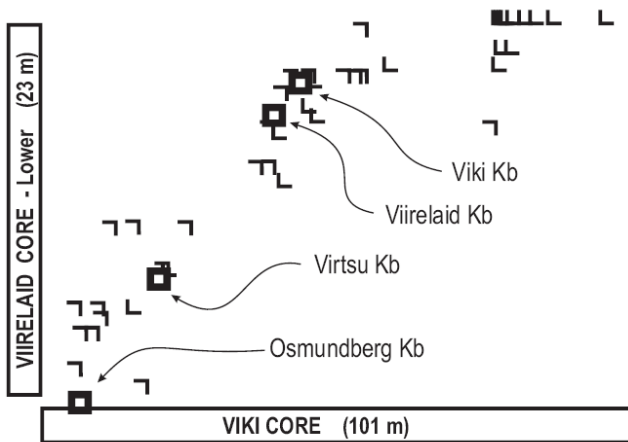
Graphic correlation would implement this concept as a rectangle through which the LOC must pass, without being required to intersect the center. CONOP implements the concept as a “boxed” event. The box is a pair of events, like a taxon range. Whereas the ends of an observed taxon range may be stretched farther apart to match the corresponding FAD and LAD in a time line, the uncertainty interval may only shrink to match the corresponding MAX and MIN in the time line.

#### EXERCISE VIII.3b

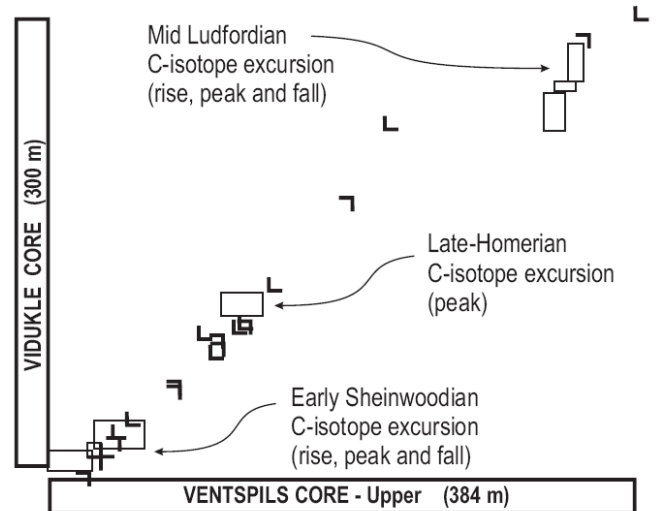
Correlation augmented with carbon isotopes

3b.1 Draw a line of correlation on the basis of the range-end data in Figure VIII.3

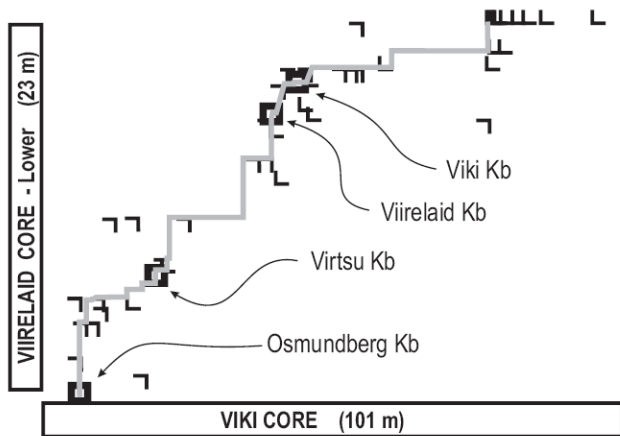
3b.2 Compare the results after consideration of the shared isotopic excursions plotted in figure VIII.6 and with the CONOP-generated LOC (Fig. VIII.8).



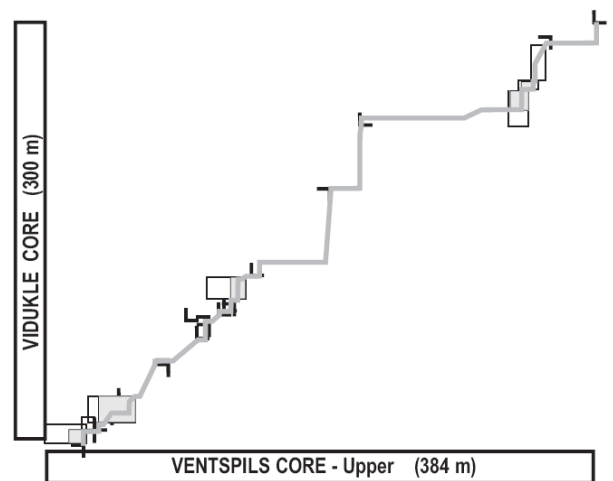
**Figure VIII.4:** Cross-plot of two Baltic wells that share four bentonites (Kb). Baltic practice names the bentonites for core locations or for depth in the Viki core. The bentonites are LOCs are constrained to pass through the center of these squares.



**Figure VIII.5:** Cross-plot of two Baltic wells that share three carbon isotope excursions. Two are each coded as three uncertainty intervals (boxes on graph) – the rise, the crest and the fall in  $\delta$  values. For the other, only the crest is coded. LOCs are constrained to intersect some part, edge or corner of all these boxes.



**Figure VIII.6** CONOP-generated line of correlation for figure VII.4. The LOC (gray) results from a consideration of all 44 sections, not merely the two in this graph.



**Figure VIII.7** CONOP-generated line of correlation for figure VII.5.

### VIII.3: Retrieving the Last CONOP Solution

At the end of a CONOP run, the details of the best-fit time line are written to several text-files named in the configuration file, CONOP9.CFG. There is also a Text-File Output menu option to write out the best fit solution in the format of a CONMAN9 section file. This option allows comparison of time lines

prepared from the same data with different objective functions. Among the text file versions of the best-fit solution that are written at the end of a CONOP search is one, named by the `STARTFILE = "..."` parameter. This is the file that CONOP will read in as the initial timeline if the `STARTYPE` parameter is set to `'FILE'`. In order to reload the previously

determined best-fit time-line and proceed to the end of a search without any mutations, start with this file and perform no trials and no cooling steps as directed in exercise VII.3.

The following exercise is intended as a rather open-ended invitation to explore the graphical output options at the end of a CONOP9 run. Many of the illustrations in this document were redrawn from these graphical output screen. Try to generate similar plots for the Baltic Silurian data. In a course, instructors may lead this exploration on screen. Otherwise, refer to the CONOP Users' Guide and Reference Manual.

#### EXERCISE VIII.3:

Restart from the last solution.

3.1 Find the folder with the Baltic Silurian run and open the CONOP.CFG file

3.2 Make sure that CONOP9.CFG contains the following parameter settings:

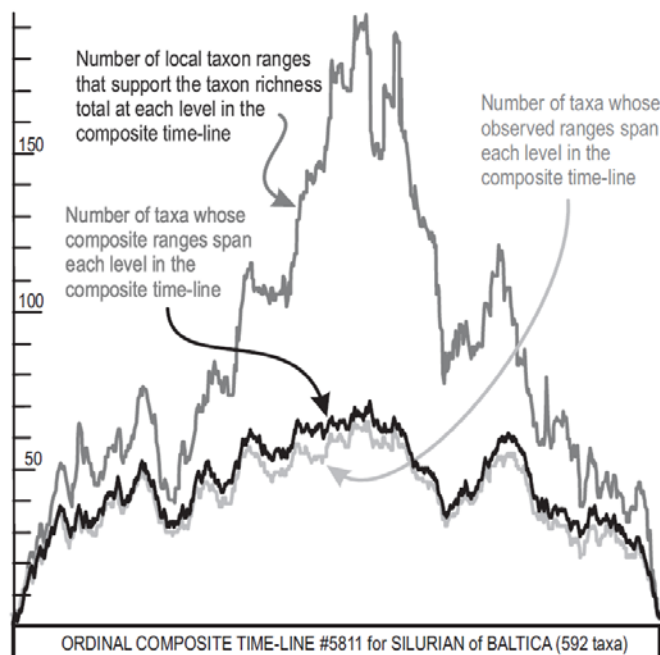
```
STEPS=0
TRIALS=0
STARTEMP=0.00

STARTYPE= ' FILE '

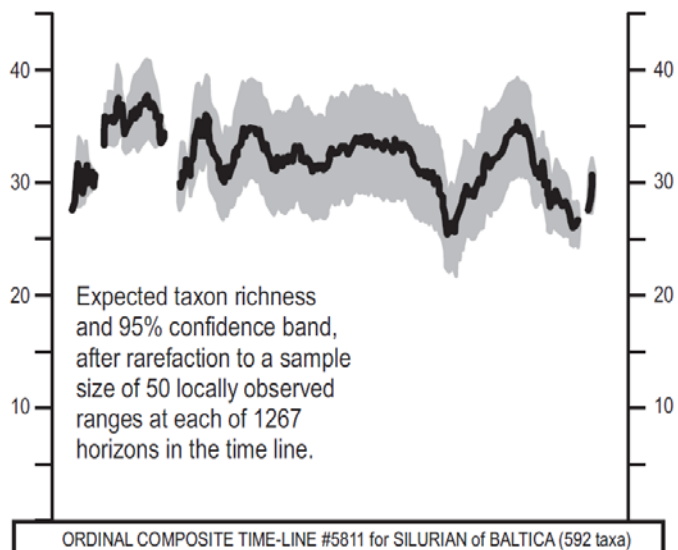
PAUSES= ' OFF '
```

3.3 Restart CONOP.EXE. At the end of the run, explore the new output files written automatically to the data folder, the options in the graphical output menu, and the option to write more text files to disk. Most graphics can be exited by striking the {Esc} key two or three times. For many options, this leads through a series of related graphs. Some plot very rapidly, others take a while to calculate.

3.4 Some graphical summaries of the Silurian Baltic Time are provided below. Try to generate similar plots on screen. The figure captions give some advice on this.



**Figure VIII.8** Taxon richness summaries for Silurian of Baltica based on a composite time line of 592 graptolite, conodont and chitinozoan taxa in 44 sections, for a total of 3052 locally observed range ends (2649 of them away from section ends). The local taxon ranges proved 10,953 pairwise coexistences of taxa and 56,052 instances of a first-appearance before a last-appearance. K-bentonites and isotopic excursions were included to better constrain the outcome. The time-line was optimized using the level objective function with a small secondary penalty for excess first-before-last relationships implied by the composite. The total misfit to the field data was 5811 levels – an average of 2.19 levels per adjustable observation. The black curve can be accessed from the CONOP “taxon richness” option of the Graphical Output menu. The two grey curves may be found in the “Composite Range Increments” option of the same menu. To reach them, wait while CONOP maps the observed ranges into the composite; set the rarefaction level to 1 (no rarefaction); and click {Esc} to advance beyond the taxon histograms. Cycle through the succeeding histograms by using the + (plus) and – (minus) keys.



**Figure VIII.9.** Taxon richness for the same time-line as figure VIII.8 after rarefaction. The gaps in the graph are intervals for which fewer than 50 locally observed ranges are available. Fortunately, this cull also removes the terminal parts of the diagram where richness is low, not because the faunas had low diversity, but because the data are artificially truncated at the limits of the Silurian.

To generate this diagram from CONOP-, repeat the process for figure VIII.8, but this time select a rarefaction level of 50 and request a text-file output. CONOP does not (yet) draw the rarefied plot. The data in the text file need to be opened in a spreadsheet program. Open the file `rarifact.txt` and read the header to understand the file content. Delete all the text lines in the header and then save the file as `rarifact.dat`. The new text-free file can be opened in a spreadsheet program, that can build graphs. The position in the ordinal composite is the first column.

## **IX. Classic Riley Formation Data Set**

The classic data set for comparison of quantitative correlation.

This exercise is described in the CONOP9 Users' Manual and uses sample data available with the CONOP program.

## **XI. Mohawkian Facies Problems**

A graptolite data set from the late Ordovician of Laurentia that is notoriously difficult to correlate.

This exercise is described in the CONOP9 Users' Manual and uses sample data available with the CONOP program.

## **XI. Liassic Ammonites of England**

A classroom exercise in biostratigraphic correlation based on very old English publications.

This exercise can be downloaded from the Spring 2009 GEO 206B website at UC Riverside [http://ilearn.ucr.edu] by clicking the following buttons and tabs:  
[Guest Login](#) > [Courses](#) > [Course Catalog](#)  
> [Geosciences\(GEO\)](#) > [Spring 2009\(1 \)](#)>  
[GEO 206B 001 09S](#)

## **XI. Carboniferous of the Harz Mountains**

A classroom exercise in graphic correlation and diachronous facies based on 1970's German publications.

This exercise can be downloaded from the Spring 2009 GEO 206B website at UC Riverside [http://ilearn.ucr.edu] by clicking the following buttons and tabs:  
[Guest Login](#) > [Courses](#) > [Course Catalog](#)  
> [Geosciences\(GEO\)](#) > [Spring 2009\(1 \)](#)>  
[GEO 206B 001 09S](#)



### XIII. References

- AGTERBERG, F. P. 1990. Automated stratigraphic correlation. *Developments in Palaeontology and Stratigraphy*, 13:1-424.
- AGTERBERG, F. P., AND GRADSTEIN, F. M. 1999. The RASC method for ranking and scaling of biostratigraphic events. *Earth Science Reviews*, 46:1-25.
- AGTERBERG, F. P., AND NEL, L. D. 1982. Algorithms for the ranking of stratigraphic events. *Computers and Geosciences*, 8:69-90.
- ALROY, J. 1992. Conjunction among taxonomic distributions and the Miocene mammalian biochronology of the Great Plains. *Paleobiology*, 18:326-43.
- ALROY, J. 1994. Appearance event ordination: a new biochronological method. *Paleobiology*, 20:191-207.
- ALROY, J. 2000. New methods for quantifying macroevolutionary patterns and processes. *Paleobiology*, 26:707-33.
- AUBRY, M.-P., 1995. From chronology to stratigraphy: interpreting the Lower and Middle Eocene stratigraphic record in the Atlantic Ocean, *SEPM Special Publication*, 54, 213-274.
- BAADSGAARD, H., LERBEKMO, J. F., WIJBRANS, J. R., SWISHER, C. C., AND FANNING, M. 1993. Multi-method radiometric age for a bentonite near the top of the *Baculites reesidei* zone of southwestern Saskatchewan (Campanian-Maastrichtian stage boundary?), *Canadian Journal of Earth Science*, 30:769-775.
- BLACKHAM, M. 1998. The unitary association method of relative dating and its application to archaeological data. *Journal of Archaeological Method and Theory*, 5:165-207.
- CARNEY, J. L. AND PIERCE, R. W. 1995. Graphic Correlation and composite standard databases as tools for the exploration biostratigraphers, p. 23-43. *In* K. O. Mann, and H. R. Lane (eds.), *Graphic Correlation*. Special Publications of the Society of Economic Paleontologists and Mineralogists, 53.
- CLARK, R. M. 1985. A FORTRAN program for constrained sequence-slotting based on minimum combined path length. *Computers and Geoscience*, 11:605-17.
- CLARK, R. M. 1995. Depth-matching using PCSLOT version 1.6. *Newsletter of the INQUA Working Group for Data-Handling Methods* 13. <http://kv.geo.uu.se/inqua/news12/n13-mc.htm>
- CODY, R. D., LEVY, R. H., HARWOOD, D. M., AND SADLER, P. M. 2008. Thinking outside the zone: high resolution quantitative diatom biochronology for the Antarctic Neogene. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 269(1-2):92-121.
- COOPER, R.A., CRAMPTON, J. S., RAINE, J.I., GRADSTEIN, F. M., MORGANS, H. E. G., ET AL. 2001. Quantitative biostratigraphy of the Taranaki Basin, New Zealand: a deterministic and probabilistic approach. *American Association of Petroleum Geologists Bulletin*, 85:1469-98.
- COOPER, R. A., AND SADLER, P. M. 2004. The Ordovician Period, p. 165-187 in F. Gradstein, J. Ogg and A. Smith (eds.) *A Geologic Time Scale 2004*. Cambridge University Press
- CRAME, J. A., FRANCIS, J. E., CANTRIL, D. J., and PIRRIE, D. 2004. Maastrichtian stratigraphy of Antarctica. *Cretaceous Research*, 25:411-423
- CUVIER, G., AND BRONGNIART, A. 1808. *Essai sur la géographie minéralogique des environs de Paris*, *Annales du Musée Histoire Naturelle de Paris*, 11:293-326.
- DARWIN, C. 1859. *The Origin of Species*. Murray, London, 490 p.
- DELL, R. F., KEMPLE, W. G., AND TOVEY, C. A. 1992. Heuristically solving the stratigraphic correlation problem. *Proceedings of the 1st. Industrial Engineering Research Conference*, 1:293-97
- D'ORBIGNY, A. 1851. *Cours Elementaires de Paleontologie et de Geologie Stratigraphiques*. Masson, Paris, 382 p.
- EDWARDS, L. E. 1978. Range charts and no-space graphs. *Computers and Geosciences*, 4:247-255
- EDWARDS, L. E. 1982. Quantitative biostratigraphy: the methods should suit the data, p. 45-60. in J. M Cubitt, and R. A. Reymont (eds.). *Quantitative stratigraphic correlation*, Chichester, Wiley.
- ELLES, G. L., AND WOOD, E. M. R. 1907. A monograph of British graptolites. *Palaeontographical Society Monograph*, 61:1-216.
- GARDNER, T. W., JORGENSEN, D. W., SHUMAN, C., AND LEMIEUX, C. R. 1987. Geomorphic and tectonic process rates: effects of measured time interval, *Geology*, 15:259-261.
- GINGERICH, P. D. 1983. Rates of evolution: effects of time and temporal scaling, *Science*, 222:159-161.
- GINGERICH, P. D. 2001. Rates of evolution on the time scale of the evolutionary process, *Genetics*, 112/113:127-144.
- GLOVER, F. 1989. Tabu search – Part I. *Operations Research Society of America Journal on Computing*, 1:190-206.
- GORDON, A. D., AND REYMENT, R. A. 1979. Slotting of borehole sequences, *Mathematical Geology*, 11:309-27.
- GRADSTEIN, F. M., AND AGTERBERG, F. P. 1998. Uncertainty in stratigraphic correlation, p. 9-29, in F.

- M. Gradstein and K. O. Sandvik (eds.), Sequence Stratigraphy: concepts and applications, Norwegian Petroleum Society Special Publication 8.
- GRADSTEIN, F., OGG, J., AND SMITH, A. 2004. A Geologic Time Scale 2004. Cambridge University Press, 589p.
- GUEX, J. 1977. Une nouvelle méthode d'analyse biochronologique. Bulletin Laboratoire Géologique Lausanne, 224:309-22
- GUEX, J. 1991. Biochronological Correlations, Springer Verlag, 252 p.
- HAMMER, O and HARPER, D. A. T. 2005. Paleontological Data Analysis. Blackwell, 368 p.
- HANCOCK, J. M. 1977. The historic development of concepts of biostratigraphic correlation, p. 3-22. In E. G. Kauffman, and J. E. Hazel, (eds.), Concepts and Methods of Biostratigraphy. Dowden, Hutchinson and Ross, Stroudsburg, Pennsylvania.
- HARRIES, P., 2003. High Resolution Approaches in Paleontology, Kluwer Academic Press
- HAY, W. W., AND SOUTHAM, J. R. 1978. Quantifying Biostratigraphic Correlation. Annual Reviews of Earth and Planetary Science, 6:353-375.
- HICKS, J. F., OBRADOVICH, J. D., AND TAUXE, L. 1999. Magnetostratigraphy, isotopic age calibration and intercontinental correlation of the Redbird section of the Pierre Shale, Niobrara County, Wyoming, USA. Cretaceous Research, 20:1-27.
- HOOD, K. C. 1986. GRAPHCOR - Interactive graphic correlation software, version 2.2: copyright 1986-1995, KC Hood.
- HUTT, J. E. 1975. The Llandovery graptolites of the English Lake District. Palaeontographical Society Monographs, 1-2:1-137.
- INGBER, L. 1993. Simulated annealing: practice versus theory. Mathematical and Computer Modelling, 11:29-57.
- KEMPLE, W. G., SADLER, P. M., AND STRAUSS, D. J. 1995. Extending graphic correlation to many dimensions: stratigraphic correlation as constrained optimization, P. 65-82, in K. O. Mann, and H. R. Lane (eds.) Graphic Correlation. Special Publications of the Society of Economic Paleontologists and Mineralogists, 53.
- KIRKPATRICK, S., GELATT, C. D., AND VECCHI, M. P. 1983. Optimization by simulated annealing. Science, 220:671-680.
- LAPWORTH, C. 1876. On Scottish Monograptidae. Geological Magazine, 2(3):308-321.
- LAUDAN, R., 1976. William Smith. Stratigraphy without Paleontology. Centaurus, 20(3):210-226.
- LIOW, L. H., SKAUG, H. J., ERGON, T., AND SCHWEDER, T. 2010. Global occurrence trajectories of microfossils: environmental volatility and the rise and fall of individual species. Paleobiology, 36(2):224-252
- LUDVIGSEN, R., WESTROP, S. R., PRATT, B. C., TUFNELL, P. A., AND YOUNG, G. A., 1986. Dual Biostratigraphy: zones and biofacies. Geoscience Canada, 13: 139-154.
- MACELLARI, C. E. 1984. Late Cretaceous stratigraphy, sedimentology, and macropaleontology of Seymour Island, Antarctic Peninsula, PhD Thesis, The Ohio State University, 598 p.
- MACELLARI, C. E. 1986. Late Campanian-Maastrichtian ammonite fauna from Seymour Island (Antarctic Peninsula). Paleontological Society Memoir 18:1-55.
- MACLEOD N., AND SADLER P. M. 1995. Estimating the line of correlation, p. 51-64, in K. O. Mann and H. R. Lane, (eds.) Graphic Correlation. Special Publications of the Society of Economic Paleontologists and Mineralogists, 53.
- MANN, K. O., AND LANE, H. R. 1995. Graphic correlation: a powerful stratigraphic technique comes of age. p. 3-13, in K. O. Mann and H. R. Lane (eds.) Graphic Correlation. Special Publications of the Society of Economic Paleontologists and Mineralogists, 53.
- MARSHALL, C. R., 1990. Confidence intervals on stratigraphic ranges. Paleobiology, 16:1-10.
- McGOWRAN, B., 2005 Biostratigraphy: Microfossils and Geologic Time. Cambridge University Press, 459 p.
- MCLAREN, D. J. 1988. Detection and significance of mass killings. Canadian Society of Petroleum Geologists Memoir, 14:1-7.
- MELCHIN, M. J., COOPER, R. A., AND SADLER, P. M. 2004. The Silurian Period, p. 188-201 in F. Gradstein, J. Ogg and A. Smith (eds.) A Geologic Time Scale 2004. Cambridge University Press.
- MURPHY, M. A., 1994. Fossils as a basis for chronostratigraphic interpretation. Neues Jahrbuch fuer Geologie und Palaeontologie Abhandlungen, 192:255-271.
- OBRADOVICH, J. D. 1988. A different perspective on glauconite for geologic time scale studies. Paleoceanography, 3:757-770.
- OBRADOVICH, J. D. 1993. A Cretaceous time scale, p. 379-396, in W. G. E. Caldwell and E. G. Kauffman (eds.), Evolution of the Western Interior Basin. Geological Association of Canada, Special Paper 39.
- OBRADOVICH, J. D., AND COBBAN, W. A. 1975. A time scale for the Late Cretaceous of the Western Interior of North America, p. 31-54, in W. G. E. Caldwell, (ed.), The Cretaceous system in the Western Interior of North America, Geological Association of Canada Special Paper 13.
- OPPEL, A. 1856-58. Die Jura Formation Englands, Frankreichs, und des südwestlichen Deutschlands; nach ihren einzelnen Gliedern eingeteilt und verglichen. Stuttgart, Ebner und Seubert, 875p. (Württembergische Naturwissenschaftliche Jahreshefte, 12:121-556, 13:141-396, 14:128-291)

- PHILLIPS, J. 1829. Illustrations of the Geology of Yorkshire. T. Wilson, York. 192p.
- PIRRIE, D., CRAME, J. A., LOMAS, S. A., AND RIDING, J. B. 1997. Late Cretaceous stratigraphy of the Admiralty Sound Region, James Ross Basin, Antarctica. *Cretaceous Research*, 18:109-137
- PRIBYL, A. 1940. Die Graptolithen fauna des mittleren Ludlows von Böhmen (oberes eß). *Věstník Státního Geologického Ústavu Československé*, 16:63-73.
- REINECK, H.-E. 1960. Über Zeitlücken im Rezenten Flachsee Sedimenten. *Geologische Rundschau*, 49:149-161.
- REZNICK, D. N., SHAW, F. H., RODD, F. H., AND SHAW, R. G. 1997. Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science*, 275:1934-1937.
- RYAN, P. D., RYAN, M. D. C., AND HARPER, D. A. T. 1999. A new approach to seriation, p. 433-49., in D. A. T. Harper (ed.) *Numerical Palaeobiology: computer-based modeling and analysis of fossils and their distributions*, Chichester, UK: Wiley
- SADLER, P. M. 1981. Sediment accumulation rates and the completeness of stratigraphic sections, *Journal of Geology*, 89:569-584.
- SADLER, P. M. 1993. Time scale dependence of the rates of unsteady geologic processes, p. 221-228 in J. M. Armentrout, R. Bloch, H. C. Olsen, and B. F. Perkins (eds.), *Rates of Geologic Processes*, SEPM Gulf Coast Section Annual Research Conference, 14.
- SADLER, P. M. 1999. The influence of hiatuses on sediment accumulation rates, p.15-40, in P. Bruns and H. C. Haas (eds.), *On the Determination of Sediment Accumulation Rates*, Trans Tech Publications, 5.
- SADLER, P. M. 2004. Quantitative Biostratigraphy - achieving finer resolution in Global Correlation. *Annual Reviews of Earth and Planetary Sciences*, v. 32, p. 187-213.
- SADLER, P. M. 2006. Composite time-lines: a means to leverage resolving power from radioisotopic dates and biostratigraphy, p. 145-170. *In* T. D. Olszewski (ed.) *Geochronology: Emerging Opportunities*, Paleontological Society Papers, 12:145-170.
- SADLER, P. M. 2009. False coincidences – a pervasive characteristic of the unevenly under-sampled fossil record? 9<sup>th</sup> North American Paleontological Convention Abstracts, Cincinnati Museum Center Scientific Contributions, 3:256-257.
- SADLER, P. M. 2010. Constrained optimization approaches to the paleobiologic correlation and seriation problems: a users' guide and reference manual to the Conop family of programs, version 7.61, copyright 1998-2010, P. M. Sadler
- SADLER, P. M., AND COOPER, R. A. 2003 Best-fit intervals and consensus sequences: comparison of the resolving power of traditional biostratigraphy and computer-assisted correlation. in P. Harries (ed.) *High Resolution Stratigraphic Approaches in Paleontology*. Kluwer-Academic Press.
- SADLER, P. M., AND COOPER, R. A., 2004, Calibration of the Ordovician Time Scale, in Webby, B., Paris, F., Droser, M.L., and Percival, I. G. (eds.), *The Great Ordovician Biodiversification Event*, Columbia University Press (*Critical Moments and Perspectives in Earth History and Paleobiology Series*) p. 48-51.
- SADLER, P. M., COOPER, R. A., AND MELCHIN, M. 2009. High-resolution, early Paleozoic (Ordovician-Silurian) time scales. *Geological Society of America Bulletin*, 121:887-906.
- SADLER, P. M., AND SABADO, J. A., 2009, Automated correlation, seriation, and the treatment of biotic dissimilarity. *Museum of Northern Arizona Bulletin* 65:21-35
- SALVADOR, A., 1994. *International Stratigraphic Guide*. Boulder, USA: Geological Society of America, 214pp.
- SAVARY, J., AND GUEX, J. 1991. Biograph: un nouveau programme de construction des correlations biochronologique basées sur les associations unitaires. *Bull. Lab. Géol. Univ. Lausanne* 313:317-40.
- SAVARY, J., AND GUEX, J. 1999. Discrete biochronological scales and unitary association: description of the BioGraph computer program. *Mémoire Géologique Lausanne*, 34:1-281.
- SCOTT, G. H., 1985. Homotaxy and biostratigraphical theory. *Palaeontology*, 28:777-782.
- SEDGWICK, A., AND MURCHISON, R. I. 1839. Classification of the older stratified rocks of Devonshire and Cornwall. *Philosophical Magazine and Journal of Science Series 3*, 14:241-260.
- SHAW, A. B. 1964. *Time in Stratigraphy*. New York: McGraw Hill. 365 pp.
- SHAW, A. B. 1995. Early history of graphic correlation. p. 15-19. *In* K. O. Mann, and H. R. Lane (eds.) *Graphic Correlation*. Special Publications of the Society of Economic Paleontologists and Mineralogists, 53.
- SMITH, W., 1799, Tabular view of the order of strata in the vicinity of Bath with their respective organic remains.
- SPENCER, H. 1864. *The Principles of Biology*, volume 1. Williams and Norgate. London. 492p.
- TIPPER, J. C. 1988. Techniques for quantitative stratigraphic correlations: a review and annotated bibliography. *Geological Magazine*, 125:475-94
- WEBSTER, M., SADLER, P. M., KOOSER, M. A., AND FOWLER, E. 2003. Combining stratigraphic sections and museum collections to increase biostratigraphic resolution: application to Lower Cambrian trilobites from southern California, in P. Harries (ed.) *High Resolution Stratigraphic Approaches in Paleontology*. Kluwer-Academic Press.

WOOD, E. M. R. 1900. The Lower Ludlow formation and its graptolite fauna. *Quarterly Journal of the Geological Society of London*, 56:415-492

ZHANG, T., AND PLOTNICK, R. E. 2001. Determining the line of correlation using genetic algorithms. *Geological Society of America Abstracts with Programs* 33(6):141 (Abstr.).

ZINSMEISTER, W. J. 2001. Late Maastrichtian Short-term biotic events on Seymour Island, Antarctic Peninsula. *Journal of Geology*, 109:213-229.

## XIV. Acknowledgements

Lucy Edwards, Marilyn Kooser and Mike Murphy introduced me to graphic correlation., Robert Dell, Bill Kemple, Ian Raine and David Strauss showed me some essential statistical and programming approaches. Funds for the development of CONOP and CONMAN were part of several grants from the National Science Foundation: EAR0518939; OPP0338274; EAR9980372; EAR9219731. Quality time to attempt the coding was provided by generous invitations to spend sabbatical leaves at GNS Science in New Zealand, MIT, and the University of Chicago. For these invitations I thank Roger Cooper, James Crampton, Sam Bowring, Mike Foote and Mark Webster.

Several colleagues provided data sets to test CONOP: Roger Cooper, James Crampton, Vladimir Davidov, Dan Goldman, Gilbert Klapper, Steve Leslie, Carlos Macellari, Joerg Maletz, Chuck Mitchell, Hugh

Morgans, Mike Murphy, Scott Samson, Percy Strong, Wang Yue and Bill Zinsmeister.

Several graduate students and interns tested early versions of the programs or helped develop the data sets: David Babst, Yanying Bai, Rosemary Cody, Phoebe Cohen, Carlos Cuartos, Robert Falero, Daniel Garson, Angela Gerhardt, Melanie Hopkins, Greg Miller, Camille Partin, Jen Sabado, Bryan Sell and Erik Stokstadt.

The nuclei for this course notebook can be found in an encyclopedia entry, three book chapters and two Paleontological Society short courses. For invitations to write these precursors I am indebted to John Alroy, Sam Bowring, Cinzia Cervato, Mary Droser, Peter Harries, Gene Hunt and Tom Olzewski. The incentive to pull these parts together was an invitation to Nanjing from Fan Junxuan.

## XV. Worksheets to Print

Larger versions of text figures  
for graphic-correlation exercises  
that draw lines of correlation (LOCs)

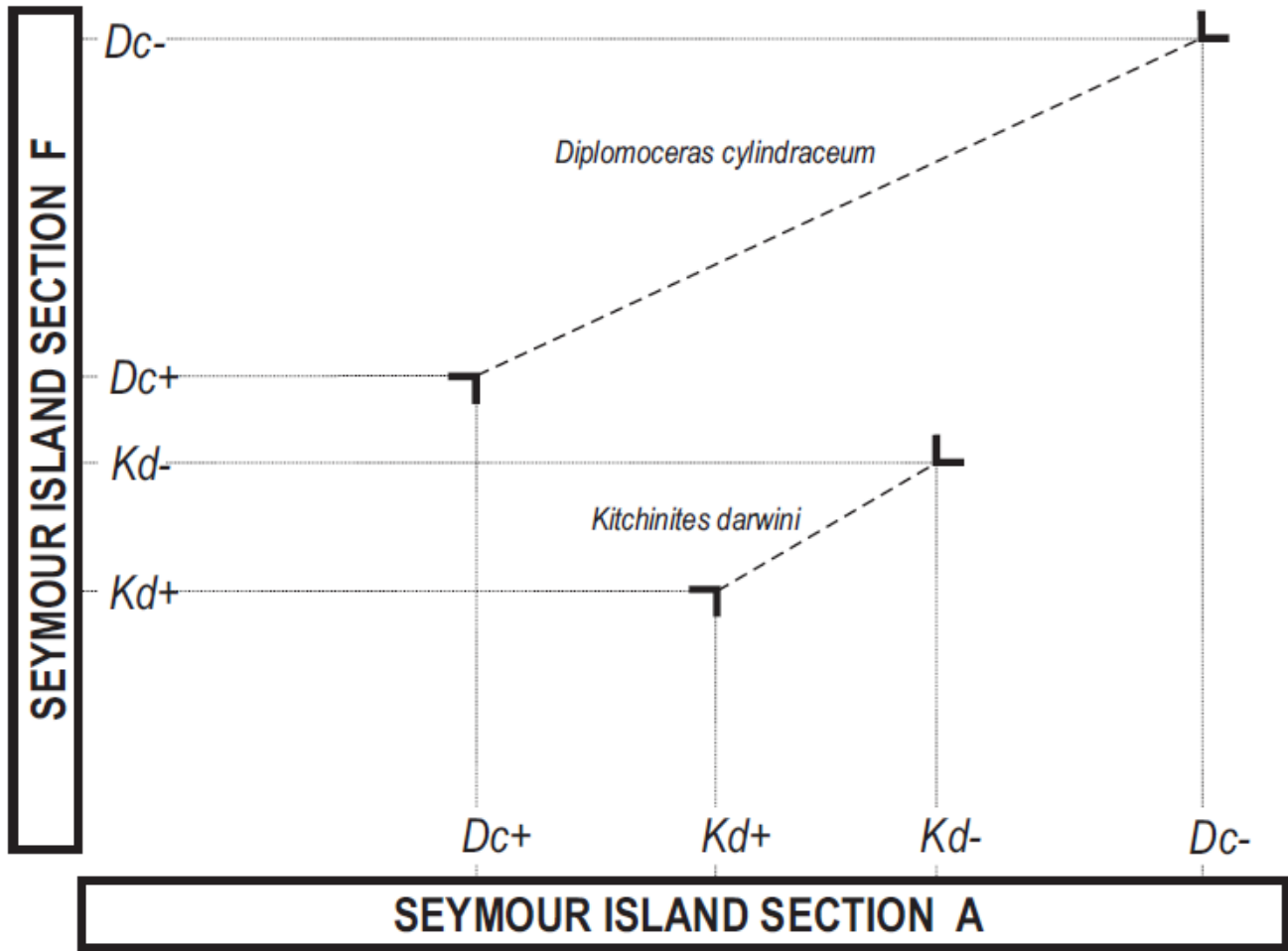


Figure VIII.3

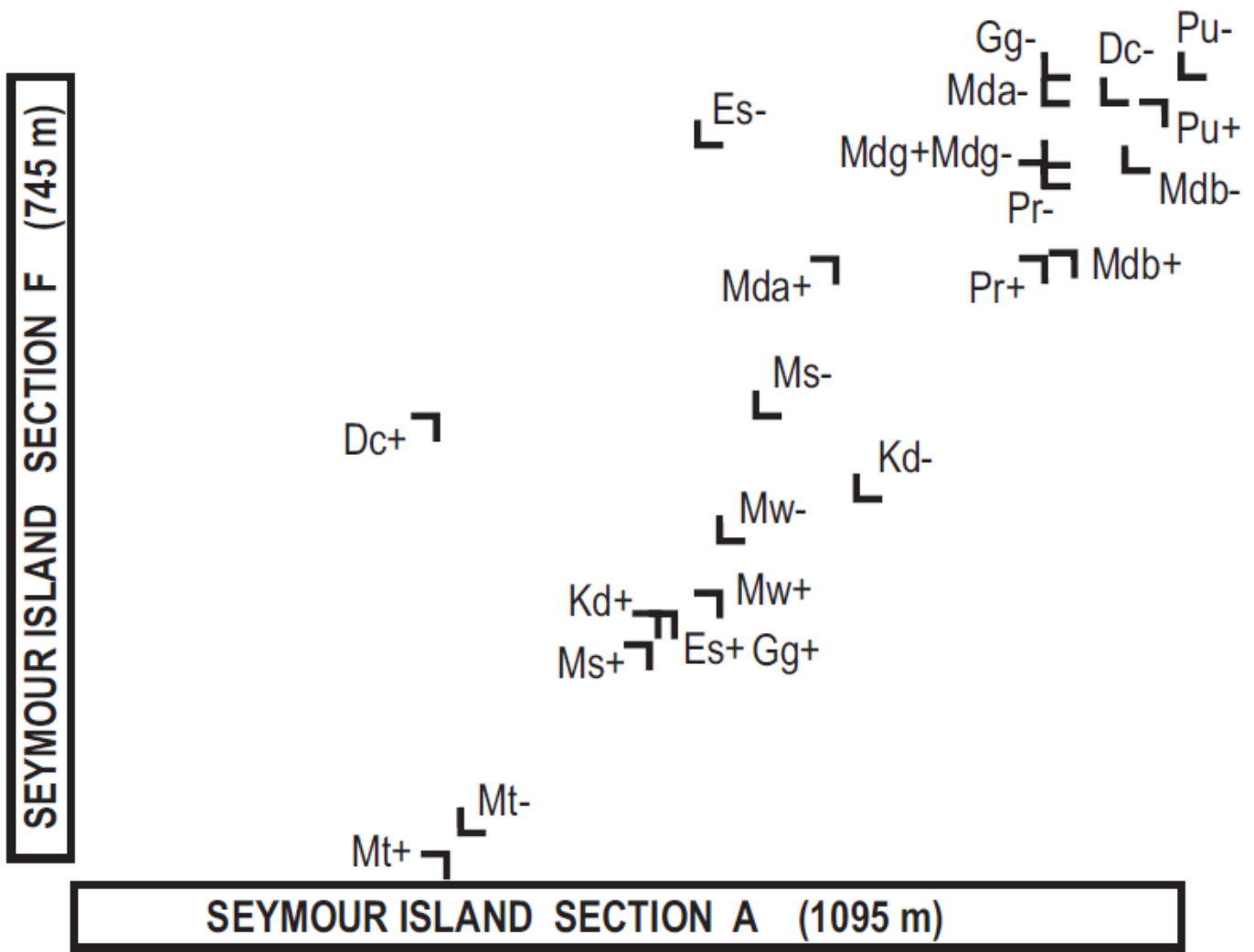


Figure VII.5

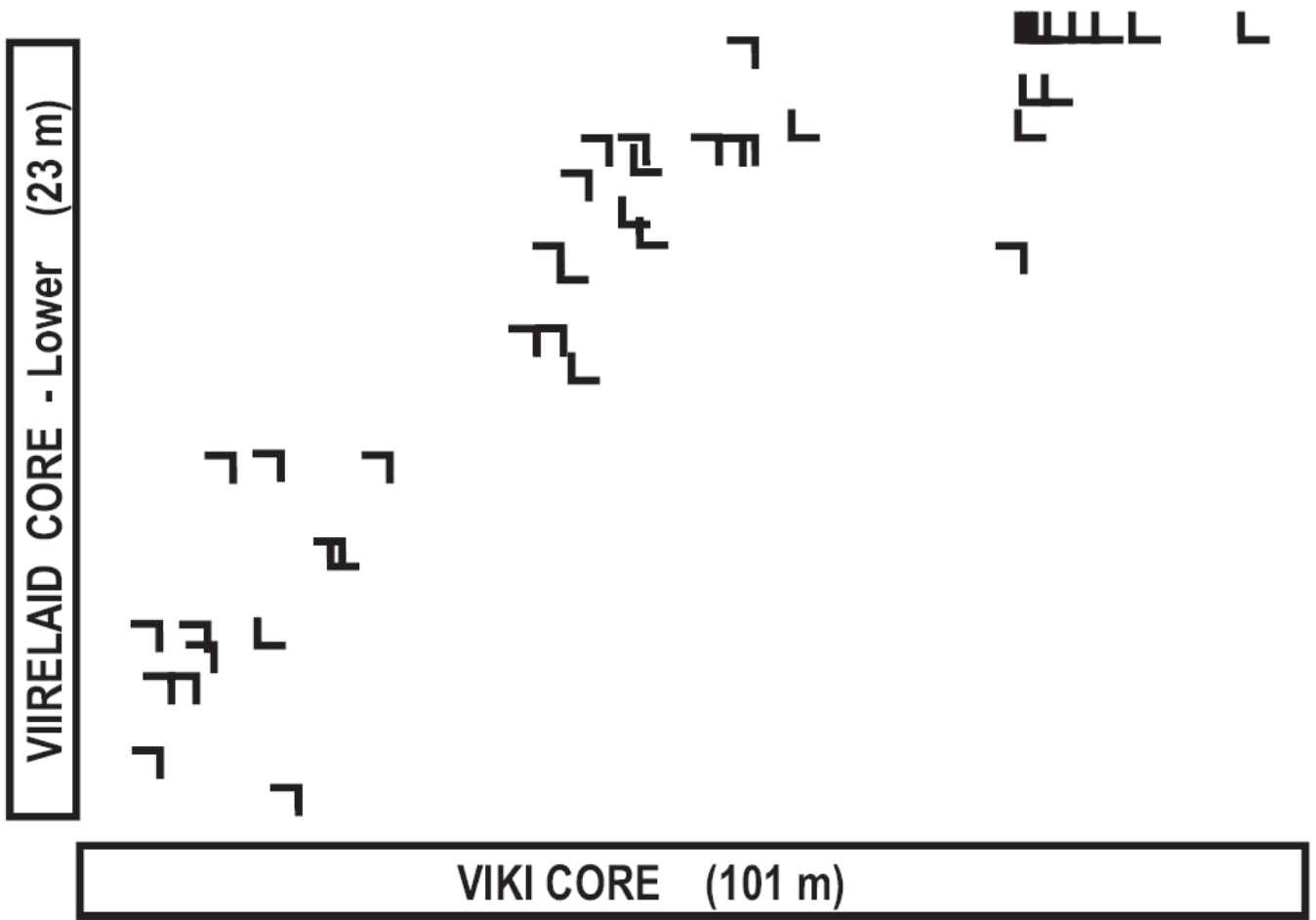


Figure VIII.2



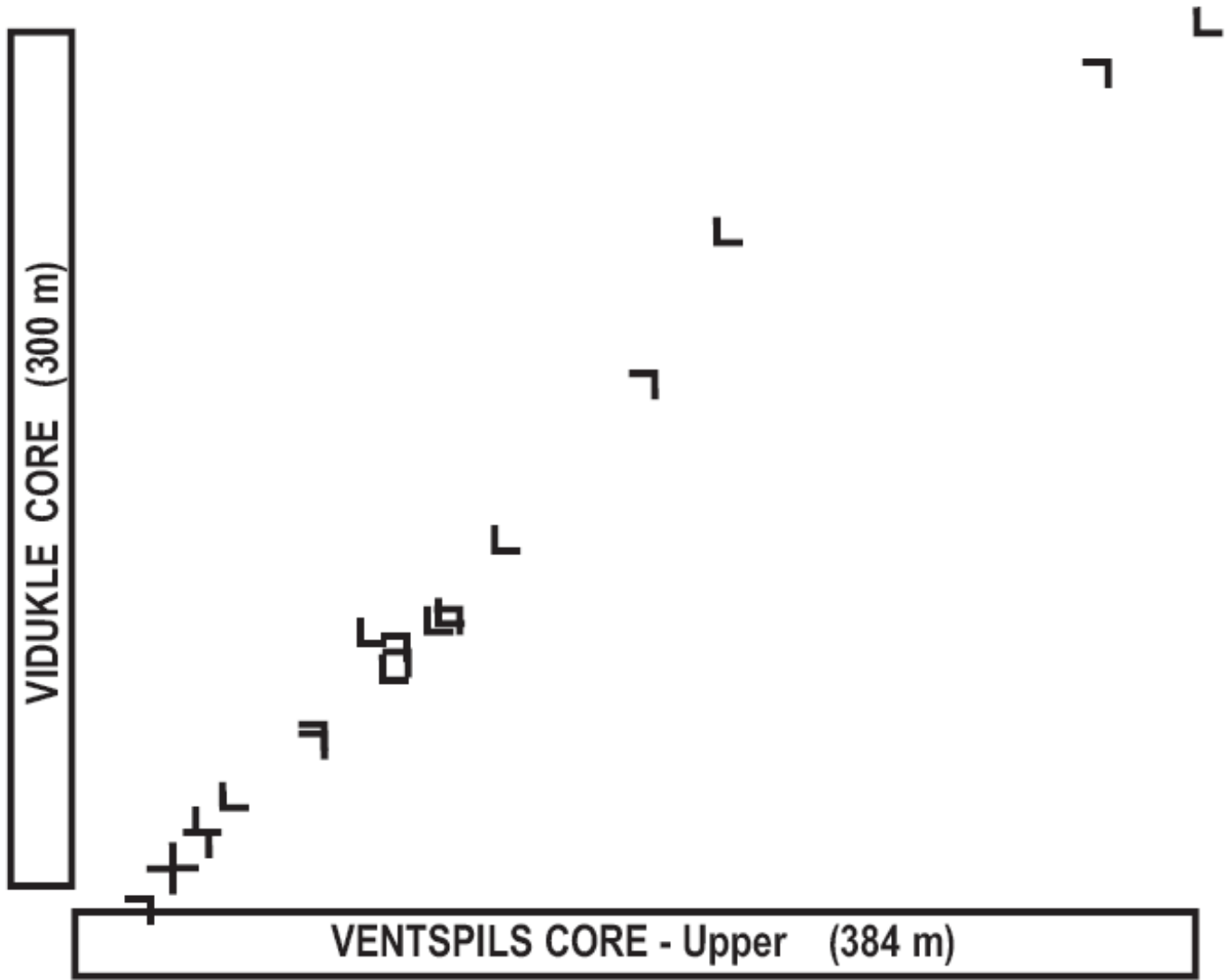


Figure VIII.3