Palaeogeography, Palaeoclimatology, Palaeoecology 285 (2010) 237–247

Contents lists available at ScienceDirect

Palaeogeography, Palaeoclimatology, Palaeoecology

journal homepage: www.elsevier.com/locate/palaeo

Neogene biochronology of Antarctic diatoms: A comparison between two quantitative approaches, CONOP and UAgraph

Federico Galster [a], Jean Guex [a,*], Oyvind Hammer [b]

[a] Department of Geology and Palaeontology, University of Lausanne, Anthropole, CH-1015 Lausanne, Switzerland
[b] Natural History Museum, University of Oslo, Norway

A R T I C L E I N F O

Article history:
Received 26 June 2009
Received in revised form 11 November 2009
Accepted 13 November 2009
Available online 20 November 2009

Keywords:
Quantitative stratigraphy
Biochronological correlations
Diatoms
CONOP
Unitary Associations

A B S T R A C T

A quantitative biochronological study by Cody et al. (2008) integrates comprehensive diatom biostratigraphy, magnetostratigraphy, and tephrostratigraphy from 32 Neogene sections around the Southern Ocean and Antarctic continental margin. A recent method, known as Constrained Optimization (CONOP), which can be viewed as a multidimensional version of graphic correlation, is applied to that very interesting database. The goal of the present paper is to discuss some theoretical aspects of quantitative biochronology and to compare the constrained optimization with the deterministic method called Unitary Associations (UAM), a graph theoretical model. We illustrate the fact that the UAM is an extremely powerful and unique theory allowing an in-depth analysis of the internal conflicting inter-taxon stratigraphic relationships, inherent to any complex biostratigraphical database.

1. Introduction

Cody et al. (2008) recently produced a convincing demonstration that quantitative stratigraphic approaches can produce results with a much higher resolution potential than empirical zonations. This approach was based on the method known as “Constrained Optimization”, using the program CONOP (Sadler, 2006), and provided a splendid example to be compared with other methods.

The goal of the present paper is to compare the constrained optimization with an alternative quantitative tool, the Unitary Associations, a graph theoretical model, and explain the differences between the outputs. It is a common belief among non specialists that all quantitative tools in stratigraphy are more or less equivalent but we do not agree with that idea. The present paper should help the users of quantitative biostratigraphical tools in their choice of one or the other method.

A short description of the two approaches is given below.

2. CONOP and UAs: a brief description

2.1. The UA method

The Unitary Associations (UAs) method is designed for the construction of concurrent range zones using a fully deterministic approach. The basic idea is to construct a discrete sequence of coexistence intervals of species. Each interval, corresponding to one UA, is of minimal duration while consisting of a maximal set of intersecting ranges. Each UA is characterised by a set of species allowing its identification in the stratigraphic sections.

The basic steps of the method are as follows. The data are compiled into a presence–absence matrix, with samples in rows and taxa in columns. From these data, maximal sets of mutually co-occurring species (maximal cliques) are constructed. Stratigraphical superpositions of maximal cliques are then inferred from the observed superpositional relationships between the taxa they contain. The longest possible sequence of superposed maximal cliques is then used to construct a sequence of UAs. Finally, the original samples are assigned to UAs whenever possible and are thus stratigraphically correlated.

The difficult part of the UA theoretical model consists in finding and resolving what we call conflicting stratigraphic relationships, i.e. cyclic structures in the graphs representing the inter-taxa stratigraphic relationships (G*, see Fig. 1B) on one side and the strongly connected components in the graph representing the sequential relationships between the maximal cliques (Gk) on the other side. A full review of the Unitary Associations method is beyond the scope of the present paper and can be found in Guex (1991).

The program used in the present paper is the latest version of UAgraph (Hammer et al., 2009, available at http://folk.uio.no/ohammer/uagraph) which is an improved version of the Biograph program (Savary and Guex, 1991, 1999).

2.2. The Constrained Optimization method

The Constrained Optimization method used by Cody et al. (2008) was originally designed to produce automated graphic correlation and

* Corresponding author.
E-mail address: Jean.Guex@unil.ch (J. Guex).

0031-0182/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.palaeo.2009.11.015

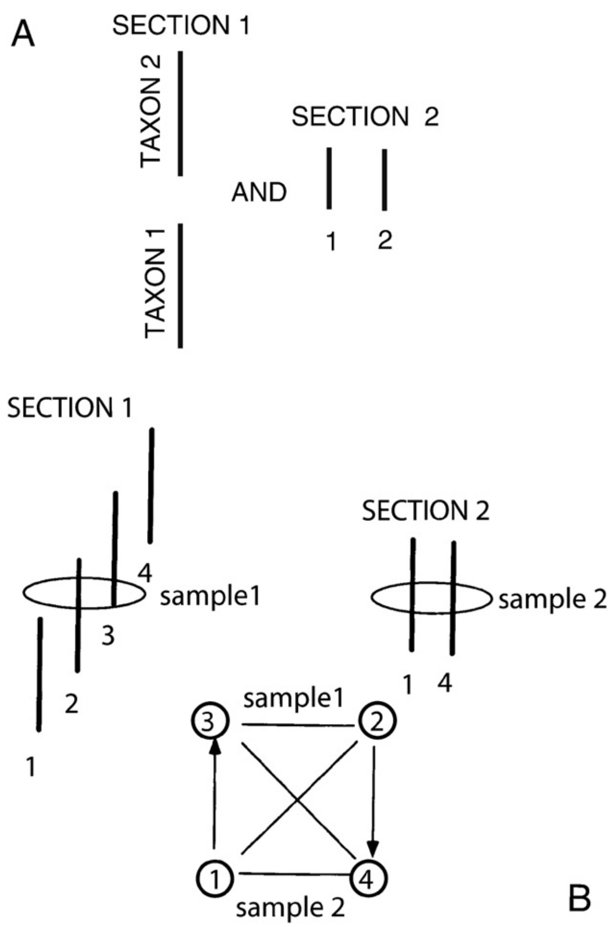

Fig. 1. A) A typical conflicting situation considered and treated by CONOP (after Sadler and Cooper, 2003). Such cases are not considered as conflicting in the UA method. B) Conflicting relationship between two samples and four taxa involved in a cycle of length 4 (Z4 in graph theoretical terminology). In the graph, straight lines (edges) between taxa represent co-occurrences, while arrows (arcs) represent superpositions. Note that sample 1 is simultaneously above and below sample 2.

sequences of events in a multidimensional space (Kemple et al., 1995) by means of the program named CONOP. The first rule used by the authors is to satisfy the observed inter-species coexistences, as in the UA method. One particular type of contradictory situation treated by CONOP is illustrated in Fig. 1A.

The second criterion is that the computed solution must satisfy the condition that adjusted FO and LO horizons lie at or below and at or above the observed FO and LO respectively. The latest version of the CONOP program (Sadler, 2006) offers multiple choices and adjustments for a variety of different parameters, such as specifying the measure of misfit and comparing the penalty of several different sequences.

CONOP considers the problem of stratigraphic correlation as NP-complete (following Dell et al., 1992; see remark below). Sadler (2006) notes that the CONOP program minimizes the simplifying assumptions and maximizes the flexibility of the choice of measures of fit between solutions and the data. This is achieved by inverting the solution process. Instead of building a solution from the data (as in the UA method), 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, a process called “inversion” by geophysicists. Cody et al. (loc. cit.p.107) use an option called LEVEL as a penalty function to minimize the corrections of local ranges, i.e. the minimal value upon which the simulated annealing is calculated (details in Sadler, 2006, see also Kirkpatrick et al., 1983 for details on simulated annealing). The present discussion is limited to this particular choice.

When judiciously configured by an experienced user, the program may be tailored to contract or extend local ranges or limit the use of the coexistence constraint by eliminating some rare co-occurrences (i.e. observed in only one sample) during the computation. Again, a full description of the model is outside the scope of the present paper and the reader is referred to Kemple et al. (1995) and Sadler (2006) for details.

In summary, the main differences between UAs and CONOP is that the goal of the first method is to construct concurrent range zones whereas the second method constructs sequences of datums. Note that UA constructions are based on sample contents and their mutual stratigraphic relationships (i.e. the biostratigraphic graph G*). Sequences of datums established by means of UAs consist of non-diachronous datums only which are constructed in a second run of the program, after a semi-empirical selection. During the correlation process, such datums are treated as intervals of uncertainty, like the total ranges of the taxa calculated by the program.

CONOP on its side is exclusively oriented towards a seriation of the events based on inter-stratigraphic sections comparisons (i.e. graphic correlation philosophy, Shaw, 1964).

We can also notice that the two methods will reach similar conclusions when applied to very complete sets of empirical information relative to inter-taxa coexistences.

3. Technical remarks about the NP-completeness of the problem

In our theoretical model, the stratigraphic correlation problem is not addressed as an NP-complete problem but as a problem of seriation of the observed samples. Sequences of first and last occurrences (FOs and LOs) are only treated when these events are not diachronous in the different localities. This explains the important differences existing between the computing time used by UAgraph and that used by CONOP. In the UA program, the number of computing operations is linearly proportional to the number of samples and to that of the taxa involved. In CONOP the growth is exponential but can be limited at the expense of precision.

3.1. The database

One major interest of the Cody et al. (2008) paper is that the complex problem which was originally addressed, the quantitative study of the biostratigraphy, magnetostratigraphy and tephrostratigraphy of 116 diatom species from 32 Neogene sections around the Southern Ocean and Antarctic continental margin has been analysed by the best specialists on the CONOP program, producing a result which can be interpreted as one of the best possible results obtainable by means of Constrained Optimization.

Another significant aspect of Cody's database is that the original taxonomy was carefully revised and homogenised before being applied to CONOP and that the doubtful local range extremities (reworking and downworking, contaminations) were consciously eliminated.

3.2. UA processing of the data

All the data considered here are strictly taken from Cody's database (Online Appendix 2 in Cody et al., 2008), which is considered, for theoretical reasons, as exact.

The goal of our application of the new UA program UAgraph (Hammer, Guex and Savary in PAST) is to produce an *integrated* solution given the complex input including the inter-species coexistences, selected moderately diachronous first and last occurrences and paleomagnetic data. The database used for the computation is given in the online Appendix 1 (1-A: local ranges; 1-B: taxa dictionary; 1-C: sections codes).

The paleomagnetic data (base/respectively top of the recognized reversals) have been included into the computed database to force an