

Restoring gradients from fossil communities: a graph theory approach

Petr ČEJCHAN

Geological Institute, Academy of Sciences of the Czech Republic, Rozvojová 135, CZ – 165 02 Praha 6 – Lysolaje, Czechia

ABSTRACT. Gradients are often involved in patterns of species distribution in space and time. They are characterised by a specific pattern of species abundance, which in turn can be used to reconstruct fossil gradients from fossil biotic communities. Rather than the classical approach to gradient retrieval, based essentially on clustering or eigenvector methods, a graph theory was used as closer to the nature of the problem. Late Devonian marine benthic communities were successfully used to search for the latent gradient corresponding to the intensity of environmental stress in the vicinity of the Kellwasser extinction event. The reconstructed gradient was then in turn applied to elucidate the history of environmental stress during the Kellwasser crisis interval.

KEYWORDS: quantitative palaeoecology, gradients, methods, graph theory, travelling salesman problem.

Prerequisites

Continuous, directed environmental change, a time trend, or other similar processes often produce a typical pattern when reflected in fossil record. Such a directed smooth change is a **gradient**. What is its typical pattern? Usually, different species meet their optimal conditions at different places along the gradient; these optima are characterised by the highest abundance of the corresponding species. In both directions away from the species optimum, its abundance usually decreases monotonically (Fig. 1). How the species optima are spaced along the gradient depends on the process itself, as well as on other controls; however, the positions of optima usually do not coincide, but tend to be spaced in a spectrum of possibilities from almost random to almost regular. Examples include a depth gradient in a marine environment: some organisms are adapted to shallower conditions, others to deep water ones, each having its respective optimum on a defined point on the gradient (maximum abundance), and its tolerance limits (towards them the abundance is decreasing to zero). However, this is a very simple example.

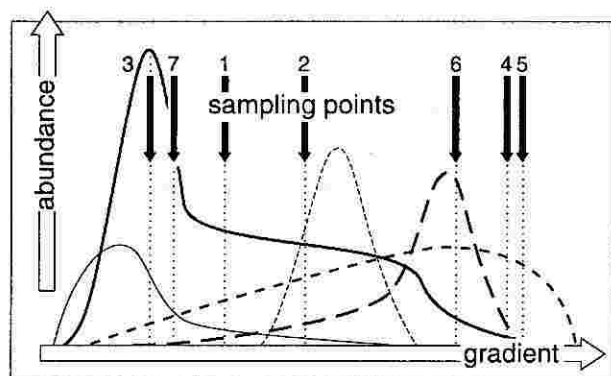


Figure 1. Typical distribution of species abundance along the gradient. Abundance curves are usually continuous, smooth, unimodal. The gradient is sampled more or less randomly at sampling points (1 to 7) of unknown position on the gradient. The sequence of samples 3-7-1-2-6-4-5 is the natural ordering, corresponding to the order of sampling points along the gradient. The aim of the study is to find the natural ordering using species abundances only.

Aims

In a fossil record, such gradients are commonly latent and are recorded only indirectly via biotic communities. This study deals with such a situation. The record may be (and commonly is) incomplete, or biased, but that is another story. Assume for now that the fossil communities preserve a usable record of the original gradient, and that the only thing we possess are samples of randomly chosen communities from different places on the assumed gradient. The crucial point is that we do not know anything about the nature of the gradient itself (i.e., the controlling process is compound or unknown), nor do we have access to its direct record. Now we would like to arrange our samples in a series that corresponds to a natural section along that gradient (Fig. 1).

Classical approach

Classically, this task was addressed via different varieties of cluster analysis, or via a primitive eigenvector procedure, the principal components analysis. Both branches evolved into more advanced techniques (e.g. fuzzy clustering, supervised clustering, etc., on the one side, and factor analysis and multidimensional scaling on the other). Yet all of these methods, although advanced, suffer from a common shortcoming: they make unrealistic assumptions about the data. While clustering methods essentially assume that distinguishable natural grouping exists in the data and believe that forcing the data into clusters will uncover the latent information, the eigenvector-based methods assume the linear response of the data to the latent factors. Both assumptions are unrealistic for our case. We chose another approach: from "realistic" model to suitable algorithm.

The Travelling Salesman: our approach

We started from the assumptions about the biotic responses to the continuous change along the gradient, stated above. According to the model, the gradient is crowded by continuous, smooth, unimodal "abundance curves" of all the species involved. The curves have their maxima at different points along the gradient. The sampling points are spaced more or less randomly on the gradient; quantitative samples of communities originate here yielding species abundances.

- (3) real abundance values are well represented by sampled data,
- (4) squared Euclidean distance is a suitable measure for assessing dissimilarities of biotic communities,
- (5) the gradient is sufficiently densely sampled, and
- (6) sum of inter-sample distances along the gradient is, or approaches, the minimum possible.

Assumption (1) is usually accepted. Violations are possible, e.g., when strong competition occurs and the stenoeccous stronger competitor outcompetes the euryeocous weaker competitor species from a part of its span over

the gradient. We do not know how serious assumption (2) is, i.e., how much does the clustering of maxima affect assumption (6). Modelling should give answer to this, but it is a matter for future research. Assumption (3) can be tested by resampling or using a binomial distribution. Neither was done during this study. We have no answer to (4), perhaps modelling will say more. We have no answer to (5) as well, no idea how to address this. Intuitively, we feel that (6) holds, but we did not prove it. Perhaps proof is possible, or we should use modelling to give us some insight.

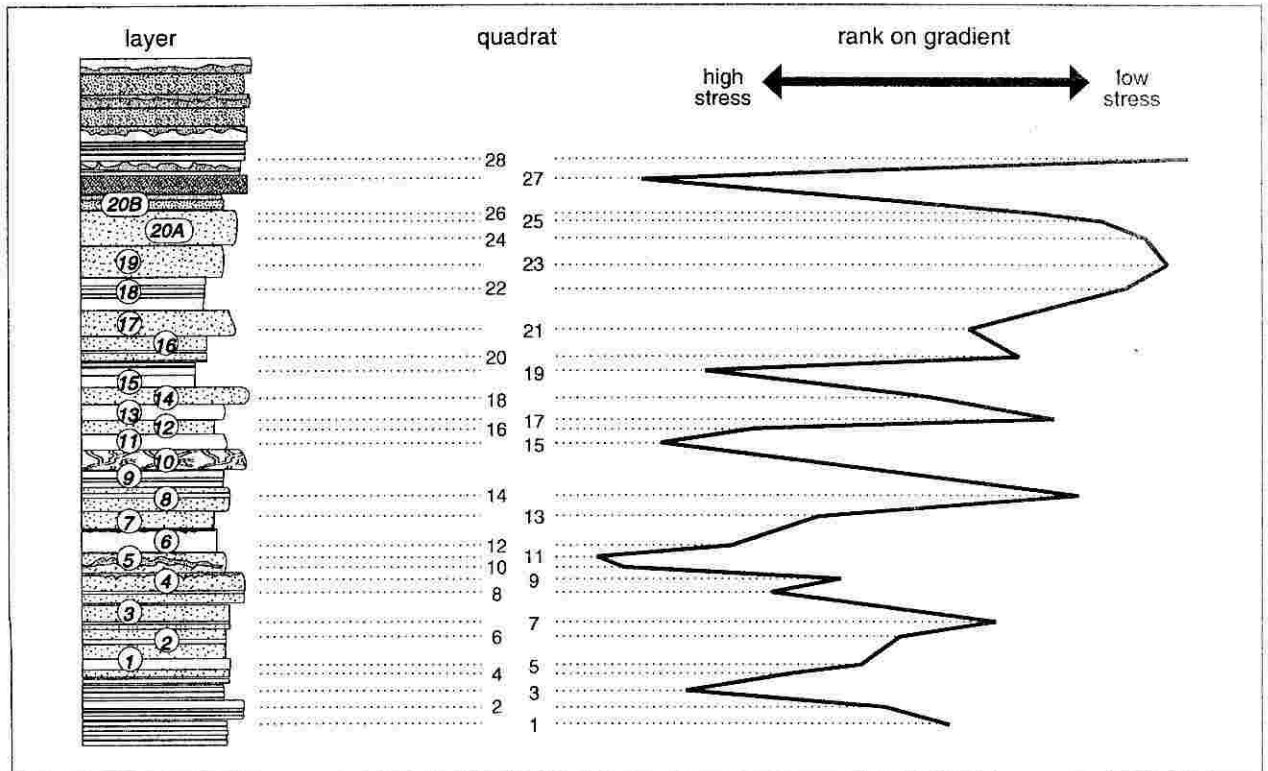


Figure 3. Mokra, Western Quarry section. Late Devonian marine limestone sequence, vicinity of the Frasnian - Famennian boundary. Ranks of quadrats on presumed extinction gradient, derived from the TSP-reconstructed order of quadrats. Squared Euclidean distance based on abundances was used as a distance measure. Abundances were measured as numbers of individuals, or colonies. The curve of ranks is supposed to show the intensity of environmental crisis. Numbers in circles (italics) are numbered layers of the section, numbers outside the column are quadrat numbers.

Conclusion

The TSP algorithm showed a good capability to recover a meaningful gradient involved in the species abundance data collected from a medium-sized pool of samples, derived from fossil benthic communities. It is the first attempt to use this method in palaeobiology known to us. It seems to be a promising addition to the traditional instruments of ordination.

Acknowledgements

The study was carried out under the support of the Grant Agency of the Czech Republic (GACR), Project 205 / 93 / 0723, and the Geological Institute of the Academy of Sciences of the Czech Republic Project 5213. It was a part of the UNESCO - IGCP Project 335. Thanks are due to Eric A. Oches (University of Minnesota, U.S.A.), Doug Erwin (Smithsonian Institution, U.S.A.), and James Ebert (University of Oneonta, N.Y., U.S.A.) for a critical review and language correction of the text.

References

CEJCHAN P. and HLADIL J. (in print). Searching for extinction / recovery gradients: the Frasnian - Famennian interval, Mokra Section, Moravia, Central Europe. In M.B. Hart (ed.): *Biotic Recovery from Mass Extinction Events*, 135-161. *Geological Society Special Publication*, 102. London.

DVORAK J., FRIAKOVA O., HLADIL J., KALVODA J. and KUKAL Z. 1987. Geology of the Palaeozoic rocks in the vicinity of the Mokra Cement Factory quarries (Moravian Karst). *Sbornik geologickych ved, Geologie*, 42, 41-88.

HLADIL J., KALVODA J., FRIAKOVA O., GALLE A. and KREJCI Z. 1989. Fauna from the limestones at the Frasnian / Famennian boundary at Mokra (Devonian, Moravia, Czechoslovakia). *Sbornik geologickych ved, Paleontologie*, 30, 61-84.