Part IV

Modelling the Record

Chapter 16

Quantitative Methods for Applied Microfossil Biostratigraphy

FELIX M. GRADSTEIN
Geology Museum, University of Oslo, N-0318 Oslo, Norway.

> "Quantitative stratigraphy uses relatively simple or complex mathematical-statistical methods to calculate stratigraphic models that with a minimum of data provide a maximum of predictive potency, and include formulation of confidence limits", (F.P. Agterberg, 1990).

16.1 Introduction

Modern biostratigraphy frequently copes with occurrence data from hundreds of fossil taxa, in thousands of samples, derived from many wells or sections in many different basins. Two challenges, particularly on a regional, basinwide scale, are:

(a) to increase stratigraphic resolution in biozonations, using a combination of eventls derived from several different micro- and/or macrofossil groups, and

Applied Stratigraphy, edited by Eduardo A. M. Koutsoukos,

(b) to calibrate these zonations to modern time scales, and extract sequence stratigraphic signals that assist with seismic mapping of prospects, and burial history for basin modelling.

This study focuses on the first challenge. Conventional stratigraphic resolution places considerable emphasis on the end pioints of a few taxa in a few sections. Since a limited number of sections are likely to have the uniform presence or consistent order of all "zonal or index" taxa, there always is a fair amount of subjective judgments as to the perceived "true" order.

New tools in stratigraphy, using semi-quantitative or quantitative methods, make it easier to build integrated zonations, and individual wells or outcrop sections may be tested for 'stratigraphic normality'. Advantages of these tools in stratigraphy, particularly in frontier regions, are:

1. Standardization during digitization of the fossil record and execution of (semi-) objective stratigraphic methods gives easy access to all data and interpretations.
2. Data sets and results are easy to communicate and are rapidly updated with new information.
3. Integration of all fossil and also physical (e.g. isotope, well-log) events in one stratigraphic solution increases resolution and practical use.
4. Methods and results (zonation + correlation) are more objective than 'hand-made' solutions.
5. Zones, events and their correlations may have error bars attached, which improve insight into true stratigraphic resolution and reliability of event correlation.
6. Interpolation of missing event positions in sections increases detail in correlations.
7. Unlike subjective stratigraphy, the new methods provide more than one possible solution of the data, depending on run conditions (multiple working hypothesis).
8. Sequence stratigraphic levels or trends may be detected and defined.
9. The new methods handle large and complex data sets, and calculate reliable solutions quickly.

That is not to say that there are no limiations to quantitative biostratigraphy. For example, the fossil record cannot be modelled a priori for spatial and temporal distributions, and it is difficult to directly weigh records (observations) in terms of stratigraphic quality. In addition, the methods are time consuming because of demands on data organization and data formating. On the other hand, modern studies are showing that benefits of the quantitative approach outweigh limiations, and enhance the quality of geologic interpretations.

16.2 Properties of Stratigraphic Data

A paleontological record is the position of a fossil taxon in a rock sequence. The stratigraphic range of a fossil is a composite of all its records. The end-points of the range are biostratigraphic events, which includes the first appearance in time, and disappearance from the geologic record. A biostratigraphic event is the peresence of a taxon in its time context, derived from its position in a rock sequence. Fossil events are the result of the continuing evolutionary trends of life on earth; they differ from physical events in that they are unique, non-recurrent, and that their order is irreversible.

Often the first and last occurrences of fossil taxa are relatively poorly defined records, based on few specimens in scattered samples. Particularly with time-wise scattered last occurrences, reworking may have locally extended the record, which may be distinguished by differentiating between the last occurrence (LO), and the last common or last consistent occurrence (LCO) of taxa.

The spacing in relative time between successive fossil events is called resolution. The greater the probability that such events follow each other in time, the greater the likelihood that correlation of the event record models isochrons. Most industrial data sets make use of sets of LO and LCO events. In an attempt to increase resolution in stratigraphy, particularly when many sidewall cores are available, efforts are made to recognize a half dozen events along the stratigraphic range of a fossil taxon (Fig. 16.1), including last stratigraphic occurrence (‘top’ or LO event), last common or consistent occurrence (LCO event), last abundant occurrence (LAO event), first abundant occurrence (FAO event), first common or consistent occurrence (FCO event) and first occurrence (FO event). Unfortunately, such practice may not yield the desired increase in biostratigraphic resolution sought after, for reason of poor event traceability.

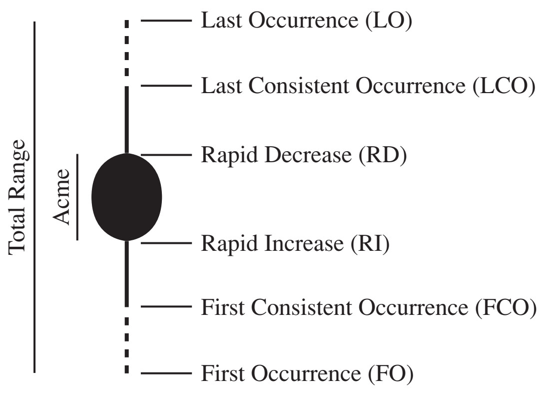


Figure 16.1 Terminology of biostratigraphic events along the total stratigraphic range of a single taxon. The first and last consistent occurrence frequently coincide with the first and last common occurrence; the rapid increase and rapid decrease encapsulate the acme of the taxon.

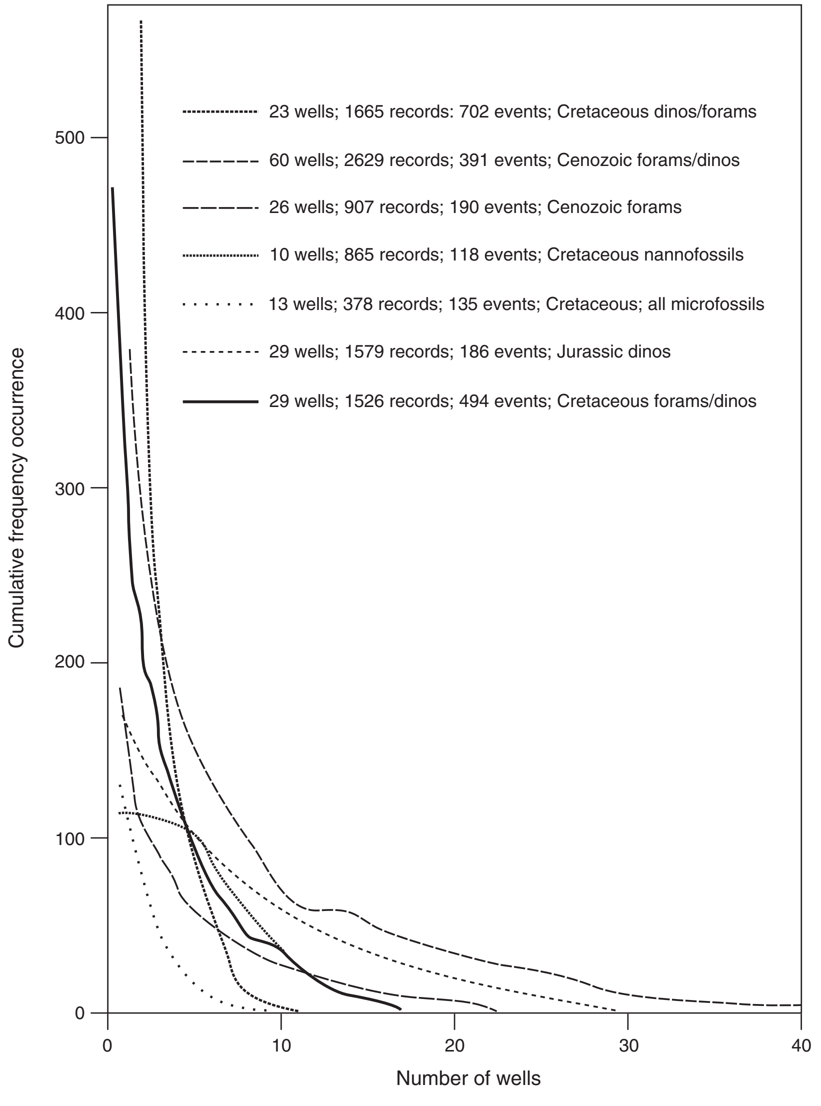


Figure 16.2 Cumulative frequency of microfossil event occurrences versus number of wells for seven subsurface data sets. Dots – 23 wells/702 events, Cretaceous of North Sea and offshore Norway, various consultants, unpublished; dashes – 60 wells/391 events (Gradstein et al., 1994; Gradstein and Bäckström, 1996, and unpublished); dot-dash – 26 wells/190 events, Cenozoic offshore eastern Canada (Gradstein et al., 1994); closely spaced dots – Upper Cretaceous nannofossils (with relatively high traceability, as seen from the hump in the curve), offshore East Canada (Doeven et al., 1982); pairs of dots – 13 wells/135 events, Indian and Atlantic Oceans Deep Sea Drilling Sites and Ocean Drilling Sites with relatively low fossil diversities and frequent hiatusses (Gradstein et al., 1992); small dashes – Jurassic dinoflagellates, Troll area, offshore southwest Norway (R. Woollam, pers. comm., 1994); solid line – 29 wells/494 events, Cretaceous foraminifers and dinocysts, offshore Norway (Gradstein et al., 1999).

Poor event traceability is illustrated in Figure 16.2, where cumulative event distributions are plotted using a wide variety of microfossils from different stratigraphic intervals in different basins. All curves are asymptotic, showing an inverse relation between event distribution and the number of wells. None of the events occur in all wells, and far fewer events occur in 5 or 6 wells than only in 1 or 2 wells; hence, the cumulative frequency drops quite dramatically with a small increase in the number of wells. Obviously, the majority of fossil events have poor traceability, which is true for most data sets, either from wells or from outcrops. Groups of microfossils with higher local species diversity, on average have lower event traceability.

Data sets with above average traceability of events are those where one or more dedicated observers have spent above average time examining the fossil record, verifying taxonomic consistency between wells or outcrop sections, and searching for 'missing' data. In general, routine examination of wells by consultants for drilling completion reports yields only half or (much) less of the taxa and events than may be detected with a slightly more dedicated approach.

There are other reasons than lack of details from analysis for why event traceability is relatively low. For example, lateral variations in sedimentation rate change the diversity and relative abundance of taxa in coeval samples between wells, particularly if sampling is not exhaustive, as with well cuttings or sidewall cores. Because chances of detection depend on many factors, stratigraphical, mechanical, and statistical in nature, increasing sampling and studying more than one microfossil group in detail is beneficial.

Although not always explicit, biostratigraphy relies almost as much on the absence, as on the presence of certain markers. This remark is particularly apt for microfossils that generally are widespread and relatively abundant, and compose many stratigraphically useful events. Only if non-existence of events is recognized in many, well-sampled sections, may absences be construed as affirmative for stratigraphic interpretations. If few samples are available over long stratigraphic intervals, the chance to find long-ranging taxa considerably exceeds the chance to find short-ranging forms. In actual practice, index fossils have a short stratigraphic range, are generally uncommon and hence easily escape detection. Therfore, interpretations based on absences should be used with caution.

16.3 Data Bookkeeping

An important aspect of quantitative stratigraphy is microfossil event input, and efficient bookkeeping of such records for many wells. The creation of datasets that provide meaningful stratigraphic answers is dependent on such. In the process, detailed checks are advisable to eradicate taxonomic errors and to remove gross outliers representing caved (in exploration wells) or geological recycled (reworked) events. There is no doubt that data input, data bookkeeping and data checking take most of the time in a project. Without a suitable computer program to digitize, organise, pre-digest and filter data of many wells or outcrop sections, such tasks can not be executed, and quantitative biostratigraphic methods cannot be applied. A key property of such a methodology is that all wells in a dataset are accessible simultaneously for queries, corrections, and modifications, so-called multi-well tasking. Both standard spreadsheet or relational database programs can be adapted to the task, and should offer some or all of the following options:

(a) bookkeeping and organisation of fossil events such that they can be queried simultaneously in all wells;
(b) calculation of simple census-type statistics;
(c) tracing of events over all well sites;
(d) finding of co-occurring events, synonyms and geographic substitutes;
(e) cross-plot events from two wells, or of well versus zonation to eliminate outliers;
(f) reformat files for direct input in quantitative stratigraphy programs;
(g) create subsets of the original data to verify local biozonal trends with selected taxa; and
(h) provide a complete printed record of all data, suitable for reports.

The curious anomaly in data processing is the lack of agreement on a standard format in which biostratigraphic data are processed and stored. This is a challenge that needs an urgent consensus.

16.4 Stratigraphic Methods

16.4.1 Introduction

Traditionally, biostratigraphic zonations are executed ‘by hand’ through a painstaking process of (mental) stacking in relative geologic time of numerous fossil events from many different outcrop or well sections. Subtle stratigraphic order relationships are evaluated, and frequent gaps are bridged by superpositional hypothesis, where data are scarce. The human mind is good at evaluating observed and virtual superpositional data and bridging data gaps.

Quantitative methods of biostratigraphy, like graphic zonation and correlation, or ranking and scaling cannot easily match the subtilities of very detailed subjective zonations, based on many, often incomplete stratigraphic sections, using much information on missing data. As mentioned above, the experienced biostratigrapher uses almost as much information on absence as on presence of data, and the former cannot be evaluated by a method. To produce a data set that is detailed and informative enough to yield quantifiable, high-resolution zonations is a considerable task. However, once such a dataset and its derived quantitative zonation is accomplished, and made available together with its raw data and data processing details, it serves as a more reliable model for correlation and chronostratigraphic calibration than a poorly documented, subjective zonation. True stratigraphic resolution improves if event spacing in relative time is assessed with standard deviations, that create an understanding as to the chance that two events are superpositional.

16.4.2 Deterministic and Probabilistic Methods

There are two principal families of quantitative stratigraphic methods: (a) deterministic, and (b) probabilistic. Deterministic methods seek the total or maximum stratigraphic range of taxa, whereas probabilistic methods calculate the most probable or average range (Fig. 16.3), accompanied by an estimate of stratigraphic uncertainty. Deterministic methods assume that inconsistencies in the stratigraphic range of a taxon from well to well are due to missing data. On the other hand, probabilistic methods assume that the inconsistencies are the result of random deviations from the most commonly occurring or average stratigraphic range. Athough this concept is relatively foreign to conventional biostratigraphy, the large (and often noisy) body of local range data for microfossils in many different basins makes this concept attractive for exploration biostratigraphers in particular.

Hood (1995) evaluated the use of average composite sections from graphic correlation, showing a model of taxon first occurrence with a localized speciation event, with delayed migration into different environments in a basin (Fig. 16.4; see also figure 49 in Thierry, 1997). The shape of the 'average' first appearance emphasises the difference between the use of maximum and average event positions for realistic zonations and

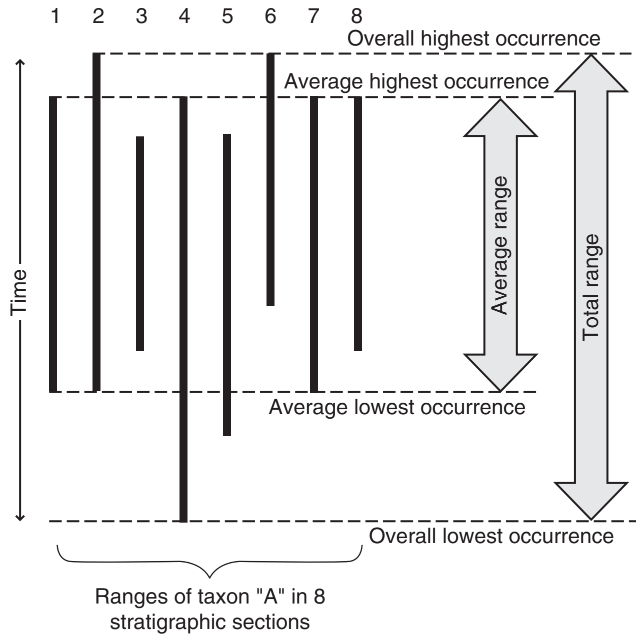


Figure 16.3 Deterministic biostratigraphy tries to find the total range of a taxon, whereas probabilitistic methods seek the average stratigraphic range. The latter may have an estimate of uncertainty attached that is a function of the spread in local ranges of the taxon in the eight sections examined (after Cooper et al., in press).

AQ: Please update

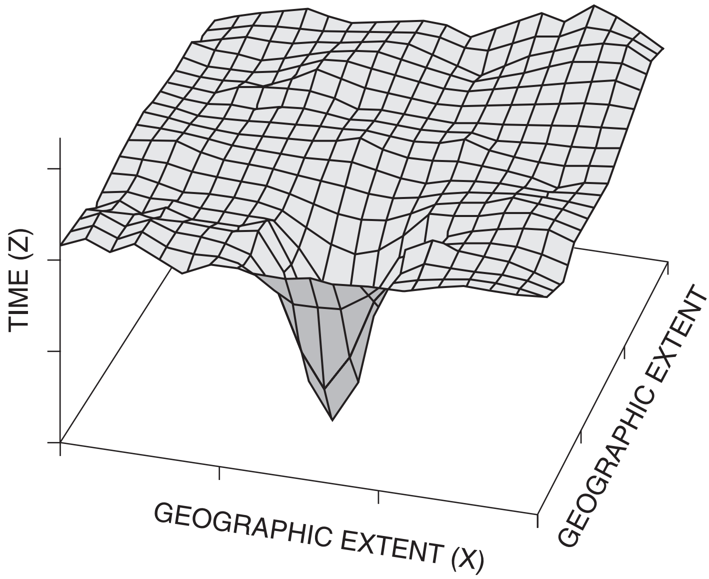


Figure 16.4 Model of the first occurrence of a taxon with a localized speciation event, with delayed migration into different environments in a basin, that may be modelled with an average first occurrence surface (redrawn after Hood, 1995).

correlations. Similar figures may be drawn for average last occurrence distributions of fossils.

Deterministic methods are traditionally simple, and most suitable for data sets of few wells or outcrop sections; they lack error analysis, are slow to execute, and sensitive to geological reworking and poor sampling. Probabilistic methods have a mathematical basis, and may be more complex, but have detailed error analysis, execute quickly, and are less sensitive to reworking of taxa or incomplete sampling. A prerequisite for both methods is good data input, and good data organisation, with probabilistic methods more effective with larger data sets. Characteristics of programs in both categories of methods are in Table 16.1.

16.4.3 Graphic Correlation

Among deterministic methods, Graphic Correlation is best known. Graphic Zonation and Correlation, also called Shaw's Method (Shaw, 1964) has become accepted and used by academic and industrial biostratigraphers as a simple, semi-objective tool to assess the fossil record for zonation and correlation purposes. This is actually a type of "crossplot" method, where a comparison is made between order and spacing of stratigraphic events in pairs of sections, using bivariate scatterplots. First, one of the stratigraphically more complete sections is selected as the reference, and a second section is crossplotted with it, using the events common to both. On one axis the events are in the order (and spacing if so desired) they have in the reference section, and along the other axis they are in the order (and spacing) that they have in the compared section. The line of correlation (LOC) between two sections is derived either from subjectively connecting points (events) in common in the two-way scattergrams, or from a statistically modeled best fit line between the scattered points. Traditionally, stratigraphers have used either single straight lines, or segmented straight lines as LOC. Using the LOC, the original section is interpolated with the second section, using a few simple rules, to produce a composite sequence. The composite of the two is then crossplotted with a third section, and again interpolated. This procedure is repeated for all (well) sections, and the final solution is called the composite standard. This composite standard is a handy and detailed zonation, expressed in composite standard units that may be used to interpret the sequence history of wells, not unlike conventional geohistory.

For theory and applications with graphic zonation and correlation the excellent book by Mann and Lane (eds., 1995) is recommended, as well as the elegant study with Ordovician graptolite data of Cooper and Lindholm (1991). The latter gives a clear illustration of the principle and utility of the method for closely controlled, high-resolution datasets (Fig. 16.5). The authors created a standard graptolite zonation with 45 FA (first appearance) and 45 LA (last appearance) events of 90 taxa, using fourteen closely sampled and closely studied sections in Australia, Texas, NW Canada, Newfoundland, S. Sweden and S.E. Norway. The Australian sequence, the richest and best known was taken as the initial reference section to be composited with the order of the other thirteen ones. Figure 16.5A shows a plot of the chosen reference section against the S. Sweden section. Figure 16.5B is a plot of the final composite standard sequence against the original Australian section, revealing the extent to which the Australian section has been modified by incorporation of the order of events in other sections. In positioning the LOC, the authors gave weight to events based on species that are relatively distinctive, relatively abundant, and relatively short ranging. The latter minimizes awkward ‘unfilled range’ situations, and allows plotting the data (by hand) on straight line segments.

16.4.4 Constrained Optimization

Recently, a new method has appeared that overcomes some disadvantages of graphic zonation and correlation. It is called CONOP (*con*strained *op*timization), and was designed by Kemple *et al.* (1995), with P.M. Sadler (University of California, Riverside) doing further development. A recent stratigraphic application is by Cooper *et al*. (2001; see below). As in graphic correlation, order and thickness spacing of events in sections are used, but the method is multi-dimensional in the sense that it treats the observations in all sections simultaneously. Like in RASC (see below) it can complete the task of sequencing (the ranking problem), before the task of scaling (the spacing problem).

Table 16.1 Properties of principal methods in quantitative biostratigraphy used in exploration biostratigraphy (modified after Cooper et al., in press). Together with their manuals, the programs GRAPHCOR, STRATCOR, CONOP, RASC and CASC provide professional and academic biostratigraphers with rapid and versatile tools to organise, explore and interpret biostratigraphic data for zonation and geological correlation, with estimates of uncertainty

Graphic correlation	*Constrained optimization*	*Ranking & scaling*
Programs GRAPHCOR, STRATCOR	Program CONOP	Programs RASC & CASC
Deterministic method: graphic correlation in bivariate plots (note: STRATCOR program can operates in a probabilistic or deterministic manner)	Mostly a deterministic method; can also simulate probabilistic solutions Constrained optimisation with simulated annealing and penalty score	Probabilistic method: ranking, scaling normality testing, and automated, most likely graphic correlation with error analysis
Uses event order and thickness spacing; works best with datasets having both first and last occurrences of taxa	Uses event order, event cross-over, and thickness spacing; datasets best have both first and last occurrences of taxa	Uses event order from well to well, and scores of cross-over from well to well for all event pairs
Best suited for small data sets, but can operate also on large datasets	Processes medium to large data sets	Processes large data sets fast; has data management and data input module
Requires selection of an initial standard section, then section by section comparison with the intermediate composite in repeated rounds	Treats all sections and events simultaneously, and works inverse through iteratively improved 'guesses' about the solution	Treats all sections and events simultaneously
Line of Correlation (LOC) fitting, in section by section plots, can be partially automated	Multidimensional LOC; automated fitting; can generate several different composites depending on the many run options	Automated execution; generates several scaled optimum sequences per dataset depending on run parameters, and tests to omit 'bad' sections or 'bad' events

Attempts to find maximum stratigraphic range of taxa among the sections	Attempts to find maximum or most common stratigraphic ranges of taxa	Finds average stratigraphic position of first and last occurrence events
Builds a composite of events by interpolation of missing events in successive section by section plots, via the LOC	Uses simulated annealing to find either the 'best' or a good multidimensional LOC and composite sequence of events	Uses scores of order relationships to find the most likely order of events, which represents the stratigraphic order found on average among the sections
Relative spacing of events is a composite of original event spacing in meters in the sections	Relative spacing of events in the composite is derived from original event spacing in meters or sample levels	Relative spacing of the events in the scaled optimum sequence derived from pair-wise cross-over frequency
No automatic correlation of sections; can be used to build time scales	**Correlates sections automatically; can be used to build a standard time scale**	**Automated correlation of sections using isochrones**
No error analysis; sensitive to geological reworking and other 'stratigraphic noise', and sensitive to order in which sections are composited during analysis	Numerous numerical tests and graphical analysis of stratigraphic results; finds best break points for assemblage zones	Three tests of stratigraphic normality of sections and events; calculates standard deviation of each event as a function of its stratigraphic scatter in wells
Interactive operation under DOS; graphic displays of scattergrams and best fit lines	Batch operation under windows; colour graphics displays shows progress of runs	Button operated under windows, fast batch runs; colour graphics of output and options for interactive graphics editing

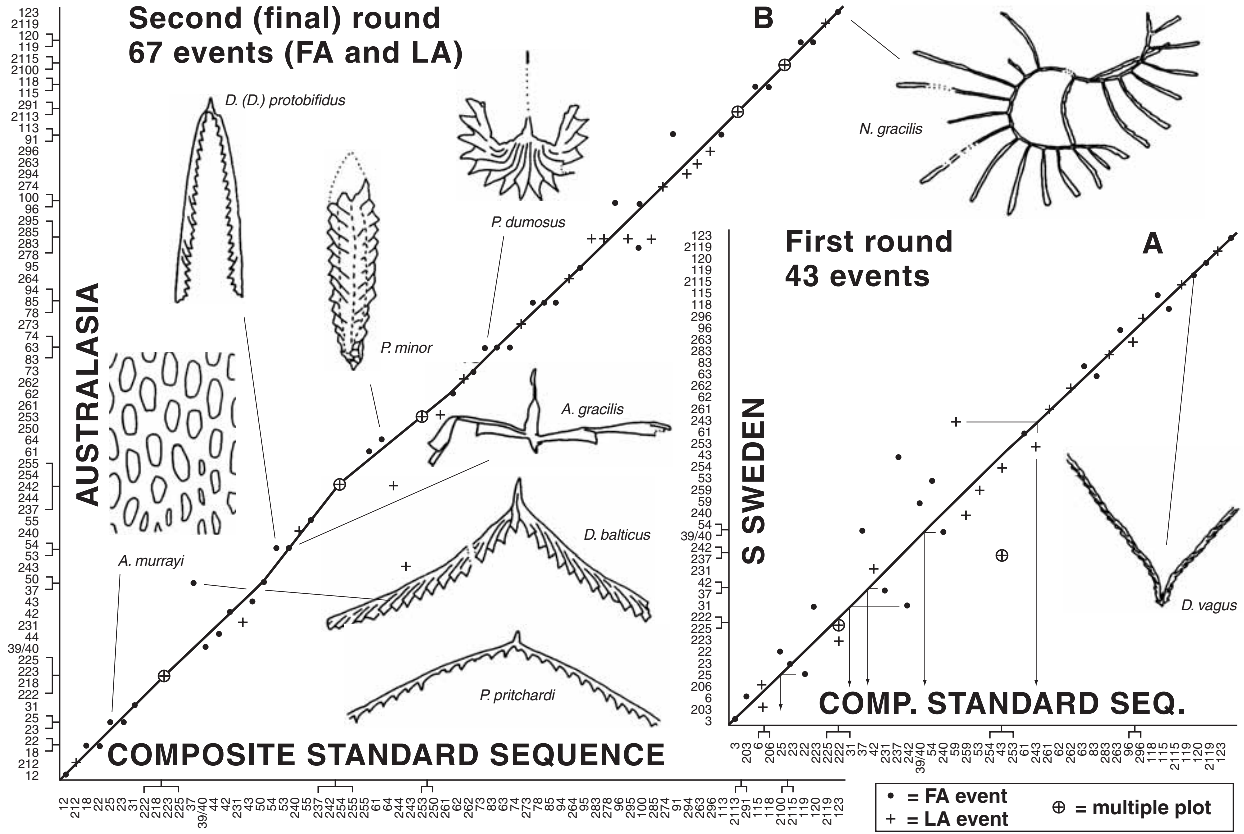


Figure 16.5 (A) First round plot of composite standard sequence standard sequence and sequence of S. Sweden, using Ordovician graptolites; the line of correlation is the diagonal, and unfilled range events are relocated as shown. (B) Final (second) round of graphic correlation of composite standard sequence and Australian sequence with a segmented line of correlation fitted (after Cooper and Lindholm, 1991).

Instead of building a solution from the data, CONOP works through a series of iteratively improved guesses about the solution. Each guess is compared with the data; the misfit between the solution and data guides the next guess. Geophysicists like to call the process 'inversion'. Unfortunately, solution time increases as 2^N, where N is the number of events. This means that for an exhaustive search of e.g. 124 events (LO and FO of 62 taxa), the searching time becomes impossible. In order to find a good solution without waiting 'forever' CONOP uses a version of the simulated annealing algorithm, using heuristic search techniques; such a technique is incapable of proof, but serves to guide to acceptable solutions.

The method is constrained in that it eliminates impossible solutions (constraint), and then searches for the best of all the possible ones (optimization). The method may be thought of as fitting a multidimensional line of correlation (LOC) simultaneously to all points in all sections. The composite 'true' section of events is that hypothetical sequence of ordered and spaced events that causes the least net disruption or penalty when the ranges of taxa in each of the well sections are adjusted to match it. Like graphic correlation, the observed tops of species in individual well sections are extended stratigraphically upwards, and bases downward to achieve a best fit. In this sense, penalty represents a measure of inconsistency of individual tops or bases among the well sections, and is expressed in meters. This penalty resembles that used in method STRATCOR (Gradstein, 1996) that keeps track of the cumulative amount (distance) over all wells that events shift from their observed position to their interpolated one. CONOP has a host of other features, and presently builds a Lower Paleozoic conodont-graptolite composite that assists with the construction of a detailed geologic time scale (R. Cooper, pers. comm., 2000).

16.4.5 Ranking and Scaling

The principal method of probabilistic biostratigraphy is called Ranking and Scaling (RASC; Agterberg and Gradstein, 1999; Gradstein *et al.*, 1985, 1999). The many options in RASC method of biostratigraphy are listed in Table 16.2. During the last two decades, RASC has been applied to a wide variety of datasets involving many tyoes of microfossils. A majority of applications are with well data sets from industry or scientific ocean drilling. Published literature on and with the method is extensive, and is listed in the literature cited.

Unlike graphic correlation, the RASC method considers the stratigraphic order of all fossil events in all wells simultaneously, and calculates the most likely (optimum) sequence of events. In this sequence, each event position is an average of all individual positions encountered in the wells.

Ranking is based on superpositional relationships between events. In general, there are three possible types of superpositional relationships for a pair of events co-occurring in the same section. An event can be observed to occur above or below another event, or the two events coexist in the same sample. In the ranked optimum sequence, which is based on a large number of sections, two events can be coeval on

Table 16.2 Products of the RASC and CASC programs for probabilistic stratigraphy

PROGRAM RASC – zonation, variance analysis and normality testing

- Bookkeeping
  - Value of input parameters
  - Sequence of wells
  - Tabulation of event records, using frequency and cumulative frequency
  - Summary of RASC run results (vital statistics)
  - Dictionaries of events – numerical and alphabetic listings
  - Occurrence table of events in all wells
- Ranking and Scaling (= probabilistic zonation)
  - Optimum sequence of events, with option to insert variances and unique (rare) events
  - Final scaling of optimum sequence of events, with option to insert unique events
- Normality 'testing' of event record in wells
  - Graphical correlation of well sequence record and optimum sequence, with estimation 'how far' events are off best fit line (cubic spline)
  - Step model per well, with penalty points for out of place events
  - Rank correlation of event well sequences with (scaled) optimum sequence
  - Normality test per well, with second order difference statistics for all events
  - Comparison of observed and expected second order difference values
- Variance analysis
  - Standard deviations of events per well
  - Event variance analysis (difference in each well between observed and stratigraphically expected event position + frequency distribution)
  - Summary of event variance analysis results
  - Estimation of event ranges – numerical and graphical representation of probable minimum–, probable maximum –, and average observed stratigraphic event positions
  - Estimation of event cross-over ranges

PROGRAM CASC – correlation of RASC zonation, with flattening option in graphics displays

- Probable position in wells of optimum sequence events, with 95% confidence limits
- Observed event positions with 95% confidence limits

the average when one of them occurs exactly as many times above the other one as it occurs below it. If an event is observed above another event in some sections but below it in others, a stratigraphic inconsistency involving these two events is indicated. The purpose of constructing the optimum sequence is to utilize and resolve such inconsistencies. In fact, there is no point in applying RASC if inconsistencies are missing when sections are compared with one another. Lines of correlation connecting observed positions of events in sections show cross-overs when there are inconsistencies, which is a normal event feature (see Fig. 16.6).

Scaling of the optimum sequence in relative time provides information on the stratigraphic clustering of events, and is a function of the frequency with which events in each each pair in the RASC optimum sequence cross-over their relative positions (observed records) from well to well. The more often any two events cross-over, the smaller their interfossil distance. Final distance estimates are expressed in dendrogram

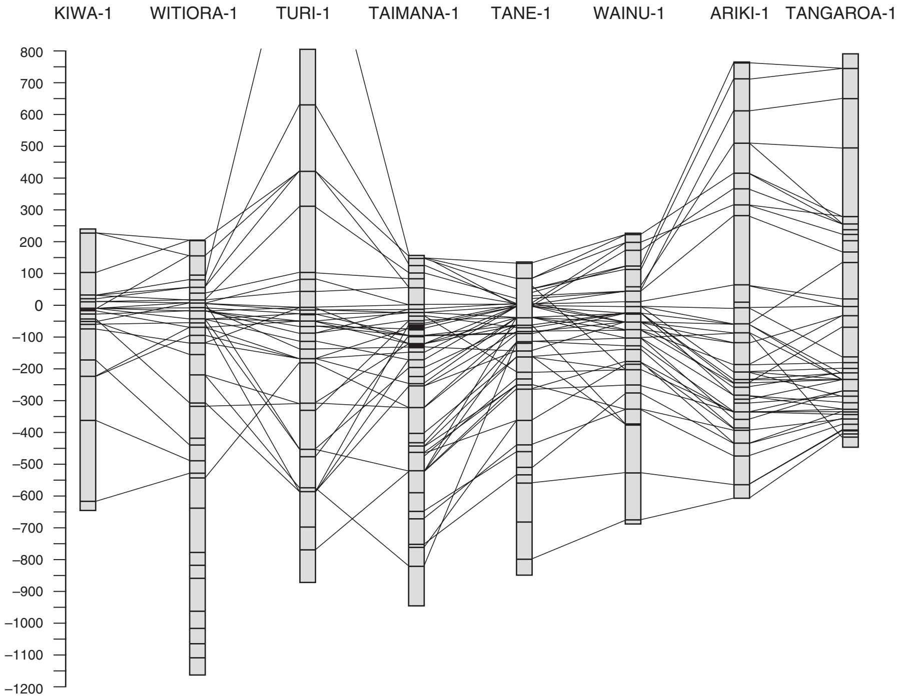


Figure 16.6 Direct correlation of the events in the CONOP Composite Sequence of Figure 16.7, based on the observed depth of the range end events, produces an intricate network with the usual cross-overs and mismatches of conventional well correlations.

format, where tightness of clustering is a measure of nearness of events along a stratigraphic scale. The scaled version of the optimum sequence features time successive clusters, each of which bundles distinctive events. Individual bundles of events are assigned zonal status. The process of zone assignment in the scaled optimum sequence is subjective, as guided by the stratigraphic experience of the users. Large interfossil distances between successive dendrogram clusters agree with zonal boundaries, reflecting breaks in the fossil record due to average grouping of event extinctions. Such extinctions occur for a variety of reasons, and may reflect sequence boundaries. From a practical point of view it suffices to say that taxa in a RASC zone on average group close together in relative time.

16.4.6 Variance Analysis

Anyone that tries to apply an event correlation framework from a zonation, quickly notices that closely spaced events tend to cross-over between wells (Fig. 16.6), indicative of some kind of uncertainty in event sampling and stratigraphic position. One way to rationalize such uncertainty is to use RASC zonations with variances, that identify more reliable stratigraphic markers (D'Iorio and Agterberg, 1989; Gradstein and Agterberg, 1998). The principle of variance calculation is straightforward. Individual well sequences are compared to the scaled optimum sequence using bivariate curve fitting to obtain the sum of differences between the observed and expected values of the events. The expected values are on the best fit line, meaning that an event would have zero variance. Once the differences between the observed and expected values are compiled for all RASC optimum sequence events in all wells, frequency distributions can be estimated. If an event is close to all lines of correlation considered it has a relatively small standard deviation; this means that it is a relatively good marker, as will be demonstrated for *Subbotina patagonica* in the lower Eocene, North Sea.

Graphical representation of differences between observed and expected positions at the well locations, such as on a map, may show that the large variance of an event is due to transgressive behaviour. Thus, variance analysis can be useful for tracking the time transgressive behaviour of events, as illustrated below.

RASC features two more tests to determine if the individual well record differs from the most likely zonation, the Stepmodel and the Normality tests (Table 16.2), both described in the literature cited. These tests, like variance analysis guide the user to outliers due to event misidentification, reworking or sample contamination, and to 'good and bad' wells.

16.4.7 Correlation and Standard Error Calculation

The preceding chapters dealt with zonation of fossil events, and tests to discern "good and bad" events and "good and bad" wells, using the RASC method. A companion method performs geologic correlation of RASC events and standard error calculation, and is named CASC, for *C*orrelation *a*nd *S*tandard Error *C*alculation. The technique is an extension of so-called "graphic correlation", as discussed in more detail in Gradstein *et al.* (1985), and in Agterberg (1990). Uncertainty limits (error bars) are calculated for both the most likely–, and the observed event positions in the wells or outcrop sections. Well correlation diagrams are both displayed in numerical and in colour graphics format. A detailed application in the Lower Cretaceous subsurface of the Grand Banks, eastern Canada, that also converts the RASC optimum sequence based on foraminifers to a RASC timescale and correlates isochrons, is in Williamson (1987); a recent application of CASC on a large Cretaceous dataset with many events using dinoflagellate cysts and foraminifers, offshore Norway is in Gradstein *et al.* (1999).

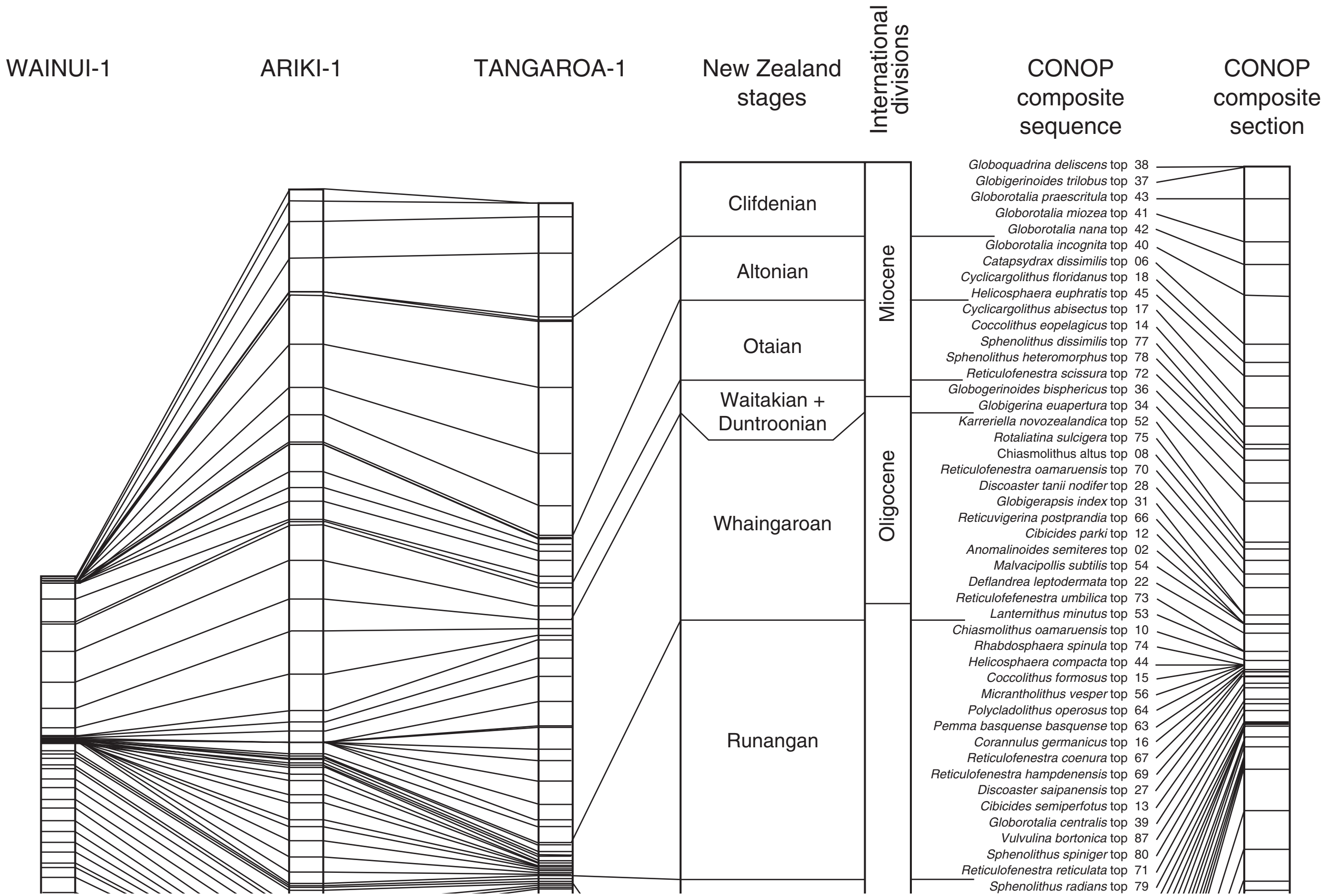
16.5 Stratigraphic Applications

16.5.1 Constrained Optimization: Taranaki Basin

The Taranaki Basin, New Zealand's producing hydrocarbon province, contains a highly fossiliferous Upper Cretaceous–Cenozoic sedimentary succession, resting unconformably on an erosional surface of varied relief that cuts across a 'basement' of Paleozoic and Mesozoic rocks. The basin has a complex depositional history with depositional breaks, condensed intervals, contemporaneous faults and folds, and lateral facies changes. Biostratigraphy from over 80 wells is an essential tool to interpret depositional history. Cooper *et al.* (2001) developed a detailed statistically based biozonation scheme for the purpose of increasing stratigraphic resolution, and assessing depositional rates across the basin. In addition, the team of stratigraphers wanted to evaluate the relative merits of deterministic and probabilistic approaches to quantitative biostratigraphic subdivision and correlation, using the methods outlined above: CONOP, RASC, and GRAPHCOR.

In all, the dataset of choice comprises 8 wells, from which 351 usable range tops of 351 foraminifers, nannofossils and palynomorphs were extracted from early Paleocene through early Miocene ages. In order to emphasize events that have correlation potential, and to calculate event variances, those events that occur in fewer than four wells were removed from the dataset. However, the methods allow unique events back in the analysis, such as index fossils or local marker horizons found in fewer than 4 wells. The dataset was thus reduced to 178 events in the 8 wells, with first stages of analysis leading to removal of 91 more events as being highly inconsistent in position from well to well, or having tops elsewhere above the youngest level sampled in the wells. The final dataset hence comprises 87 events, with 508 records. Direct correlation of the events in the CONOP composite, based on the observed depth of the range-end tops, produced an intricate network with the usual cross-overs and mismatches (Fig. 16.6) of conventional well correlations. In Figure 16.7, the CONOP ordinal and scaled composite sequences are displayed; the same figure also shows correlation of the calculated zonation to the regional scheme of stages using key markers, and correlation of the zonation through three wells with interpolated depths, like in graphic correlation. Note that the scaled composite section (right column of Fig. 16.7) is in arbitrary units based upon interpolation and extrapolation of the stratigraphic thicknesses of the interevent-units in all well sections.

The main conclusion of the study is that the RASC probable sequence (not shown) and the CONOP composite sequence are remarkably similar, and the two compare well with classical graphic correlation that cannot be executed in automated batch mode. Several stratigraphically promising species events were detected, not generally used for conventional biostratigraphy. The CONOP composite section gives the best estimate of the 'true' stratigraphic tops of taxa, based on the 8 well sections. It is consistent with the aims of conventional biostratigraphy based on range-end events, which are to establish zonations and correlations schemes based on the (maximum) ranges of species. Hence, it relates best to the conventional regional stratigraphy of New Zealand. The RASC scaled optimum sequence, on the other hand, gives the most probable order of events and its spacing and is particularly useful as a predictive zonation and correlation tool for future exploration drilling in the basin. The probabilistic and deterministic techniques are experienced as complimentary in order to best understand biostratigraphic potential of a dataset.


WAINUI-1
ARIKI-1
TANGAROA-1
New Zealand stages
International divisions
CONOP composite sequence
CONOP composite section
Clifdenian
Altonian
Otaian
Waitakian + Duntroonian
Whaingaroan
Runangan
Miocene
Oligocene
Globoquadrina deliscens top 38
Globigerinoides trilobus top 37
Globorotalia praescritula top 43
Globorotalia miozea top 41
Globorotalia nana top 42
Globorotalia incognita top 40
Catapsydrax dissimilis top 06
Cyclicargolithus floridanus top 18
Helicosphaera euphratis top 45
Cyclicargolithus abisectus top 17
Coccolithus eopelagicus top 14
Sphenolithus dissimilis top 77
Sphenolithus heteromorphus top 78
Reticulofenestra scissura top 72
Globogerinoides bisphericus top 36
Globigerina euapertura top 34
Karreriella novozealandica top 52
Rotaliatina sulcigera top 75
Chiasmolithus altus top 08
Reticulofenestra oamaruensis top 70
Discoaster tanii nodifer top 28
Globigerapsis index top 31
Reticuvigerina postprandia top 66
Cibicides parki top 12
Anomalinoides semiteres top 02
Malvacipollis subtilis top 54
Deflandrea leptodermata top 22
Reticulofefenestra umbilica top 73
Lanternithus minutus top 53
Chiasmolithus oamaruensis top 10
Rhabdosphaera spinula top 74
Helicosphaera compacta top 44
Coccolithus formosus top 15
Micrantholithus vesper top 56
Polycladolithus operosus top 64
Pemma basquense basquense top 63
Corannulus germanicus top 16
Reticulofenestra coenura top 67
Reticulofenestra hampdenensis top 69
Discoaster saipanensis top 27
Cibicides semiperfotus top 13
Globorotalia centralis top 39
Vulvulina bortonica top 87
Sphenolithus spiniger top 80
Reticulofenestra reticulata top 71
Sphenolithus radians top 79

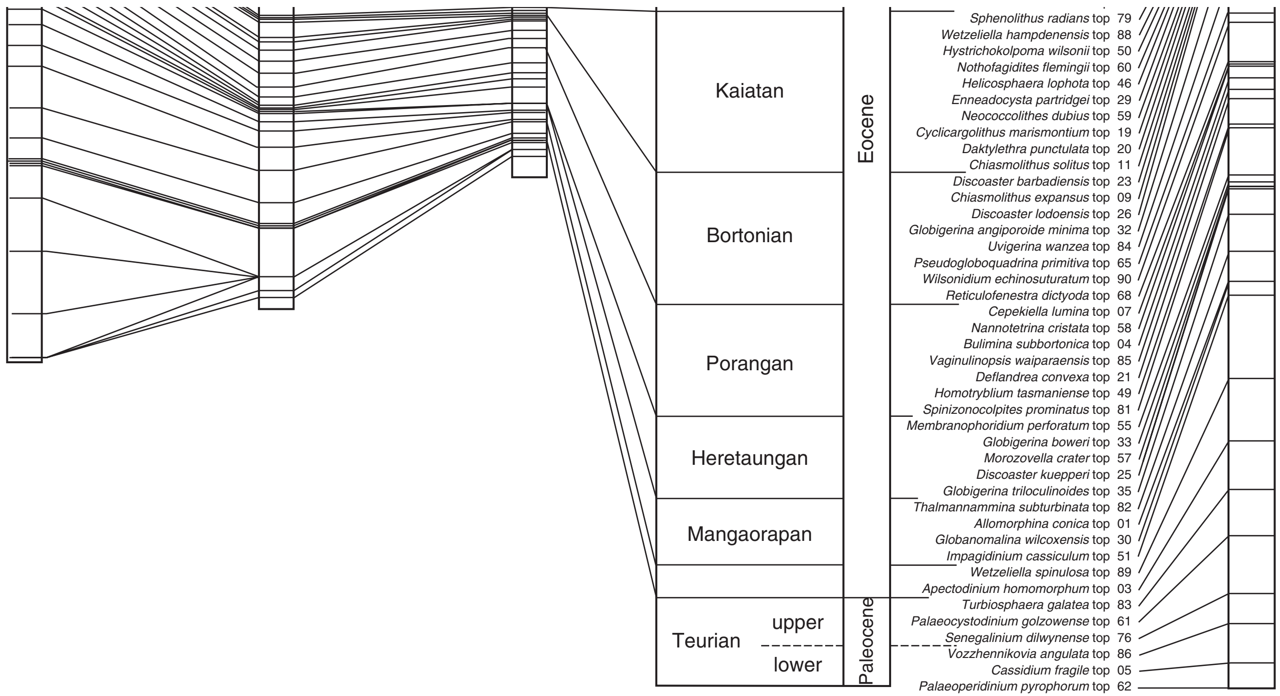


Figure 16.7 The CONOP Composite Sequence, both ordinal and scaled, its correlation to the regional scheme of stages using key markers, and correlation of the zonation through three wells with interpolated depths, Paleocene to Miocene of the Taranaki Basin, New Zealand. Each event has a dictionary number behind it.

16.5.2 Ranked and Scaled Optimum Sequences with Variances: North Sea

The North Sea region, a prolific petroleum province, contains remnants of stratigraphically superimposed sedimentary basins of Late Paleozoic through Cenozoic age, like stacks of half eaten pancakes. The regional history is complex; differential subsidence and uplifts are related to extensive mobilization of the North Atlantic rift systems. Deeper water, bathyal sediments, including minor and major gravity flow, siliciclast wedges, of middle Cretaceous through Paleogene age, are widespread and contain diversified agglutinated benthic foraminifera assemblages (Jones, 1988; Gradstein and Bäckström, 1996). In the southern part of the central North Sea, where deep water conditions prevailed into Miocene, the (DWAF) assemblage accordingly extends stratigraphically upwards (see below). The assemblages assist with biostratigraphy and paleobathymetry in exploration and exploitation wells, and DWAF taxa are prominent in the regional RASC zonation, as shown below.

The large scale deposition of basaltic ash (Balder unit) during earliest Eocene coincides with the eruption of major flood basalts in eastern Greenland and Rockal, at the onset of seafloor spreading in the Norwegian Sea. The ash is a prominent North Sea seismic reflector. Due to the flood-basalt outpourings, the North Sea became restricted, as reflected in the widespread distribution of diatoms, including the pyritized pillbox *Fenestrella antiqua* (Rank position 69 in Fig. 16.8, and Zone NSR3 in Fig. 16.9), and virtual absence of bottom fauna in the severely dysaerobic basin. Surface water salinity may have been abnormally low.

Correlations between the onshore NW Europe and North Sea Basin succession in the Paleocene and Eocene is achieved by dinoflagellate cyst biostratigraphy, integrated with the biostratigraphy provided by the calcareous plankton (foraminifera and nannoplankton) benthic foraminifera, magnetostratigraphy and volcanic ash stratigraphy. In this way a correlation network has been established over NW Europe, which serves as the background against which the probabilistic zonation was developed shown below. This probabilistic zonation serves as a template for range charts of DWAF in the petroleum basin.

Figure 16.8 shows the RASC optimum sequence with standard deviations, calculated with the variance analysis method (see section on Variance Analysis), using 1430 event records in 30 wells, based on the LO and LCO occurrences of 289 benthic and some planktonic foraminifera and dinoflagellates, most or them analysed 'in-house', which greatly enhances taxonomic consistency. In addition, North Sea log markers were incorporated (NS Log B–G) for which precise well depths data are available (see Gradstein *et al.*, 1994 for details). Each of the 88 events in the zonation occurs in at least 7 wells, except for 16 unique events (marked with two asterisks in Figs. 16.8 and 16.9) that occur in fewer than seven wells, and are inserted to complement the zonation, and/or assist with age calibration. Forty-two of the events in the optimum sequence, including many DWAF taxa, have standard deviation below average, which is quite a good number for an industrial type dataset.

The RASC scaled optimum sequence with zones assigned is shown in Figure 16.9. There are 18 zones and subzones assigned, named NSR1–13 (NSR = North Sea

RASC), of early Paleocene through early Pleistocene age. Large breaks (at events 129, 50, 206, 6, 266 and 23) indicated transitions between natural microfossil sequences, and/or hiatuses, and are candidates for sequence stratigraphic breaks if corroborated by regional seismic analysis. The zones contain 33 DWAF events (32 LO and 1 LCO events) for 32 taxa. On average, event observation in the wells may be closer to the average stratigraphic position than the last occurrence end points in regional range charts.

The average range end of DWAF in the North Sea (coarse agglutinated spp. in Fig. 16.9) is generally in mid-Cenozoic, and falls in zone NSR8A, late Oligocene. It occurs in 27 out of 30 wells in the dataset, and has an sd of 2.402, which is above average (Fig. 16.8). With variance analysis technique in RASC it is possible to display the well deviations (Fig. 16.10). Since the wells in this figure are arranged from north (left) to south (right), the method confirms that the DWAF LO is time-transgressive, with younger occurrences southward, as can be readily observed in southern (Central Graben) wells. Since in subregions of the dataset the DWAF LO is a reasonable good marker, a histogram of stratigraphic deviations (Fig. 16.10a), although rather wide still shows a reasonable central (normal Gaussian) distribution.

A more marked central tendency, and a much lower sd of 1.287, despite a sample size of only 14 and not 27 as for DWAF, is calculated for the distinctive planktonic foraminifer *Subbotina patagonica* of Zone NSR4, Lower Eocene (Fig. 16.10).

Given a large enough dataset (10 or more wells is a reasonable rule), the variance analysis technique allows to rapidly rank events in the RASC solutions on stratigraphic fidelity.

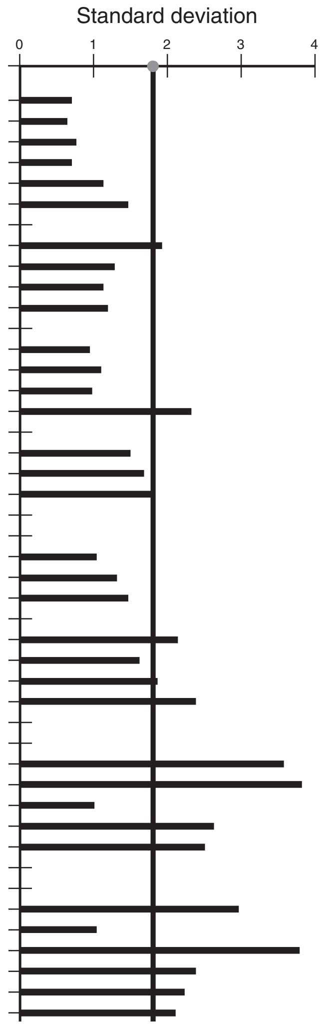
16.6 Computer Programs

A PC desktop computer version under DOS of the original method of graphic correlation is program GRAPHCOR (Hood, 1995; K.C. Hood, 9707 Arrowgrass Dr., Houston, TX 77064, USA). It was converted by Amoco Oil Co for interactive operation on a Unix workstation. An example of adaptation of graphic correlation to probabilistic stratigraphy is found in program STRATCOR (Gradstein, 1990), which has hybrid features to RASC.

Desktop PC program CONOP under DOS, has many features to analyse medium-size stratigraphic datasets. It has colour graphics displays, and the progress of search for the optimal range chart may be watched on screen, which is an instructive option. The program is actively being developed by Peter M. Sadler (Department of Earth Sciences, University of California, Riverside, CA 92521, USA).

The probabilistic stratigraphy programs that perform ranking, scaling, correlation and standard error calculation operate as a single module under MS Windows, and are called RASC & CASC. One windowing master menu controls the operation of the programs and their results that also include the data input and re-organising module called MAKEDAT, and correspondence analysis program COR (Hill, 1979; Bonham-Carter

N¼ & UI				N	SD
1		77	*Elphidium* spp.	20	0.675
2		228	*Cassidulina teretis*	20	0.559
3		31	*Cibicidoides scaldisensis*	8	0.724
4		270	*Cibicidoides grossa*	17	0.656
5		1	*Neogloboquadrina pachyderma*	12	1.068
6		171	*Buccella frigida*	7	1.421
7		318 **	*Monspeliensis pseudotepida*		
8		316	*Trifarina fluens*	7	1.909
9	0 1	23	*Sigmoilopsis schlumbergeri*	18	1.219
10	–10	4	*Globorotalia inflata*	9	1.123
11	0 1	269	*Neogloboquadrina atlantica*	14	1.162
12		74 **	*Neogloboquadrina humerosa*		
13	–10	266	*Globorotalia puncticulata*	8	0.914
14	0 2	130	*Siphonina advena*	9	1.083
15	–10	6	*Neogloboquadrina acostaensis*	7	0.929
16	0 1	219	*Martinotiella cylindrica*	11	2.347
17		301 **	*Bolboforma metzmacheri*		
18	–11	91	Diatoms/radiolarians LCO	12	1.475
19	–11	207	NS Log G	15	1.653
20	–10	282	*Uvigerina* ex.gr. *semiornata*	7	1.777
21		58 **	*Neogloboquadrina atlantica* (dext)		
22		411 **	*Globocassidulina subglobosa* var.		
23		123	*Globigerinoides trilobus*	8	1.022
24		236	*G.* ex.gr. *praescitula zealandica*	16	1.281
25		285	*Caucasina elongata*	10	1.478
26		362 **	*Cannosphaeropsis utinensis*		
27		125	*Neogloboquadrina continuosa*	8	2.150
28		17	*Asterigerina gurichi*	21	1.625
29		18	*Spiroplectammina carinata*	7	1.867
30		15	*Globigerina praebulloides*	18	2.369
31		528 **	*Catapsydrax parvulus*		
32		348 **	*Sphaeroidinella disjuncta*		
33		138	*Tenuitella angustiumbilicata*	7	3.568
34		111	*Globigerina ex.gr. officinalis*	7	3.855
35		413	*Silicious biofacies*	8	1.026
36		20	*Gyroidina girardana*	12	2.633
37		25	*Coarse agglutinated* spp.	28	2.512
38		369 **	*Rhombodinium draco*		
39		368 **	*Chiropteridium mespilanum*		
40		97	*Cyclammina placenta*	24	2.968
41		182	*Spirosigmoilinella compressa*	22	1.052
42		262	*Karrerulina horrida*	12	3.780
43		259	*Ammodiscus latus*	9	2.407
44		183	*Reticulophragmium rotundidorsatum*	15	2.240
45		24	*Turrilina alsatica*	11	2.126
46		140	*Rotaliatina bulimoides*	14	1.572



47		261		*Haplophragmoides walteri*	22	2.722
48		321		*Dorothia seigliei*	13	1.713
49		289		*Adercotryma agterbergi*	13	1.255
50		148	**	*Globigerapsis index*		
51		12		*Areosphaeridium diktyoplokus*	8	1.864
52		535		*Annectina biedai*	8	3.061
53		229		*Recurvoides* ex.gr. *walteri*	8	2.960
54		206		NS Log F	14	2.066
55		29		*Reticulophragmium amplectens*	28	1.511
56		32		*Ammosph.pseudopauciloculata*	17	2.085
57		263		*Ammomarginulina aubertae*	14	2.073
58		260		*Haplophragmoides kirki*	14	2.113
59		68		*Spiroplectammina spectabilis* LO	24	1.932
60		37	**	*Acarinina pentacamerata*		
61		264		*Karrerulina conversa*	19	3.433
62		205		NS Log E	15	1.337
63		14	**	*Eatonicysta ursulae* LCO		
64		47	**	*Planorotalites planoconicus*		
65		54		*Spiroplectammina navarroana*	19	1.721
66		279		*Recurvoidella lamella*	9	1.464
67		50		*Subbotina patagonica*	14	1.287
68	0 1	277		*Tuff*	7	1.089
69	–10	22		*Fenestrella antiqua*	12	1.323
70		78		*Trochamminoides coronatus*	10	3.165
71		204		NS Log D	15	0.983
72		203		NS Log C	15	1.111
73		105		*Rzehakina minima*	8	0.894
74		343	**	*Cenosphaera* sp. (disc or flat)		
75		310		*Saccammina placenta*	11	2.162
76		76		*Reticulophragmium pauperum*	19	0.897
77	0 2	283		*Reticulophragmium garcilassoi*	7	2.083
78	–10	57		*Spiroplectammina spectabilis* LCO	20	0.972
79		129		*Ammoanita ruthvenmurrayi*	11	1.839
80		134		*Caudammina excelsa*	11	2.512
81		532		*Cystammina sveni*	12	2.860
82		531	**	*Ammoanita ingerlisae*		
83		55		*Gavelinella beccariiformis*	15	1.245
84		107		*Remesella varians*	8	1.670
85		202		NS Log B	12	1.143
86	0 1	61		*Subbotina pseudobulloides*	18	1.046
87	–10	60		*Planorotalites compressus*	12	0.554
88		253		*Subbotina triloculinoides*	10	0.593

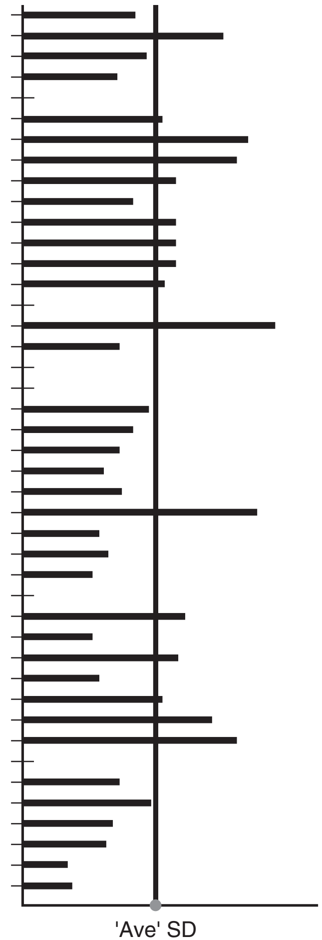
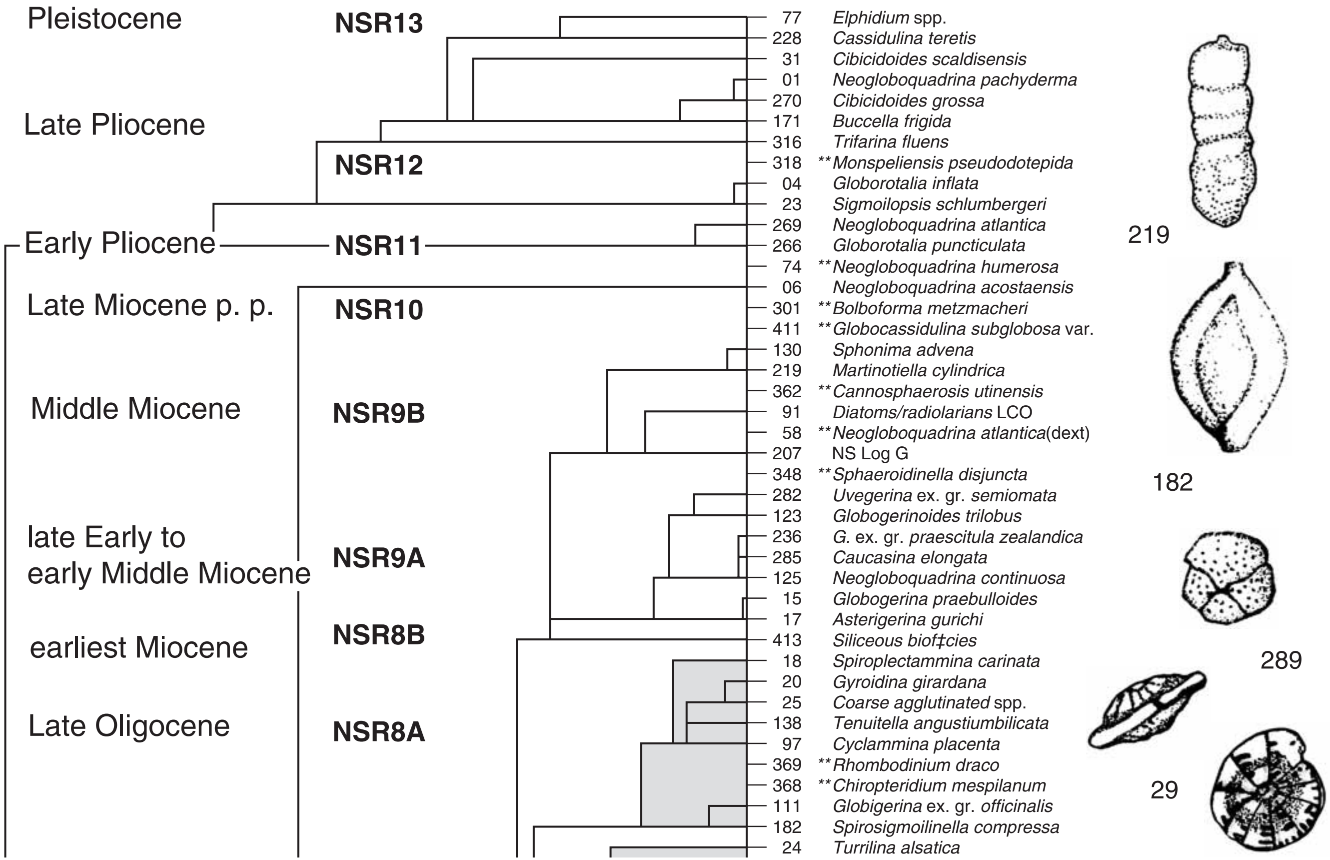


Figure 16.8 RASC Optimum Sequence for the Cenozoic of the North Sea; low standard deviations are an indication of good stratigraphic markers; ave SD = average event standard deviation; N = event occurrence in wells; UI = uncertainty interval on event position (in this case +/− 1 or 2 positions); each event has a dictionary number in front.


Pleistocene
NSR13
Late Pliocene
NSR12
Early Pliocene
NSR11
Late Miocene p. p.
NSR10
Middle Miocene
NSR9B
late Early to
early Middle Miocene
NSR9A
earliest Miocene
NSR8B
Late Oligocene
NSR8A
77 Elphidium spp.
228 Cassidulina teretis
31 Cibicidoides scaldisensis
01 Neogloboquadrina pachyderma
270 Cibicidoides grossa
171 Buccella frigida
316 Trifarina fluens
318 **Monspeliensis pseudodotepida
04 Globorotalia inflata
23 Sigmoilopsis schlumbergeri
269 Neogloboquadrina atlantica
266 Globorotalia puncticulata
74 **Neogloboquadrina humerosa
06 Neogloboquadrina acostaensis
301 **Bolboforma metzmacheri
411 **Globocassidulina subglobosa var.
130 Sphonima advena
219 Martinotiella cylindrica
362 **Cannosphaerosis utinensis
91 Diatoms/radiolarians LCO
58 **Neogloboquadrina atlantica(dext)
207 NS Log G
348 **Sphaeroidinella disjuncta
282 Uvegerina ex. gr. semiomata
123 Globogerinoides trilobus
236 G. ex. gr. praescitula zealandica
285 Caucasina elongata
125 Neogloboquadrina continuosa
15 Globogerina praebulloides
17 Asterigerina gurichi
413 Siliceous biof‡cies
18 Spiroplectammina carinata
20 Gyroidina girardana
25 Coarse agglutinated spp.
138 Tenuitella angustiumbilicata
97 Cyclammina placenta
369 **Rhombodinium draco
368 **Chiropteridium mespilanum
111 Globigerina ex. gr. officinalis
182 Spirosigmoilinella compressa
24 Turrilina alsatica
219
182
289
29

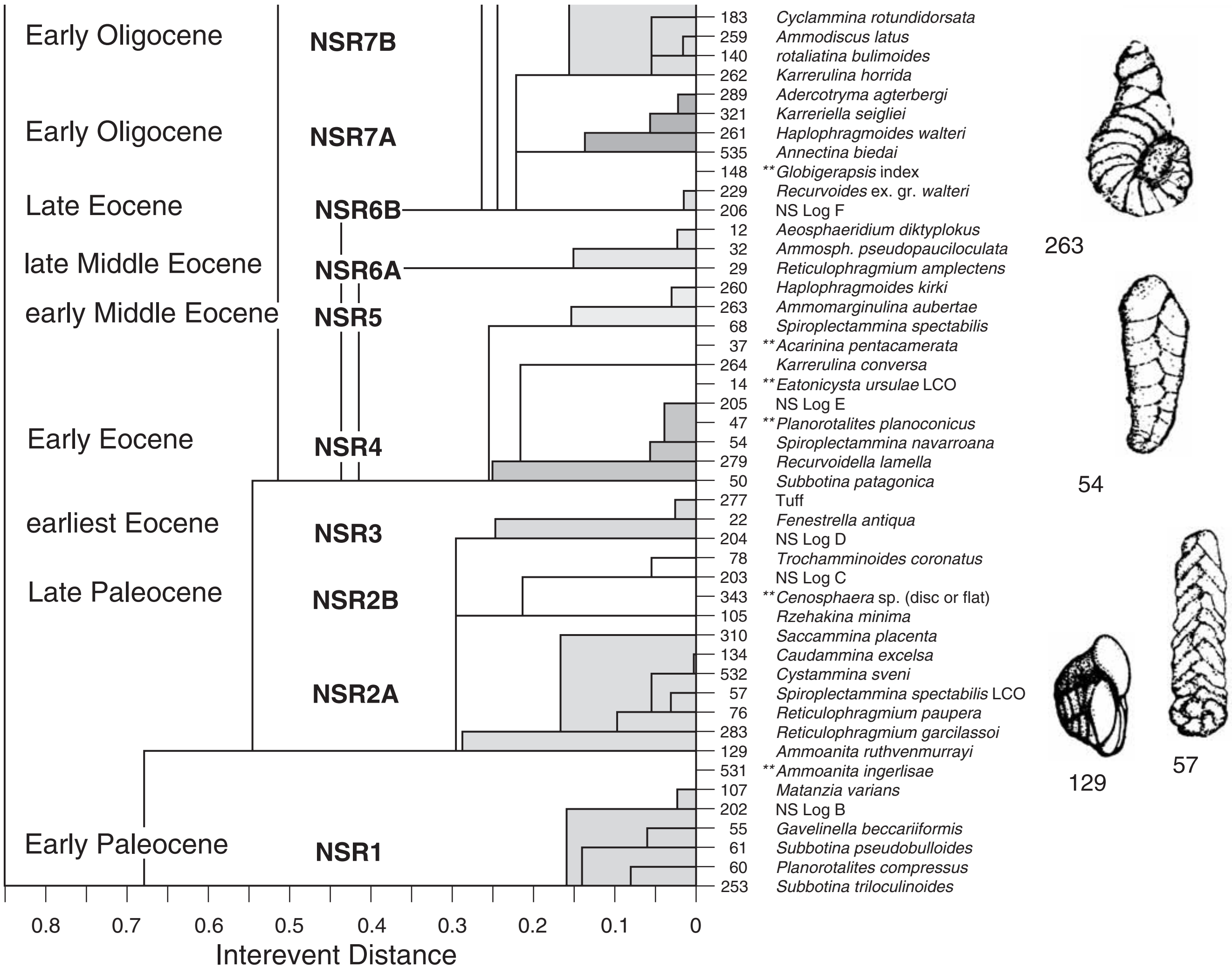
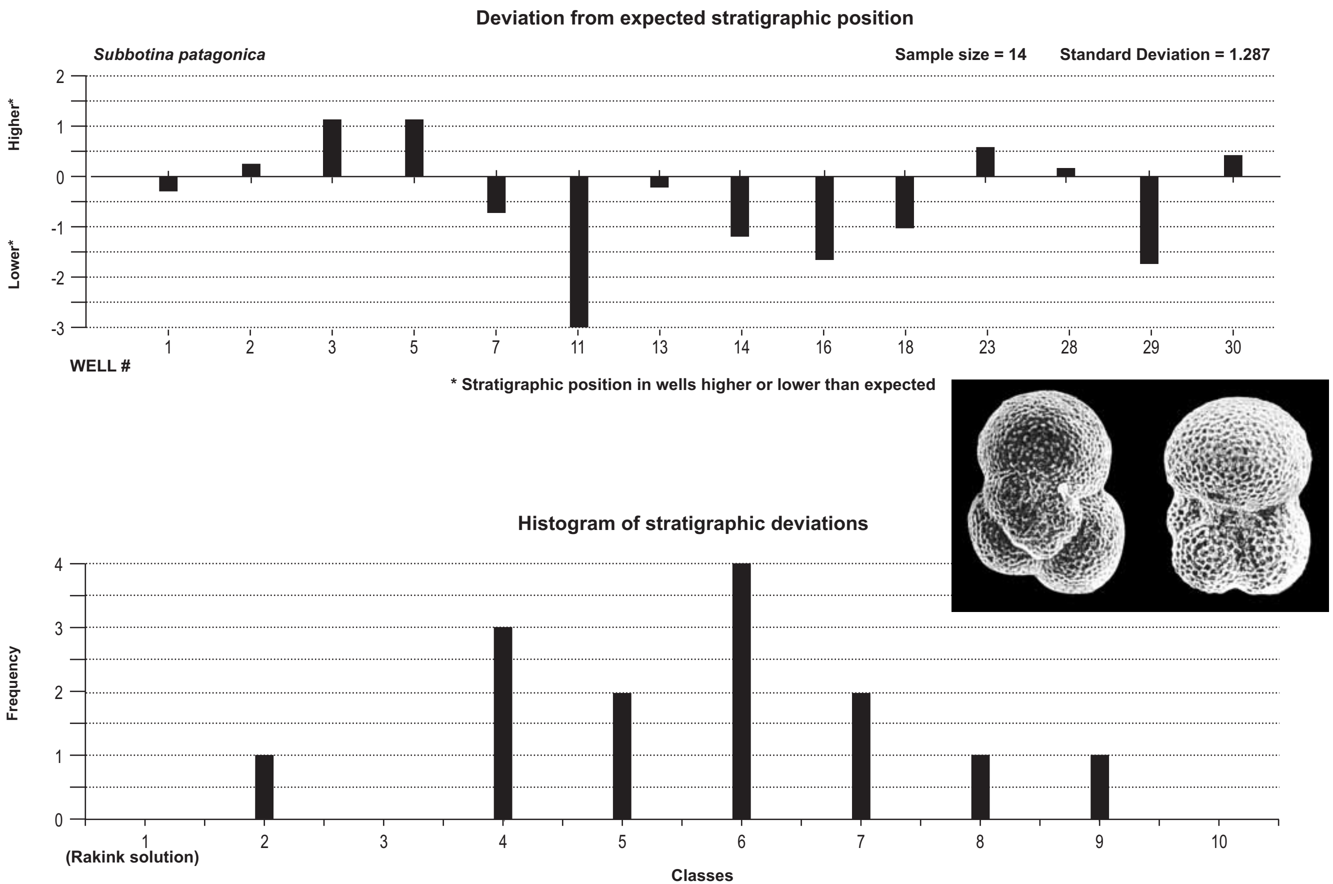


Figure 16.9 Scaled version of the RASC Optimum Sequence of Figure 16.8, with eighteen interval zones assigned (NSR1 – 13) of early Paleocene through Pleistocene age.


Deviation from expected stratigraphic position
Subbotina patagonica
Sample size = 14
Standard Deviation = 1.287
Higher*
Lower*
2
1
0
-1
-2
-3
1
2
3
5
7
11
13
14
16
18
23
28
29
30
WELL #
* Stratigraphic position in wells higher or lower than expected
Histogram of stratigraphic deviations
Frequency
4
3
2
1
0
1
(Rakink solution)
2
3
4
5
6
7
8
9
10
Classes

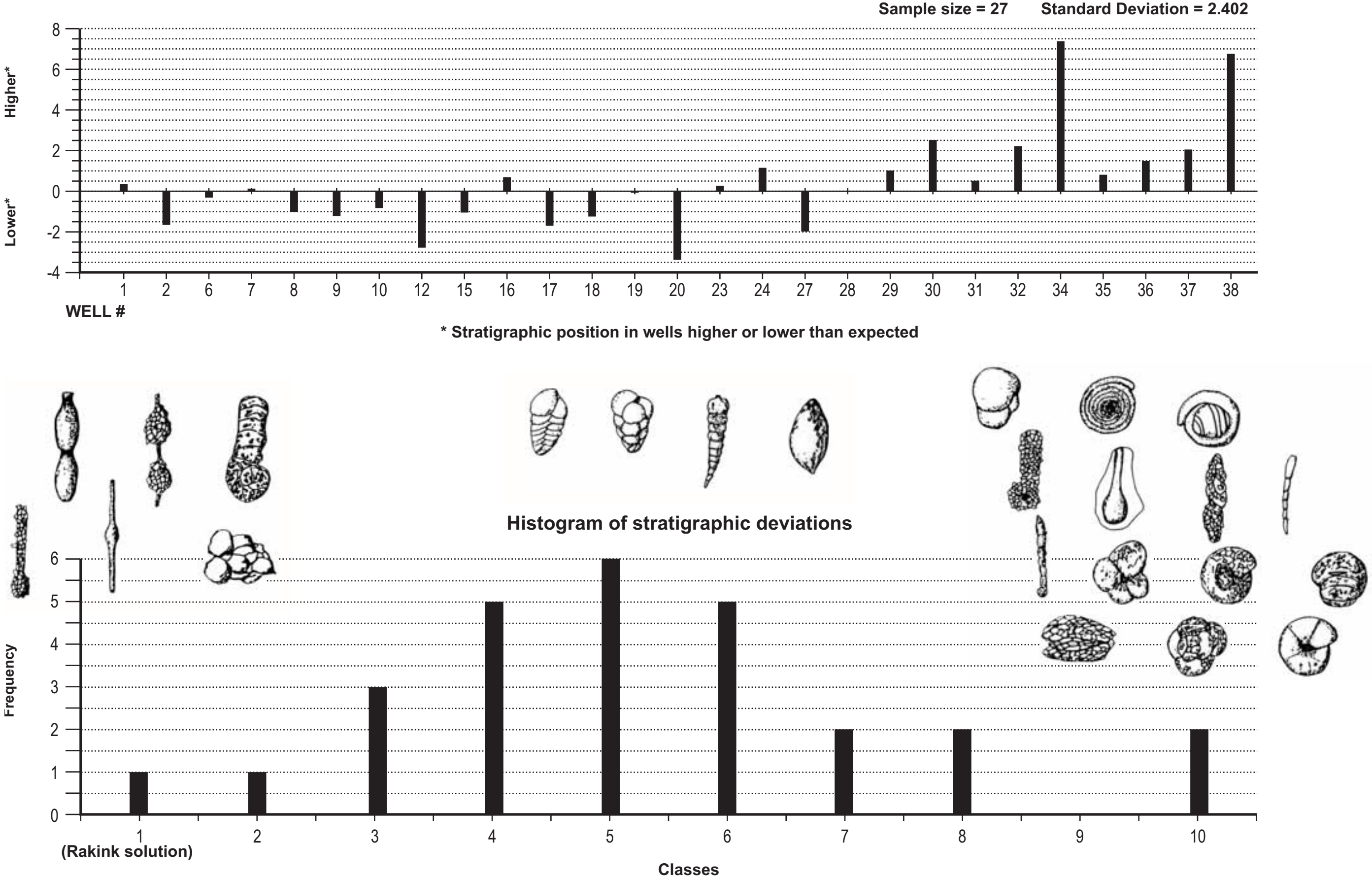


Figure 16.10 Deviations from expected stratigraphic position and histograms of stratigraphic deviations for the average LO of Deep Water Agglutinated Foraminifera, and of the planktonic foraminifer Subbotina patagonica*, North Sea.*

et al., 1986). Graphics results are displayed in colour, and may be modified and edited with a build-in 2D chart control program, and colour printed or plotted from the screen displays. The program is actively being developed by F.P. Agterberg and F.M. Gradstein. More details may be found on websites *www.q-strat.org, www.geocities. com/rasc_casc,* and *www.stratigraphy.org.*

16.7 Acknowledgements

Thanks are due to Frits Agterberg, Peter Thompson and an anonymous reviewers for constructive comments that improved this manuscript, and to H.V. dos Santos, Petrobras-Cenpes, for excellent assistance with drafting of several figures. Roger Cooper and Peter Sadler provided valuable scientific advice and permission to utilize two figures.