A NEURAL NETWORK FOR SPECTRAL ANALYSIS OF STRATIGRAPHIC RECORDS*

M. Brescia[1,2] B. D'Argenio[3,4] V. Ferreri[4] G. Longo[2]
N. Pelosi[3] S. Rampone[5,7] R. Tagliaferri[6,7]

Abstract

Astronomically controlled variations in the climate induce cyclic trends in the sedimentary process if being effectively registered by the depositional process, are easy to detect in shallow marine carbonate rocks. One of the main difficulties to be solved in order to choose among the registered periodicities is the conversion from the spatial (i.e. recurrent variations along the strata sequence) to the temporal domains of the astronomically induced frequencies present in the rock record. We discuss here how this problem can be circumvented by teaching a neural net to recognize armonics in the signal. The application to two sequences in the Cretaceous of Southern Italy has shown this approach to be particularly effective, and has confirmed the existence of Milankovic'-type periodicities in the examined records, where climate, sediments and biota concomitantly react to the variation of the solar constant induced by secular perturbations of the Earth's orbital elements.

1 Introduction

In this paper we illustrate a neural network methodology to the recognition of cyclicities in stratigraphic signals stressing the genetic connection of the orbital perturbations induced by the variability of Earth's thermal budget (Milankovic' cyclicity) with the climate, sedimentation and biosphere.

2 The data

The data used for the present analysis refer to two stratigraphic sequences observed along exposed open air quarries near Mount Raggeto ([5], [2] and Mount

*[1]Dipartimento di Informatica ed Applicazioni, Università di Salerno, Italy [2]Osservatorio Astronomico di Capodimonte, Napoli, Italy [3]Istituto di Ricerca GEOMARE sud, CNR, Napoli, Italy, [4]Dipartimento di Scienze della Terra, Università di Napoli "Federico II", Italy, [5]Dipartimento di Fisica Teorica, Università di Salerno, Italy, [6]Dipartimento di Informatica ed Applicazioni, Università di Salerno, Italy, [7]INFM, Unit of Salerno, Italy

Tobenna [4]. The geological characterization of these sequences has been discussed at length [5], [4] and we refer to these papers for further details. We want, however, to stress a few points:

- The data were obtained by sampling at centimetric scale textures directly from the rock outcrop and lithofacies with supplementary examination by means of thin sections;
- Both sequences refer to carbonate strata formed in shallow water environments; *id est* at a depth never exceeding -on the average- few meters or tens of meters and, therefore, in environments very sensitive to global sea level oscillations (eustatism);
- The sedimentation rates in these environments are usually fairly high and, therefore, a sampling at centimetric scale provides a high time resolution.

The exposed part of the Mount Raggeto sequence covers ca. 270 m, while the Mount Tobenna sequence extends ca 800 m of which 60 m have been studied in the detail required by the present study. The sequence related to Mount Raggeto is illustrated in Fig.1 as an example.

2.1 Application to stratigraphic records

In this paper we address the specific problem of how to train a neural net to recognize the existence of those periodic signals which have been known for a long time to exist in stratigraphic signals. The used neural net is a Multy Layer Perceptron with 2 hidden layers, with back propagation learning [6]. Several authors have suggested that these signals might be linked to the secular variations of the Earth's orbit. The main problem to be solved remains, however, how to convert the spatial frequencies observed in the signal to the time frequencies predicted by the orbital theory. Starting from the quite accurate time frequencies predicted for a given epoch, by Berger and his collaborators on the basis of N-body simulations of the solar system, we trained the neural network described above to recognize the existence of such periodic signals (regardless their amplitude) in real data, once a rough estimate of the sedimentation rate S_{est} has been provided.

This was achieved by producing a set of synthetic data string based on S_{est}, and the set of frequencies estimated by Berger as the most likely for the assumed age of the considered strata.

These strings were then used to train the net. In the following section we shall give more details on the application of the method to two specific cases.

3 Data analysis

Data analysis consisted of two main steps: (i) preprocessing aimed to reduce the noise level in the data and (ii) spectral analysis needed to have a first guess on the spatial armonics present in the signal. The first goal was achieved by means of a simple running mean algorithm. The use of more refined filters would have been unappropriate due to the tipology of the signal. The second step makes use, instead, of the so called Modified Scargle Algorithm [3], [2] and leads to a rough estimate of S_{est}.

4 Description of the experiments: Mount Tobenna and Mount Raggeto Data

The stratigraphic records obtained for the exposed sequences in Mount Tobenna and Mount Raggeto were processed according to the previous section.

This is the most delicate step of any experiment done with neural networks and on it depend the choices of the training and test sets and the best structure for the input data. The aim of our experiment was to teach the network how to recognize the possible existence in a very noisy spatial record of periodic signals, choosing between six possible classes, namely the six main periodicities expected for Milankovic' type phenomena. In order to train the network we produced a series of simulated records having the following characteristics:

- all possible combinations of the six Milankovic' periodicities;
- additive noise;
- same length of the real stratigraphic sequence;
- similar square-like shape;
- same sampling rate of the real data set; due to the different domains of the real (*id est* space) and simulated (*id est* time) series it is first necessary to find the conversion factor (*id est* sedimentation rate) α.

In order to estimate α we proceeded as follows: by a preliminary spectral analysis of the data we derive the most significant peak and assumed it to be related to the highest Milankovic' frequency, thus obtaining an estimate of the sedimentation rate α. This value gave the sampling rate for the synthetic signals, which were therefore degraded by adding randomly generated additive noise. In Fig.2 is shown an example of synthetic signal, containing all six Milankovic' frequencies, created by our procedure.

We subdivided every simulated data set into smaller sub-sequences long enough to contain the whole information. Every sub-sequence is slightly overlapping in order to simulate a sort of continuity in submitting the signals to the net. For each sub-sequence, a Fourier transform allows to identify the six amplitude values which represent a vector. A final pattern input is obtained by the juxtaposition of 2 adjacent vectors, related to 2 successive subsequences of the signal, to obtain a 12-item record. So every signal is submitted to the network likewise a pattern set. The number of input patterns for each signal depends on the number of subsequences obtained by the signal itself. The training set is composed of 30 synthetic signals opportunely choosen between the 64 available. Therefore an entire learning cycle is performed by submitting all the 30 synthetic signals in cascade. For the test phase we apply the same treatment to the real signal currently examined. For the test phase we use, in addition to the real stratigraphic signal, some synthetic signals not included in the original training set.

Tables 1, 2 and 3 give the most significant results of the esperiments.

4.1 Comparison with classical methods

In terms of comparison between our method and those actually used by geologists, the new approach has to be considered a more refined alternative way to analyse geological data. The so called "ratio method" [3,4] makes a series of comparisons between 2 tables, respectively containing ratios between all the spatial periodicities, obtained by Scargle's spectral analysis, and all the Milankovic' periodicities, calculated by Berger, in order to find the eventual correlations between the 2 sets of data and finally the correlation coeffient (i.e. the sedimentation rate of the geological succession). The last approach is characterized by an expensive amount of calculations, and it requires to take into account the whole Scargle's periodicities set, while our method has been developed in order to reduce human work in terms of calculations and decisions, and to introduce the possibility to realize a completely automatic system to analyse, directly on site, with a high releability degree, the geological succession.

5 Conclusions

We have presented a neural net based approach to the recognition of periodicities in stratigraphic record. With respect to the traditional power spectrum based techniques (PST), this method offers several methodological and practical advantages which render the detection of periodicities much more reliable. Infact, instead of searching blindily for peaks in the periodogram, the neural nets approach presented here makes use of PST, only to obtain a first-order estimate of the sedimentation rate. The existence of periodic components in the stratigraphic record is then performed on yes/not base by training the net to recognize signals of given frequency. The time and effort required by the training of the net is largely compensated by the effectiveness of the method and by the relatively short computing time required by the following processing steps. The method has been tested on two stratigraphic records obtained with centimetric accuracy along two exposed sections in the open air sequences of Mount Raggeto and Mount Tobenna, which had already been studied by the authors with traditional PST's [11], [12]. The results confirmed the existence of Milankovic type signals in both sequences and in particular:

- Eccentricity (long and short), Obliquity (long and short) and Precession (long) for the Mount Tobenna;
- Eccentricity (long and short), Obliquity (long and short) for the Mount Raggeto.

Acknowledgements

This work has been partly sponsored by CNR (Italian National Research Council) to Geomare Sud and Dept. of Earth Sciences Univ. of Naples; 40% MURST through the Salerno Unit of the INFM; ASI (Italian Space Agency) and, finally, by a 5 years contract between the CNR and the IIASS E. R. Caianiello.

References

(1) B. D'Argenio, V. Ferreri, A. Raspini, A cm-scale study of shallow water Cretaceous deposits formed under high frequency eustatic regime. Monti di Sarno (Southern Italy) A sedimentologic approach to microstratigraphy, Bollettino Società Geologica Italiana, 111, 399-407, 1992.

(2) G. Longo, B. D'Argenio, V. Ferreri, M. Iorio, Fourier evidence for high frequency astronomical cycles recorded in Lower Cretaceous carbonate platform strata. Monte Maggiore Southern Apennines, Italy, in: IAS Spec. Publ. on Orbital forcing and cyclic sequences (eds. De Boer P.L. & Smith D.G.), Blackwell, Oxford, 19, 77-85, 1994.

(3) N. Pelosi, A. Raspini, Analisi spettrale della ciclicità di alta frequenza in successioni carbonatiche neritiche di limitato spessore. Il caso dei Monti di Sarno (Campania), Giornale di Geologia, 55, 37-49, 1993.

(4) V. Ferreri, F. Ardillo, F. Buonocunto, N. Pelosi, Il Cretacico superiore del Monte Tobenna (Salerno). Studio sulla ciclicità di alta frequenza in depositi carbonatici di piattaforma. Rendiconto Accademia Scienze Fisiche e Matematiche in Napoli, Serie IV - Vol. LX - Anno CXXXII, 165-193, 1993.

(5) B. D'Argenio, V. Ferreri, F. Ardillo, F. Buonocunto, Microstratigrafia e stratigrafia sequenziale. Studi sui depositi di piattaforma carbonatica, Cretacico del Monte Maggiore (Appennino Meridionale), Bollettino Società Geologica Italiana, 112, 739-749, 1993.

(6) D.E. Rumelhart, J.L. McClelland, Parallel Distributed Processing, MIT Press, Cambridge, 1986.

Table 1: Training phase. Network topology contains the information on the number of layers and the number of neurons per each layer, respectively: "in-hid1-hid2-out".

Synthetic signal	network topology	cycles	rms train
Tobenna	12-6-6-6	77.000	0.013117
Raggeto	12-6-6-6	95.000	0.009991

Table 2: Test phase. The names F_i, $i = 1, ...6$ are the six MIlankovic' periodicities, from the highest (F1=400 ky) to the lowest (F6=18 kyr).

Signal	RMS -Test	Target periodicities	recognized periodicities
Tobenna	0.124958	F1,F2,F3,F4,F5	F1,F2,F3,F4,F5
Raggeto	0.031333	F1,F2,F3,F4	F1,F2,F3,F4

Table 3: Test phase. Level of confidence for positive detections for the six Milankovic' frequencies. A value higher than 50.0 meaning that the frequency has been detected.

Signal	Freq. 1	Freq. 2	Freq. 3	Freq. 4	Freq. 5	Freq. 6
Tobenna	99.48	97.77	87.69	63.51	57.44	0.38
Raggeto	82.17	85.96	55.18	57.84	41.22	16.53

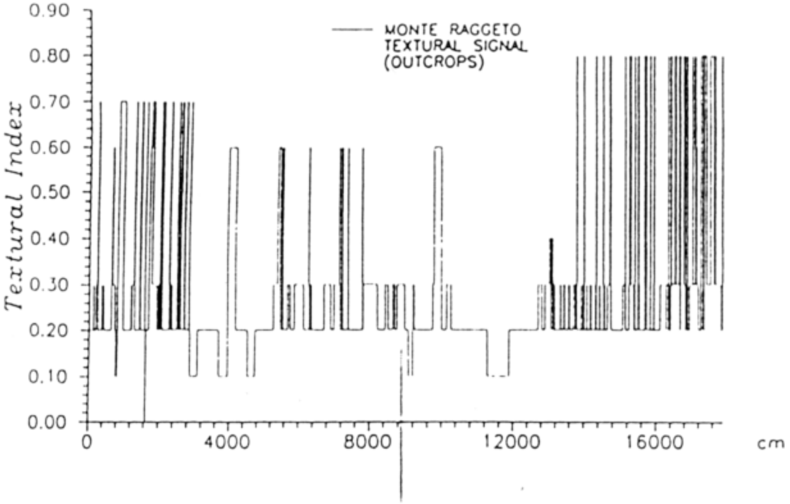

Figure 1 Textural signal extracted from mount Raggeto.

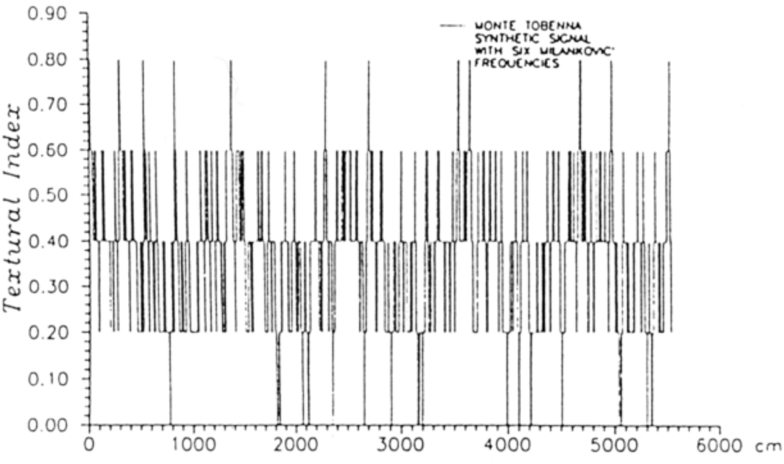

Figure 2 Synthetic signal obtained for the mount Tobenna texture record by combining the six Milankovic' cycles.