Paired Comparison Models in Biostratigraphy

Carolyn Brauer Hudson and Frederik P. Agterberg

Paired comparison models can be used to study the order of stratigraphic events along a relative time scale when differences in order are observed to occur in stratigraphic sections or wells. They give parameters of (1) a binomial distribution for the probability that a given event occurred earlier than another event, or (2) a trinomial distribution for the probability of earlier, later, or simultaneous occurrences of events. In the Bradley-Terry and Davidson models differences in positions of events along a logarithmic scale are assumed to satisfy a logistic frequency distribution. In the recently developed RASC model, a normal frequency distribution along a linear scale is used. The properties of these three models are reviewed and computer algorithms used to obtain practical solutions are discussed. The three models have each been applied to three published data sets (10 highest and lowest occurrences of Eocene nannofossils in 9 wells in California; 16 lowest occurrences of Cambrian trilobites and other fossils in 7 stratigraphic sections in Texas; and 41 highest occurrences of Cenozoic foraminifera and other microfossils in 16 wells, Canadian Atlantic Margin). Results of the three models are not significantly different. However, only the RASC model can be applied to large data sets because its computer algorithm avoids the use of time-consuming iterative processes required when other models are used.

KEY WORDS: Paired comparison models, ranking techniques, biostratigraphy, paleontology.

INTRODUCTION

Statistical methods have been used in the analysis of biostratigraphical events. Overviews include Shaw (1964), Agterberg and Gradstein (1981), and Brower (1981). One of the goals of these studies has been to establish a geochronological ordering of biostratigraphic events, such as highest and lowest occurrences of fossil taxa.

The statistical method of paired comparisons has frequently been applied to data ranking. Reviews of paired comparison techniques can be found in Torgerson (1958), Brunk (1960), David (1963), and Kendall (1975). Not all of these

2 Department of Geology, University of South Carolina, Columbia, S.C. 29208.
3 Geological Survey of Canada, 601 Booth Street, Ottawa, Canada, K1A 0E8.
procedures are valid for biostratigraphical data because the number of paired comparisons cannot be controlled. For example, Worsley and Jorgens (1977) have shown that an optimum sequence is difficult to determine when many fossils are missing in observed sequences. Methods applied to biostratigraphical data include those developed by Hay (1972), Edwards and Beaver (1978), and Gradstein and Agterberg (in press).

In this paper several applicable models are discussed. The paper concludes with a comparative analysis of three of these models: (1) the Bradley-Terry model in which the probability that an event (highest or lowest occurrence of a fossil taxon) occurs above another event is estimated for all events (Bradley and Terry, 1952); (2) the Davidson model which is an expansion of the Bradley-Terry model in that the probability that two events are simultaneous is also estimated (Davidson, 1970); and (3) the recently developed RASC model (for ranking and scaling) in which the events are ordered along a linear scale with estimation of "distances" between them (Gradstein and Agterberg, in press). The purpose of defining distances between events can be illustrated as follows. Two events which are coeval or nearly coeval on the average are likely to be assigned approximately zero distance between them on this linear scale. On the other hand, an event that nearly always occurs above another event will be separated by a longer distance from this other event. It will be shown that each distance can be transformed into the probability of one event occurring above the other in a stratigraphic section.

The three models yielded similar results for three published data sets used for the comparative analysis; however, the RASC model was found to be the most expedient in terms of computer time when large sets of data were analyzed.

**OCCURRENCES OF BIOSTRATIGRAPHIC EVENTS**

In a stratigraphic section, the first and last occurrences of a fossil taxon correspond in a general way to the time during which the taxon existed. The range may appear to be longer than average because of sediment mixing that moved fossils upward or downward. Sediment mixing can also shorten the apparent range. Other factors which can cause this are poor preservation, slow dispersion, early extinction in localities studied, facies control, incomplete sampling, and misidentification.

For any given series of biostratigraphic events, a true order of first and last occurrences is assumed. Each event \( E_i \) (where \( i = 1, 2, \ldots, N; N = \) the number of events) is associated with a value \( \pi_i \) on a scale which may be a relative time scale. If these events are observed in several sections, statistical inferences about their true order are possible by considering the order relationship between each pair of events.

If two events \( E_i \) and \( E_j \) occur in the same section, three order relations in